Does trait-based joint species distribution modelling reveal the signature of competition in stream macroinvertebrate communities?

Merja Elo1,2 | Jenny Jyrkänkallio-Mikkola3 | Otso Ovaskainen1,4,5 | Janne Soininen3 | Kimmo T. Tolonen6 | Jani Heino6

1Department of Biological and Environmental Sciences, University of Jyväskylä, Jyväskylä, Finland; 2School of Resource Wisdom, University of Jyväskylä, Jyväskylä, Finland; 3Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland; 4Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland; 5Department of Biology, Centre for Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, Norway and 6Finnish Environment Institute, Freshwater Centre, Oulu, Finland

Abstract

1. The occupancy and abundance of species are jointly driven by local factors, such as environmental characteristics and biotic interactions, and regional-scale factors, such as dispersal and climate. Recently, it has been shown that biotic interactions shape species occupancies and abundances beyond local extents. However, for small ectothermic animals, particularly for those occurring in freshwater environments, the importance of biotic interactions remains understudied. Species-to-species associations from joint species distribution models (i.e. species associations while controlling for environmental characteristics) are increasingly used to draw hypotheses of which species possibly show biotic interactions.

2. We studied whether species-to-species associations from joint species distribution models show signs of competition using a hypothesis testing framework in stream macroinvertebrate communities at regional extent.

3. We sampled aquatic macroinvertebrates from 105 stream sites in western Finland encompassing a latitudinal gradient of c. 500 kilometres. We hypothesized that if competition drives these associations (H1) functionally, similar species are mostly negatively associated, whereas functionally dissimilar species show random associations. We further hypothesized that the relationship between functional dissimilarity and the strength of association is more pronounced (H2) for abundances rather than occupancies, (H3) at small grain (i.e. stream site) rather than at large grain (i.e. river basin), and (H4) among species having weak dispersal ability than among species with high dispersal ability.

4. Stream macroinvertebrates showed both negative and positive species-to-species associations while controlling for habitat characteristics. However, the negative associations were mostly at large grain (river basin) rather than at small grain

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Journal of Animal Ecology published by John Wiley & Sons Ltd on behalf of British Ecological Society
1 | INTRODUCTION

Understanding the factors determining species occupancies and abundances’ spatial arrangements has been a longstanding core topic of ecology and biogeography (Mittelbach, 2012; Wisz et al., 2013). These patterns are crucial not only to conceptual issues on the origin, maintenance and distribution of biodiversity but also to more applied research such as conservation biology. For instance, to predict species range shifts caused by ongoing global change, it is essential to identify the determinants of species occupancy and abundance. At broad spatial extents (i.e. landscapes, regions, continents and beyond), species occupancy and abundance patterns have been attributed mainly to climate and other abiotic factors, whereas the effects of biotic interactions have often been neglected (Pearson & Dawson, 2003; Wisz et al., 2013). Recently, this view has been challenged with increasing evidence showing how biotic interactions both within and across trophic levels contribute to these patterns beyond local extents (Araujo & Luoto, 2007; Blois et al., 2013; Gotelli et al., 2010; Kissling & Schleuning, 2014; Wisz et al., 2013). While it is certainly true that abiotic factors and dispersal affect the occupancy and abundance of species, an additional interesting question is how biotic interactions influence these patterns.

Determining the effect of biotic interactions through statistical methods from observational data began with studies of species co-occurrence patterns on islands (Connor & Simberloff, 1983; Diamond, 1975). The key idea was that competition should lead to the lack of co-occurrence of the competing species (Diamond, 1975) compared to co-occurrences predicted by a null model (Connor et al., 2013; Connor & Simberloff, 1983). Recently, the growing interest in biotic interactions has led to a myriad of different statistical methods to study whether species are found together less often (negative association in occupancy; segregation) or more often (positive association in occupancy; aggregation) than predicted by random chance only (see Dormann et al., 2018 for an overview). These techniques include single species distribution and dynamic occupancy models in which other species’ occupancy (Heikkinen et al., 2007), phylogenetic relationships (Morales-Castilla et al., 2017) and/or ecological similarity (Beaudrot et al., 2019) are included as covariates. Other examples include partial correlation networks (Garcia-Girón et al., 2020), and different ordination methods, such as distance-based redundancy analysis (Bottin et al., 2016). A technique that is becoming increasingly popular is joint species distribution models (JSDMs; Ovaskainen & Abrego, 2020; Ovaskainen et al., 2017; Pollock et al., 2014; Thorson et al., 2015). JSDMs enable the simultaneous modelling of multiple species’ responses to environmental variables, and the residuals of the model indicate whether species-to-species associations are positive, negative or random, while controlling for environmental variables (Warton et al., 2015). Considering the effect of environmental variables is crucial since positive associations might be due to preferences for similar habitat characteristics (reflecting species’ convergence), especially among close relatives. Likewise, negative associations may be due to differences in habitat preference (reflecting species’ divergence) rather than competition.

However, meaningful interpretation of non-random associations is often complicated, and mere non-random associations, even after controlling for environmental factors, may not allow one to separate between abiotic and biotic factors reliably (Dormann et al., 2018). For example, failure to include biologically important environmental variables in a JSDM can also produce non-random associations (Dormann et al., 2018). Hence, species-to-species associations from JSDMs merely represent hypotheses suggesting which species potentially show biotic interactions (Ovaskainen et al., 2017). Experiments are by far the most reliable way for testing whether species-to-species associations indeed result from biotic interactions (Fayle et al., 2015; Siepielski et al., 2010) but also hypothesis testing framework is a useful method to explore these associations (D’Amen et al., 2018; Kohli et al., 2018). Within hypothesis testing frameworks, the effects of biotic interactions and habitat preferences can be separated.
for instance, by linking observed species-to-species associations to species ecological and functional traits (Dormann et al., 2018; Kohli et al., 2018). If ecologically and functionally similar species with similar habitat preferences are negatively associated, this can be interpreted as a result of interspecific competition (Kohli et al., 2018; Mönkkönen et al., 2017). Indeed, the notion that ecologically similar species should compete more intensively than ecologically dissimilar species is not new (Darwin, 1859), and ecological similarity has been used as a proxy for the probability that species compete with one another (Beaudrot et al., 2019). Functional and phylogenetic information can be included in the modelling framework of JSDMs (Ovaskainen et al., 2017) to study if species that share functional traits respond in a similar fashion to specific environmental factors (Pollock et al., 2012). However, the dependence of species-to-species associations on species functional dissimilarity cannot (yet) be incorporated into the JSDM framework.

We present an approach that combines JSDMs with a hypothesis testing framework. We build a priori hypotheses of the relationship between species functional dissimilarity and species-to-species associations when biotic interactions affect communities. We then test these hypotheses with species-to-species associations derived from JSDMs, that is, associations independent of the in situ measured relevant habitat characteristics. A hypothesis testing framework has been used previously with a context of separating possible biotic interactions from other factors on observational data as a form of dichotomous logic-trees (D’Amen et al., 2018; Kohli et al., 2018). Each individual species pair has been tested separately whether the hypothesis is rejected or accepted. By contrast, we evaluate the association of each species pair in relation to all other species pairs (Mönkkönen et al., 2017) to obtain a robust perspective. If negative associations of functionally similar species occur with the same probability as negative associations of functionally dissimilar species, controlling for environmental variables, this raises doubts whether competition drives segregation. In addition, we extend our approach by adding hypotheses predicting at which spatial grains, within which species, and whether using occupancy or abundance species-to-species associations and functional dissimilarity should show a relationship. Hence, we get a deeper sight whether biotic interactions, and competition in particular, drive species-to-species associations derived from JSDMs.

Specifically, we examined species-to-species associations of stream macroinvertebrates at a regional extent. The studies of biotic interactions at regional extents have focused on the endothermic terrestrial fauna or flora (e.g. Wisz et al., 2013) while small ectothermic animals, particularly those living in freshwater environments, have received less attention. In general, it has been suggested that small ectothermic animals show more random occupancy patterns in contrast to endothermic animals (Gotelli & McCabe, 2002). However, the findings may stem from differences in species richness among datasets (Ulrich & Gotelli, 2010, 2013), and the results obtained thus far for insect communities range from a low degree of competition (D’Amen et al., 2015) to severe competition (Fayle et al., 2015). Distributions of stream macroinvertebrates have been found to show spatial segregation (Heino, 2009), although the effect might be dependent on the environmental context (McCreadie & Bedwell, 2013), accompanied by spatial scaling issues (Heino & Grönroos, 2013). At local extents (i.e. multiple sites within one stream), segregation may be interfered by high dispersal rates, the effects of which may be less important at regional extents (Heino & Grönroos, 2013).

We hypothesized that (H1) if competition drives the species-to-species associations, functionally similar species will show negative associations, whereas functionally dissimilar species show random associations (Figure 1). In addition, we hypothesized that the relationship between functional dissimilarity and the strength of the association is more pronounced (H2) for abundances rather than occupancies. Studies of potential biotic interactions at regional extents typically use presence–absence data, which is also true for aquatic macroinvertebrate studies (Heino, 2009; McCreadie & Bedwell, 2013; but see García-Girón et al., 2020). However, the effect of biotic interactions must be very strong to lead to total exclusion of other species from a site or a species inability to occupy a site without the other species (D’Amen et al., 2015), especially in aquatic environments where animals move easily with currents (Tonkin et al., 2018). By contrast, changes in abundances might reveal more subtle patterns remaining undetected in occupancy data, due to, for example, mass effects (Leibold et al., 2004). These effects may be strong especially at local extents (e.g. multiple sites within one stream; Heino & Grönroos, 2013). Abundance data may reveal that a species may be less abundant in the presence of another species at high abundance, and vice versa. These associations in abundances could arise from biotic interactions that take place in macroinvertebrates, at least at grain sizes typically less than 1 m² (Holomuzki et al., 2010). Hence, while occupancy data are more often available and more robust than abundance data, abundance may better help to reveal possible biotic interactions.

We hypothesized that the relationship between functional dissimilarity and species-to-species association is more pronounced (H3) at small spatial grains rather than at large spatial grains, and (H4) among species having weak dispersal ability than among species with high dispersal ability (Figure 1). Species with good dispersal ability may be especially prone to mass effects which would lead to more spurious patterns than species with weak dispersal ability. In comparison to competition, positive biotic interactions among macroinvertebrates have been rarely studied in freshwater settings (Holomuzki et al., 2010; Silknetter et al., 2020), possibly reflecting the long-term tendency in ecology to emphasize more negative interactions than facilitation (Bruno et al., 2003). Positive associations can occur among functionally very different organisms (Silknetter et al., 2020), and they may be most likely due to indirect facilitation by ameliorating environmental conditions (Holomuzki et al., 2010). Thus, we do not have specific predictions which species, if any, show positive associations in occupancy or abundance.
(H1) Functionally similar species are mostly negatively associated, whereas functionally dissimilar species show random associations.
(mS/m) in the field using YSI device model 556 MPS (YSI Inc.), and took additional water samples during the field sampling. Water samples were subsequently analysed for (3) total nitrogen (TN; µg/L), (4) total phosphorus (TP; µg/L) and (5) water colour (Pt mg/L) following Finnish national standards (SFS-EN ISO 1568–1 (2004) for TP, EN ISO 11905–1 (1998) for TN and SFS-EN ISO 7887 (2012) for water colour). The wide geographical variation of the sampling sites and the aim of covering as much environmental conditions as possible resulted in a large variation of the measured habitat characteristics. The mean values and ranges of the environmental values are shown in Supporting Information 2.

To sample benthic macroinvertebrates, we took a 2-min kick-net sample (net mesh size: 0.5 mm) at each of the 105 stream sites surveyed. The sample consisted of four 30-s samples that were obtained to cover the main microhabitats (i.e. variations in current velocity, depth, particle size and macrophyte cover) at a riffle site of c. 25–50 m². The four samples were pooled, preserved in ethanol in the field and taken to the laboratory for further processing. In the laboratory, all invertebrates were separated from debris and they were identified mostly to species level (76% of taxa). However, young instars and individuals lacking species-level taxonomic keys were identified to genus level. Chironomidae, Ceratopogonidae, Simuliidae, Hydracarina, Oligochaeta and Turbellaria were identified to a coarser taxonomic level, and were thus excluded from the analyses.

We selected the macroinvertebrate species occupying at least 10 sites and for which we found at least 100 individuals, resulting in a set of 40 species for the statistical analyses. We classified the species according to their primary substrate attachment mode (i.e. burrower, crawler, semi-sessile, swimmer) and functional feeding guild (i.e. gatherer, filterer, predator, scraper, shredder). We calculated mean potential maximum size (dry weight mg) of the aquatic stage of species using the length–weight relationships obtained from the literature for body size measure (e.g. Benke et al. 1999). We estimated dispersal mode primarily following Tachet et al. (2010) who classified species to aerial active, aquatic active and aquatic passive dispersers. The traits and literature references of the trait classifications for all the study species are presented in Supporting Information 2.

2.2 | Statistical analyses

2.2.1 | Joint species distribution models

For estimating species co-occurrences, we applied Bayesian joint species distribution models, namely Hierarchical Modelling of Species Communities (HMSC; Ovaskainen & Abrego, 2020; Ovaskainen et al., 2017). HMSC is a flexible framework in which both species occupancies and abundances can be modelled in response to different environmental factors. Residual variation (i.e. variation not explained by the fixed factors) in species occurrences is captured by latent variables. These latent variables can be estimated at any level of the hierarchical sampling scheme. Species-to-species residual association matrices, (Ω) unique for each level of the hierarchical sampling scheme, are then estimated as variance–covariance matrices of the loadings of the latent variables. Furthermore, species-to-species association networks can be translated to the level of the correlation matrix R, defined as $R_{ij} = \Omega_{ij}/\sqrt{\Omega_{ii} \Omega_{jj}}$, where $\Omega_{ij}$ describes the amount of covariation among the species $j_1$ and $j_2$. The element $R_{ij}$ measures to what extent species $j_1$ and $j_2$ are found together more or less often than expected by chance in the scale from −1 to +1.

We modelled the occupancies and abundances of the 40 species in the 105 sites with a hurdle model consisting of two parts: we modelled the occupancy of the species by a probit model, and conditionally on the presence, we modelled the abundance (log-transformed count, normalized to zero mean and unit variance within each species) with a normal model. In both parts of the hurdle model, we included random effects to reflect the hierarchical sampling scheme as stream sites (105) nested within a river basin (21; five stream sites in each). The hierarchical level (i.e. grain for now on ‘river basin’) was modelled as a spatially explicit random effect, using the mean coordinates of the five stream sites as the location. Thus, species-to-species associations are modelled at two grains, river basins and stream sites, while the extent remains constant.

We had altogether 12 potentially important candidate variables measured directly at the stream sites: mean current velocity (m/s); mean depth (cm); mean cover of mosses (%); mean grain size (mm); mean stream width (m); shading (%); cover of riparian deciduous trees (%); pH; conductivity (mS/m); total nitrogen (TN; µg/L); total phosphorus (TP; µg/L) and water colour. As TN and TP were correlated ($r = 0.631$, $p < 0.001$), we used only TN, thus resulting in 11 environmental variables.

We ran the models with all the 11 environmental variables using R-package HMSC 3.0 (Tikhonov et al., 2020). The package uses the Bayesian framework with Gibbs Markov chain Monte Carlo (MCMC) sampling. We assumed the default prior distributions, and sampled the posterior distribution with two chains, both for 1,000 samples (1,000,000 iterations with thinning of 1,000), using transient phase of 500,000, and adaptation (the number of MCMC steps at which the adaptation of the number of latent factors is conducted) of 400,000. We evaluated the chain mixing by assessing the effective size of the posterior sample as well as with a potential scale reduction factor. To ensure that the full model did not contain unimportant variables leading to overfitting, we ran also a reduced model containing only the six most important variables based on preliminary variance partition of the full model. These variables were cover of riparian deciduous trees, water colour, pH, cover of mosses, TN and conductivity for occupancy model, and current velocity, depth, grain size, stream width, TN and conductivity for abundance model. We compared the preliminary explanatory and the predictive power based on twofold cross-validation of the full and the reduced model (calculated as Tjur’s $R^2$ and $R^2$ for occupancy and abundance models, respectively). Since the reduced model did not lead to improved predictive ability (Supporting Information 4), we continued our analyses with the full model.
2.2.2 | Effect of species traits

We divided aquatic macroinvertebrate species according to their primary substrate attachment mode, functional feeding guild and body size which are key traits of aquatic macroinvertebrates (Tachet et al., 2010). The selected 40 species covered a range of biological traits (Supporting Information 3). Most of the species were classified as crawlers in their substrate attachment mode (29 from the total of 40 species), scrapers as their functional feeding guild (14 species) and aerial active as dispersal type (21 species).

With species substrate attachment mode, functional feeding guild and log_{10}-transformed body size as traits, we calculated pairwise functional dissimilarity using Gower distance (function ‘daisy’ of the R package ‘cluster’; Maechler et al., 2016). If competition drives the associations, we expect negative associations between the most similar species, whereas dissimilar species would show random associations (H1). This would lead to a positive relationship between species-to-species associations and species dissimilarity. Moreover, we expect that this positive relationship is more pronounced between species having weak dispersal ability than between species having high dispersal ability (H4) because species with high dispersal ability may be especially prone to mass effects leading to more spurious patterns in comparison to species with weak dispersal ability. Thus, we divided species pairs to three groups differing in their dispersal ability: (a) both partners in the pair are aerial dispersers (highest dispersal ability), (b) both are aquatic dispersers (lowest dispersal ability) and (c) one species is an aerial disperser and the other is an aquatic disperser (medium dispersal ability). We applied a simple Spearman rank correlation between species-to-species association and species dissimilarity, for the three dispersal groups separately. This was repeated for each of the four species-to-species association measures (model: occupancy/abundance; grain: river basin/stream site). Because of the interdependence of the observations, assessing the level of statistical support through, for example, the p values associated with the linear regressions would not be adequate. We thus generated a null distribution by randomizing the order of species when calculating dissimilarity and re-calculated the correlations with these randomized data. We repeated the procedure 1,000 times and calculated 2.5% and 97.5% percentiles of the resulting parameter values. In addition, we repeated the procedure without predatory species, as for predator-prey relationships, the pattern of increased segregation with increased similarity was not expected.

All analyses were carried out with R version 3.5.1 (R Development Core Team, 2017). We provide the codes for HMSC and the following dissimilarity analyses in Supporting Information 5.

3 | RESULTS

The mixing of MCMC chains was good for the abundance model as the potential scale reduction factor was smaller than 1.1, and the effective sample sizes were close to 2,000 for the studied parameters (Supporting Information 5). For the occupancy model, the mixing was adequate for omega-parameters, although for some pairs in stream sites, the convergence was not ideal (Supporting Information 5). For beta-parameters (species-specific regression parameters describing the response to the environmental data that are not considered in the current study), the effective sample sizes were smaller than 2,000. The potential scale reduction factors were sometimes much larger than 1.1 (Supporting Information 5). This is not an unusual phenomenon with non-normally distributed, large and multidimensional data (Tikhonov et al., 2020). For occupancy, the mean explanatory power (measured as Tjur’s R^2) of the model was 0.340 and the mean predictive power was 0.146. For abundance, the mean explanatory power (measured as R^2) of the model was 0.457 and the mean predictive power was 0.085. There was much variation among species in the explanatory and predictive powers of both parts of the hurdle model (Figure 2).

We found at least moderate (posterior probability at least 80%) species-to-species associations in occupancy for a large number of species pairs at both grains (Figure 3a,b). At the river basin grain, we found both positive and negative associations (Figure 3a), whereas at the stream site grain, the majority of them were positive (Figure 3b). The negative associations at the stream site grain were related mainly to two species (Asellus aquaticus, Plectrocnemia conspersa; Figure 3b).

In abundance model, there were virtually no strong associations at the river basin grain (Figure 3c), and at the stream site grain, the few positive associations were confined to pairs consisted of only

![FIGURE 2 Explanatory and predictive power of the model for species occupancy (a) and abundance, conditionally on presence (b)](image-url)
sis (H2) stating that species-species associations should be more pronounced for abundance than occupancy. Thus, without even taking species functional similarity account, we can already discard our hypothesis (H2) stating that species-species associations should be more pronounced for abundance than occupancy.

Disagreeing with our hypothesis (H1), there was no positive relationship between species functional dissimilarity and species-to-species association’s strength and sign (Figure 4a,b). This also disagrees with the hypotheses stating that the relationship should
be more pronounced at the stream site grain (H3) and for species with low dispersal ability (H4). However, at the stream site grain, the relationship between association and species functional dissimilarity depended on the species dispersal, as predicted. For aquatic dispersers (i.e. with low dispersal ability), the most similar species were positively associated and the most dissimilar species were randomly associated while for species pairs with aerial dispersal (i.e. high dispersal ability) or mixed dispersal there was no relationship (Figure 4b; Table 1). Exclusion of the predatory species did not qualitatively affect the results (Supporting Information 6).

Since all but five of the associations in abundance were weakly supported (Figure 3c,d), we did not execute further analyses for them.

## DISCUSSION

Stream macroinvertebrates showed both negative and positive species-to-species associations while controlling for in situ measured habitat characteristics. However, the negative associations were mostly at large grain (river basin) than at small grain (stream site), in occupancy rather than abundance, and not related to species functional dissimilarity. Thus, we rejected all of our hypotheses considering possible competition, and we can conclude that competition is not a major driving force behind the negative associations at the spatial scales studied. However, the analyses revealed interesting patterns showing grain dependency of positive associations and their relationships with species functional dissimilarity on specific dispersal modes. This raises the question of the key mechanism behind positive associations.

Our analyses showed that a significant proportion of the species-to-species associations in occupancy were non-random when accounting for habitat characteristics, thus disagreeing with the findings of random co-occurrences of stream macroinvertebrates at regional extents (Heino & Grönroos, 2013; McCreadie & Bedwell, 2013). Also, the sign and strength of the species-to-species associations were dependent on the spatial grain: at the larger river basin grain, there were both negative and positive associations, whereas at the smaller stream site grain, the associations were more frequent and predominantly positive. This is in line with a study of mammals showing the increase in prevalence of positive co-occurrences at local versus landscape scales (Kohli et al., 2018).

There were both positive and negative associations in occupancy at the river basin grain, but the associations were not related to

---

**TABLE 1** Spearman’s rank correlation coefficient for species dissimilarity and species-to-species associations in occupancy at different two different grains (river basin, stream site). The coefficient is calculated separately for each of the species’ dispersal ability groups (both species aquatic dispersers, n = 171; the other species aquatic and the other aerial disperser, i.e. mixed dispersers, n = 399; both aerial dispersers, n = 210). The observed correlation coefficient value as well as the 2.5% and 97.5% percentiles from randomized data are shown. The observed parameter values that fall outside of the 2.5% and 97.5% percentiles obtained from the randomized data are shown in bold.

| Grain     | Dispersal group | Observed | Quantiles       |
|-----------|-----------------|----------|-----------------|
|           |                 |          | 0.025 0.975      |
| River basin| Aquatic        | −0.022   | −0.219 0.127    |
|           | Mixed          | −0.001   | −0.119 0.106    |
|           | Aerial         | 0.005    | −0.212 0.100    |
| Stream site| Aquatic       | −0.458   | −0.357 0.237    |
|           | Mixed          | 0.086    | −0.202 0.185    |
|           | Aerial         | −0.029   | −0.239 0.194    |
species functional dissimilarity. As the distributions of the majority of the species cover the whole of western Finland, and a large part of Europe (www.freshwaterecology.info), it is not likely that these associations are due to dispersal limitation. The drainage basins showed a wide variety of agricultural-dominated catchments to almost entirely forested areas. Adding catchment-scale variables may improve the models based on only field-measured variables in stream settings (Johnston et al., 2017). Thus, we cannot rule out the effects of unmeasured environmental variables, such as catchment land-use characteristics and soil type, at the river basin grain. Including these variables could have explained some of the associations between species pairs. However, we measured a set of key local environmental variables for stream macroinvertebrates (Grönroos & Heino, 2012; Malmqvist & Mäki, 1994), and local environmental conditions and catchment features are typically correlated (e.g. agricultural land use is associated with water nutrient levels; Soininen et al., 2015). In addition, Perez Rocha et al. (2018) used partly the same data as described here, and showed that local environmental variables and spatial structuring, rather than catchment and climate variables, were superior in explaining beta diversity of aquatic macroinvertebrates. Hence, we suggest that the mechanism behind these associations at the river basin grain might be related to the species’ metapopulation dynamics, in which spatial dynamics might result in clumped distributions of species detected as positive and negative associations. These spatial dynamics may also be related to temporal variations of the habitat characteristics (e.g. chemistry and flow) which go unnoticed with the snap-shot measures done using field surveys.

At the stream site grain, positive associations in occupancy constituted the vast majority of the non-random associations. The reason for these positive associations is likely to be related to microhabitat variation. Although our modelling procedure included 11 environmental variables, which have previously shown to be important for stream macroinvertebrates in boreal regions (Grönroos & Heino, 2012; Malmqvist & Mäki, 1994), there are important environmental characteristics which tend to vary at very fine scales (e.g. algal biomass, fine organic material, and leaves of different aquatic macrophytes and shoreline trees). Unfortunately, it is not feasible to measure these characteristics in such a large-scale study as presented here due to logistic and time constraints. Support for the similar small-scale habitat preferences comes from the result showing that for species pairs with low dispersal ability, the functionally similar ones had the most positive associations in occupancy. This pattern remained undetected in species with higher dispersal ability, possibly because of the strong mass effects. Interestingly, some species pairs had negative associations at large (river basin) grain and positive association at small (stream site) grain. Although this may seem counterintuitive, it stems from those species that occupy different drainage basins more often than predicted by chance, given the habitat characteristics. Still, when they do co-occur in a drainage basin, they also co-occur in a stream site.

The negative associations in occupancy at the stream site grain occurred only for two species, Asellus aquaticus and Plectrocnemia conspersa, which tended to have negative associations with almost all other species. The mechanisms behind these negative associations for the two species are likely to differ. A. aquaticus is able to tolerate various environmental conditions, and it is also often abundant in streams with relatively poor water quality (Maltby, 1991), where most species are not likely to occur. P. conspersa, in turn, is an effective predator, particularly in small stream sites (Edington & Hildrew, 1995), and may thus suppress other species. The general lack of negative associations at the small grain may be because of two reasons. First, the selected species are relatively generalist and able to shift their resource use (Mihuc, 1997), and thus they may not actually compete for resources. Second, the actual grain size was relatively large (c. 50 m², on average) compared to the size of the studied animals, which allows for spatial segregation of the species within a study grain.

Contrary to our predictions, species-to-species associations in abundance were mostly weak, irrespective of the spatial grain. This may be related to mass effects where individuals disperse from a good site (i.e. source) to other sites (i.e. sinks), regardless of environmental characteristics at sink sites (Leibold et al., 2004). A good (source) site is a location where a species can maintain a viable population, which is opposite to an unsuitable (sink) site where a species goes extinct without a flux of immigrants. The characteristics of a good (source) site depend on each species’ ecological preferences. In addition, weak species-to-species associations in abundance are at least partly related to the modelling approach: the abundance was modelled conditionally on presence and thus absences were considered as missing data. This obviously leads to the decrease of data in comparison to the occupancy model, and is also shown as a rather weak predictive power of the abundance model. Another limitation of our approach is the number of species as we were only able to include only 40 of the most abundant species, thus excluding other rare species. Moreover, it may also be that biotic interactions occur among species groups (García-Girón et al., 2020; Schwirwth et al., 2016) and within the same genus, rather than among individual species.}

To conclude, our results corroborate the potentiality of a trait-based approach when interpreting species-to-species associations based on joint species distribution models (Dormann et al., 2018; Kohli et al., 2018). Although we found many strong associations based on occupancy information, they were not generally linked to species traits in a way that would support biotic interactions, at least competition. Thus, species traits were used to interpret if species-to-species associations were the result of biotic interactions. The use of two spatial grain sizes in this study enabled the formulation of hypotheses to separate habitat characteristics and biotic interactions as a cause for non-random species-to-species associations. We see value in conducting further research to disentangle whether biotic interactions and abiotic environmental factors jointly affect community assembly even across large spatial extents. Our results add to the previous studies showing that distributions and abundances of stream macroinvertebrates are governed by abiotic environmental conditions (Heino & Mendoza, 2016; de Mendoza et al., 2018). They emphasize using of
novel trait-based methods to investigate the factors affecting species occupancy and abundance at different spatial scales.

ACKNOWLEDGEMENTS
The authors declare no conflict of interest. We thank K. Eyvindson for correcting our language and Kone Foundation (for M.E.) for funding the study. Financial support was also provided by grants no. 273557 (to J.H.), no. 273560 (to J.S.) and no. 309581 (to O.O.) from the Academy of Finland.

AUTHORS’ CONTRIBUTIONS
M.E. and J.H. conceived the idea of the study; J.H., J.J.-M., J.S. and K.T.T. designed the field work and collected the data; M.E. and O.O. analysed the data; M.E. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA AVAILABILITY STATEMENT
Data available from the Dryad Digital Repository https://doi.org/10.5061/dryad.k98sf7m5h (Elo et al., 2021).

ORCID
Merja Elo https://orcid.org/0000-0003-4045-5002
Otso Ovaskainen https://orcid.org/0000-0001-9750-4421
Janne Soininen https://orcid.org/0000-0002-5853-3137
Kimmo T. Tolonen https://orcid.org/0000-0001-9674-7691
Jani Heino https://orcid.org/0000-0003-1235-6613

REFERENCES
Araújo, M., & Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography, 16(6), 743–753. https://doi.org/10.1111/j.1466-8238.2007.00359.x
Beaudrot, L., Acevedo, M. A., Lessard, J.-P., Zvoleff, A., Jansen, P. A., Shell, D., Rovero, F., O’Brien, T., Larney, E., Fletcher, C., Andelman, S., & Ahumada, J. (2019). Local temperature and ecological similarity drive distributional dynamics of tropical mammals worldwide. Global Ecology and Biogeography, 28(7), 976–991. https://doi.org/10.1111/geb.12908
Benke, A. C., Huryn, A. D., Smock, L. A., & Wallace, J. B. (1999). Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeastern United States. Journal of the North American Benthological Society, 18(3), 308–343. https://doi.org/10.2307/1486447
Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C., & Finnegan, S. (2013). Climate change and the past, present, and future of biotic interactions. Science, 341(6145), 499–504. https://doi.org/10.1126/science.1237184
Bottin, M., Soininen, J., Alard, D., & Rosebery, J. (2016). Diatom cooccurrence shows less segregation than predicted from niche modeling. PLoS ONE, 11(4), 1–18. https://doi.org/10.1371/journal.pone.0154581
Bruno, J. F., Stachowicz, J. J., & Bertness, M. D. (2003). Inclusion of facilitation into ecological theory. Trends in Ecology & Evolution, 18(3), 119–125. https://doi.org/10.1016/S0169-5347(02)00045-9
Connor, E. F., Collins, M. D., & Simberloff, D. (2013). The checkered history of checkerboard distributions: Reply. Ecology, 94(11), 2403–2414. https://doi.org/10.1890/12-1471.1
Connor, E. F., & Simberloff, D. (1983). Interspecific competition and species co-occurrence patterns on islands: Null models and the evaluation of evidence. Oikos, 41, 455–465. https://doi.org/10.2307/3544105
D’Amen, M., Mod, H. K., Gotelli, N. J., & Guisan, A. (2018). Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. Ecography, 41, 1233–1244. https://doi.org/10.1111/ecog.03148
D’Amen, M., Pradervand, J. N., & Guisan, A. (2015). Predicting richness and composition in mountain insect communities at high resolution: A new test of the SESAM framework. Global Ecology and Biogeography, 24(12), 1443–1453. https://doi.org/10.1111/geb.12357
Darwin, C. (1859). On the origin of species. J. Murray.
De Mendoza, G., Kaivosoja, R., Grönroos, M., Hjort, J., Ilmonen, J., Kärnä, O., Paasivirta, L., Tokola, L., & Heino, J. (2018). Highly variable species distribution models in a subarctic stream metacommunity: Patterns, mechanisms and implications. Freshwater Biology, 63(August), 33–47. https://doi.org/10.1111/fwb.12993
Diamond, J. M. (1975). Assembly of species communities. In M. L. Cody & J. M. Diamond (Eds.), Ecology and Evolution of communities (pp. 342–444). Belknap Press.
Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., Moretti, M. D., Pagel, J., Pinkert, S., Schleuning, M., Schmidt, S. I., Sheppard, C. S., Steinbauer, M. J., Zeuss, D., & Kraan, C. (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. Global Ecology and Biogeography, 27(9), 1004–1016. https://doi.org/10.1111/geb.12759
Edington, J. M., & Hildrew, A. G. (1995). A revised key to the caseless caddis larvae of the British Isles with notes on their ecology. Freshwater Biological Association.
Elo, M., Jyrkänkallio-Mikkola, J., Ovaskainen, O., Soininen, J., Tolonen, K. T., & Heino, J. (2021). Data from: Does trait-based joint species distribution modelling reveal the signature of competition in stream macroinvertebrate communities? Dryad Digital Repository, https://doi.org/10.5061/dryad.k98sf7m5h
EN ISO 11905-1. (1998). Water quality – Determination of nitrogen - Part 1: Method using oxidative digestion with peroxodisulfate. Swedish Institute for Standards.
Fayle, T. M., Eggleton, P., Manica, A., Yusha, K. M., & Foster, W. A. (2015). Experimentally testing and assessing the predictive power of species assembly rules for tropical canopy ants. Ecology Letters, 18(3), 254–262. https://doi.org/10.1111/ele.12403
García-Girón, J., Heino, J., García-Criado, F., Fernández-Aláez, C., & Alahuhta, J. (2020). Biotic interactions hold the key to understanding metacommunity organisation. Ecography, 43(8), 1180–1190. https://doi.org/10.1111/ecog.05032
Gotelli, N. J., Graves, G. R., & Rahbek, C. (2010). Macroeological signals of species interactions in the Danish avifauna. Proceedings of the National Academy of Sciences of the United States of America, 107(11), 5030–5035. https://doi.org/10.1073/pnas.0914089107
Gotelli, N. J., & McCabe, D. J. (2002). Species co-occurrence: A metaanalysis of J. M. Diamond’s assembly rules model. Ecology, 83(8), 2091–2096. https://doi.org/10.2307/3072040
Grönroos, M., & Heino, J. (2012). Species richness at the guild level: Effects of species pool and local environmental conditions on stream macroinvertebrate communities. Journal of Animal Ecology, 81(3), 679–691. https://doi.org/10.1111/j.1365-2656.2011.01938.x
Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G., & Körber, J. H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. Global Ecology and Biogeography, 16(6), 754–763. https://doi.org/10.1111/j.1466-8238.2007.00345.x
Heino, J. (2009). Species co-occurrence, nestedness and guild-environment relationships in stream macroinvertebrates. Freshwater Biology, 54(9), 1947–1959. https://doi.org/10.1111/j.1365-2427.2009.02250.x

Warton, D. I., Blanchet, F. G., O’Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. C. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology and Evolution*, 30(12), 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J.-A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E., Schmidt, N. M., … Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews of the Cambridge Philosophical Society*, 88(1), 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Elo M, Jyrkänkallio-Mikkola J, Ovaskainen O, Soininen J, Tolonen KT, Heino J. Does trait-based joint species distribution modelling reveal the signature of competition in stream macroinvertebrate communities? *J Anim Ecol*. 2021:90:1276–1287. https://doi.org/10.1111/1365-2656.13453