

RESEARCH
PAPER



Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules?

Anna Astorga^{1*}, Jari Oksanen¹, Miska Luoto², Janne Soininen³,
Risto Virtanen¹ and Timo Muotka^{1,4}

¹Department of Biology, PO Box 3000, FI-90014, University of Oulu, Finland,

²Department of Geosciences and Geography, PO Box 65, FI-00014, University of Helsinki, Finland, ³Department of Environmental Sciences, PO Box 65, FI-00014, University of Helsinki, Finland, ⁴Finnish Environment Institute, Research Department, PO Box 413, FI-90014, University of Oulu, Finland

ABSTRACT

Aim An intensively debated issue in macroecology is whether unicellular organisms show biogeographic patterns different from those of macroorganisms. One aspect of this debate addresses beta diversity, that is, do microbial organisms exhibit distance-decay patterns similar to those of macroorganisms? And if so, is the decay of community similarity caused by spatially limited dispersal or by niche-related factors? We studied the community similarity of stream diatoms, macroinvertebrates and bryophytes across the same set of sites in relation to environmental and geographic distance.

Location A geographical gradient of c. 1100 km in Finland.

Methods We first identified the subset of environmental variables that produced the highest correlation with community similarities for each taxonomic group. Based on these variables, we used partial Mantel tests to separate the independent influences of environmental and geographical distance for distance decay of community similarity, separately for diatoms, bryophytes and macroinvertebrates. Finally, macroinvertebrates were divided into three groups based on their different dispersal categories and a partial Mantel test was used to assess whether each of these groups were differently affected by environmental versus geographic distance, i.e. is dispersal a key factor in tests of niche versus neutral models.

Results The level of environmental control was by far the strongest for diatoms; however, all groups were controlled more by environmental factors than by limited dispersal. Macroinvertebrate species with low dispersal ability were significantly related to geographic distance, while more effective dispersers showed no relationship to geography but were instead strongly related to environmental distance.

Main conclusions Our results suggest that patterns between macro- and microorganisms are not fundamentally different, but the level of environmental control varies according to dispersal ability. The relative importance of niche versus dispersal processes is not simply a function of organism size but other traits (e.g. life-history type, dispersal capacity) may obscure this relationship.

Keywords

Benthic macroinvertebrates, BIO-ENV, bryophytes, diatoms, dispersal, Finland, Mantel test, streams.

*Correspondence: Anna Astorga, Department of Biology, PO Box 3000, FI-90014, University of Oulu, Oulu, Finland.
E-mail: anna.astorga@oulu.fi

INTRODUCTION

A recurrent pattern in the similarity of species composition between ecological communities is that it typically decreases with increasing distance, a phenomenon known as distance

decay (Nekola & White, 1999). Two main explanations have been proposed for this pattern. Niche theory predicts that community similarity decreases with environmental distance, irrespective of geographic proximity, as a result of species differences along environmental gradients (Tilman, 1982).

Neutral theory, by contrast, predicts that the decay of community similarity is caused by spatially limited dispersal, independent of environmental differences between sites (Hubbell, 2001). In reality, niche control and neutrality are the opposite ends of a continuum, and most natural communities lie somewhere between these theoretical extremes (Gravel *et al.*, 2006; Thompson & Townsend, 2006). Furthermore, environmental and geographic distances are typically inter-correlated, rendering tests of these hypotheses problematic (Gilbert & Lechowicz, 2004).

During the past few years, there has been intensive debate on whether unicellular organisms exhibit biogeographic patterns different from those of macroorganisms. The traditional view holds that, being small and extremely abundant, unicellular organisms are ubiquitous dispersers, flourishing wherever they find a suitable environment ('everything is everywhere, but the environment selects'). Thus, unlike most macroorganisms, they lack well-defined biogeographic patterns (Finlay *et al.*, 1996; Godfray & Lawton, 2001). This generalization has now been challenged by a growing body of evidence showing that many microbial organisms have restricted distributions with well-structured spatial patterns of assemblage composition (Green *et al.*, 2004; Foissner, 2006). Thus, patterns of microbial diversity may not be fundamentally different from those of macroorganisms (Horner-Devine *et al.*, 2004; Martiny *et al.*, 2006).

Regressing community similarity against environmental and spatial distance provides an effective means to determine the relative roles of local environmental structuring versus regional control of community composition (Tuomisto *et al.*, 2003; Green *et al.*, 2004; Fierer & Jackson, 2006; Martiny *et al.*, 2006). However, only rarely have the distributions of micro- and macroorganisms been compared across the same set of sites, although such an approach could clearly contribute to our understanding of the importance of dispersal limitation versus local constraints for spatial turnover of ecological communities (Soininen *et al.*, 2007). In one of the few exceptions, Beisner *et al.* (2006) used data from 18 lakes in Canada, showing that the variability of community structure of less easily dispersed species (zooplankton and fish) are better predicted than bacteria and phytoplankton by the spatial distribution of lakes and their connections on the landscape. Bryant *et al.* (2008) documented that patterns of elevational distribution were profoundly different for vascular plants and soil bacteria, suggesting that these communities are not shaped by the same factors.

We studied the community similarity of three taxonomic groups present in most streams: benthic diatoms, macroinvertebrates and aquatic bryophytes. We expected the relative importance of geographic versus environmental distance to exhibit predictable variation across organisms with different dispersal ability, diatoms as unicellular organisms being the best dispersers and macroinvertebrates the poorest dispersers of the studied groups. As the among-stream dispersal of aquatic bryophytes occurs mainly via spores (Stream Bryophyte Group, 1999) that may occasionally disperse long distances (Rydin, 2009), we expected bryophytes to be less dispersal-limited than macroinvertebrates. Thus, we hypothesized that the importance

of geographic distance (dispersal limitation) to species turnover should increase from diatoms to macroinvertebrates, and the importance of environmental distance (niche partitioning) should increase from macroinvertebrates to diatoms. We then divided benthic macroinvertebrates in three groups according to their dispersal capacity to provide a more direct test of the importance of dispersal to the relationship between community similarity and geographic versus environmental distance. Most stream macroinvertebrates possess a winged adult stage, and they could thus be expected to be effective dispersers at the among-stream scale. However, this group exhibits wide variation in dispersal capacity, some species being confined to the aquatic environment, and they might therefore express varying degrees of dispersal limitation (Thompson & Townsend, 2006). We hypothesized only the weakest dispersers to show strong dispersal limitation in their distance-decay pattern.

MATERIALS AND METHODS

Study area

We sampled 14–15 stream sites in each of eight drainage systems in Finland, spanning a north–south gradient of 1100 km. All sampling was conducted during low-flow conditions in late autumn (September to October) 2000–04, the majority of sites being sampled in 2000. To restrict our analyses to a single habitat type, we focused on near-pristine (no point sources of pollution, no obvious signs of human impact in the riparian zone or the stream channel, < 9% of catchment, on average, under anthropogenic impact) streams with a base flow < 0.6 m³ s⁻¹. The catchments sampled spanned a broad environmental gradient, from oligotrophic to mesotrophic conditions [range of total P (TP): 9.0–65.1 µg l⁻¹]. All study sites were in small forest streams of orders one to three, with mean catchment sizes ranging between 13.0 and 43.8 km². The streams drained mainly forested catchments (mean percentage of forests 34.9–68.6%). Mean elevation range varied from 40.9 to 162.0 m a.s.l. More information about the studied catchments and streams is given in Astorga *et al.* (2011).

Biotic sampling

We sampled benthic macroinvertebrates by taking a 2-min kick-net sample (net mesh size 0.3 mm) at each site, aiming to cover most microhabitats present in a stream section of approximately 100 m². Macroinvertebrates and associated material were preserved in 70% alcohol, and they were sorted and identified to the lowest feasible taxonomic level, usually species, in the laboratory. Our analysis includes all other major groups of macroinvertebrates apart from chironomids, which were not identified to species level and were thus excluded from all analyses.

Bryophytes were sampled in 10 randomly placed 50 cm × 50 cm quadrats in those areas of the stream bed that remain wetted at base flow. All bryophyte species and their percentage cover were recorded in the field. For diatom sampling, we divided each study site into five or ten cross-stream transects,

depending on stream width. Ten stones were selected randomly in each transect, and diatoms were scraped off the stones from a predefined area (3.1 cm²) using a plastic template. Subsamples were then pooled into a composite sample for each site. Almost all of the diatom cells had cytoplasmic contents, and thus acid combustion (HNO₃ : H₂SO₄; 2:1) was used to clean frustules of organic material. Cleaned diatoms were mounted in Dirax or Naphrax. Five hundred frustules per sample were identified using phase contrast light microscopy (magnification 1000×), using the keys of Krammer and Lange-Bertalot (1986–91) and Lange-Bertalot & Metzeltin (1996).

Environmental variables

At each site, we measured several local environmental variables. Depth and current velocity (at 0.6× depth) were measured at 40 random locations along evenly spaced cross-channel transects covering the whole study section. Canopy cover was measured at 20 locations along the same transects. Substratum particle size was measured in the same quadrats using a modified Wentworth scale (see Table S1 in Supporting Information for categories and their size). The proportion of each size class was estimated for each quadrat, and these estimates were then averaged to give the weighted mean of substratum size for a site. Water samples were collected and analysed for pH, conductivity, TP and water colour according to the methods of the National Board of Waters and the Environment (1981). Elevation range was used as a proxy for topographic heterogeneity, and it was calculated as the range between the maximum and minimum altitude for the catchment above a site. We also included two climatic variables: (1) mean summer (1 June–30 September) air temperature, and (2) coefficient of variation (CV) of monthly precipitation. For mean temperature, we downscaled a 10-year data set (1996–2005; Finnish Meteorological Institute, Venäläinen & Heikinheimo, 2002) from the original 10 km² grid to a 1 km² grid by relating mean summer temperature to latitude, longitude and altitude of each study site through multivariate regression (r^2 of the model: 0.97). The CV of monthly precipitation was downscaled to a 1-km² grid by kriging interpolation on adjusted precipitation data (Van der Linden & Christensen, 2003). Climatic variables were thus fine-tuned to better describe local variation in climatic conditions, although they still reflect regional variation in climate. Air temperature was used as a proxy of water temperature, and CV in monthly precipitation was used to approximate regional variability in hydrological disturbance (Vinson & Hawkins, 2003).

Data analysis

We calculated similarity in community composition between all site pairs using the Sørensen coefficient on presence–absence data. This coefficient ranges from 0 to 1, high values indicating closely similar species composition between two sites. The Sørensen index is much used in community ecology, and a recent study comparing the performance of several similarity indices recommended the use of this index for analysing binary

community data (Cao & Epifanio, 2010). To examine the relative importance of geographic and environmental distance for each taxonomic group, we first used BIO-ENV (Clarke & Ainsworth, 1993) to identify the subset of environmental variables that produced the highest (nonparametric) correlation with community similarities. Typically, random variation ('noise') increases when poor environmental predictors are used among the important ones and, consequently, a small subset of variables has a higher correlation with community structure than all variables combined (Clarke & Ainsworth, 1993). The basic idea of BIO-ENV is to calculate the Pearson correlation coefficient between distances in the response matrix (in our case, community similarity matrix) and the environmental distance matrix. BIO-ENV calculates the correlation coefficient for every possible combination of predictor variables until it identifies the 'best' fit (i.e. the combination of predictor variables whose distance matrix yields the highest correlation). This best subset of environmental variables was subsequently used to calculate the environmental distance matrix for each group based on Euclidean distances between sites.

Geographical distances were calculated as Euclidean distances, ranging from 0.6 to 1100 km. Simple Euclidean distance has been shown to serve as an approximation for constraints to movement within and between stream networks for some stream invertebrates (e.g. Finn *et al.*, 2006). However, other authors have argued for the use of network instead of overland distances (e.g. Urban *et al.*, 2006; Brown & Swan, 2010). Of the groups studied by us, macroinvertebrates may indeed use stream networks for dispersal (but see Briers *et al.*, 2004, and Macneale *et al.*, 2005, for examples of overland dispersal), but bryophytes and diatoms are much less likely to do so. Obviously, these organisms are carried downstream by water flow, but in our study each stream was represented by only one sampling site. Therefore, it is much more likely that any dispersal between sites was aerial and can be approximated by linear distances. For comparative purposes, and lacking exact knowledge about dispersal pathways in each group, we think it is important that all three groups be analysed similarly; we therefore opted to use linear distances, acknowledging the potential shortcomings of this method. Furthermore, at the across-watershed extent, distances across stream networks are inappropriate, as the watersheds studied by us span a gradient of 1100 km.

The plots of similarity versus distance formed clouds with higher variance at short distances and lower variance at longer distances. For a clearer graphical output, we grouped the community similarity observations into 100-km bins for the geographic distance and 1.0-unit Euclidean distance for the environmental distance. We compared through regression the relationship between community similarity and geographical distance separately for each taxonomic group using: (1) slope of the relationship (distance versus community similarity), (2) initial similarity (similarity at 1-km distance), and (3) halving distance. These three measures represent slightly different aspects of compositional turnover with distance (see Tuomisto, 2010). The slope of the distance-(dis)similarity relationship is one of the most commonly used measures of beta diversity in

ecological studies (e.g. Nekola & White, 1999; Soininen *et al.*, 2007; Qian, 2009). Soininen *et al.* (2007), however, suggested halving distance and initial similarity as measures that consider both the rate of compositional turnover and the initial level of similarity. Halving distance indicates how far from a site you need to go, on average, for the compositional similarity between two sites to have decreased to half of the initial similarity, assuming that the rate of decrease in similarity is constant per unit increase in geographical distance. Halving distance is a biologically feasible measure that can be calculated for any type of regression between community similarity versus geographic distance and is therefore highly suitable for comparative purposes (Soininen *et al.*, 2007). Initial similarity reflects turnover at small spatial extents, with high similarity values indicating low small-scale turnover.

We tested the significance of the regression slopes using a randomization procedure with 1000 iterations. We also tested statistically for differences between the taxonomic groups in all three parameters using a taxon reshuffling procedure. First, we formed a pooled species matrix combining all taxa, and randomly resampled this matrix to form groups of 391 (diatoms), 109 (macroinvertebrates) and 83 (bryophytes) 'species', respectively. A new similarity matrix was then calculated for each of the randomly generated groups, maintaining the geographical distances between sites. This procedure was repeated 999 times, and differences in parameter values in the original data were compared with those of the randomized data sets to obtain the statistical significance of the observed differences.

Given that the studied groups were sampled with different sampling techniques, a legitimate concern might be that differences in patterns of community similarity could be caused by differences in sampling intensity (e.g. diatoms being less effectively sampled than the other two groups). We addressed this concern by examining the effect of random species loss (i.e. lower sampling coverage) by taking random subsamples of our original data. Subsamples of half of the original counts were subsampled from the original data (without replacement, replicated 999 times), then transformed into binary data (as in the original analysis). We then repeated the analysis and estimated initial similarity, slope and halving distance for each subsample. The subsample estimates thus obtained were compared against the observed estimates from the full data. The *P*-value (two-sided) gives the proportion of simulated estimates more extreme than the observed estimate; thus, a low *P*-value means that the observed and simulated estimates differ. Distance-decay rates may still be influenced by the size of the sampling unit (grain) in complicated ways, perhaps an inevitable outcome when comparing organisms of widely different type and size (aquatic plants, insects, unicellular algae). Because of the fixed sample area for all three groups, we were unable to directly control for the effect of area on the decay functions observed. Although sampling half of an area is not strictly the same as taking a random subset of half of the individuals in an area, we used this approach to mimic the situation where sampling intensity is weaker, that is, a greater number of species are missed in sampling. We therefore believe that the effect of variable sample area

on the comparison of distance-decay functions would be closely similar to that of variable number of individuals, i.e. to influence the sampling variance of dissimilarity indices, but not the overall relationship between community dissimilarities and geographic or environmental differences. We used only macroinvertebrates in this exercise because of the three organism groups studied by us they are probably best described by our sampling, and this subsampling procedure thus generates data sets that should more closely resemble other organism groups in sampling intensity.

A relatively strong correlation between environmental and geographic distance was present in our data (Mantel $r = 0.40$), with adjacent sites being most similar in terms of environmental conditions (see Fig. S1). We therefore used partial Mantel tests (with 1000 permutations) to assess the influence of environmental distance on community similarity while holding geographical distance constant, and vice versa. We only used the subset of environmental variables identified by BIO-ENV for each taxonomic group. The R package *vegan* (Oksanen *et al.*, 2008) was used for BIO-ENV, regressions and partial Mantel tests.

To further control the distance–environment covariance, we classified pairs of sites into four categories (quartiles) based on the environmental distance between them. Of these categories, we selected the first and the fourth one for each taxonomic group to represent (1) closely similar environments (sites with low environmental distance, i.e. the lowest quartile, irrespective of geographical distance) versus (2) highly dissimilar environments (sites with high environmental distance; the highest quartile). We then examined the relationship between community similarity and geographic distance within each category, fitting locally fitted weighted sums of squares (lowess) and linear regressions to the data. Significance of the regression slopes was tested using a randomization procedure with 1000 iterations.

Finally, we classified macroinvertebrates into three groups based on their dispersal capacity, using information from the published literature and expert opinion. Dispersal potential was categorized as low (< 30 m; species unlikely to disperse to another stream), intermediate (30–500 m; species that often leave the immediate surroundings of the stream but rarely reach another stream), or high (> 500 m; species that often disperse far enough to reach an adjacent stream or watershed). We were able to provide categorization for 90 invertebrate taxa (see Table S2). We then used partial Mantel tests to assess the relationship between community similarity and geographic or environmental distance (see above) separately for each dispersal category. Environmental distances were based on BIO-ENV analysis to identify the best subset of environmental variables for each dispersal group.

RESULTS

Our data consisted of 109 macroinvertebrate, 83 bryophyte and 391 diatom taxa. The mean number of taxa per site was 17.4 (± 0.64 , 1 SE) for macroinvertebrates, 6.2 (± 0.37) for bryophytes and 48.3 (± 1.52) for diatoms.

Table 1 Set of environmental variables identified for each taxonomic group by BIO-ENV analysis and the overall correlations (Pearson) for each taxonomic group.

Taxa	Environmental variables	Overall correlation
Macroinvertebrates	pH	-0.453
	Water colour	
	Temperature	
	Elevation range	
Bryophytes	Conductivity	-0.422
	Total P	
	Water colour	
	Precipitation (CV)	
Diatoms	Elevation range	-0.611
	pH	
	Conductivity	
	Total P	
	Precipitation (CV)	
	Temperature	
	Elevation range	

CV, coefficient of variation.

BIO-ENV analysis identified different sets of environmental variables for each taxonomic group, although water chemistry variables, particularly pH, conductivity, water colour and TP, were consistently selected for each taxonomic group. Elevation range, a surrogate of stream habitat heterogeneity, was also important for all three groups (Table 1). The overall correlation with environmental factors was stronger for diatoms than for macroinvertebrates and bryophytes.

Community similarity was negatively related to both environmental and geographic distance between sites in all three groups (Fig. 1). The relationship was best approximated by a logarithmic model in each case (see Table S3 for coefficients of determination of all explored models). Initial similarity and regression slopes along geographic distance were closely similar among the three groups (Table 2). Diatoms showed the highest average halving distances (902 km) and bryophytes the lowest (181 km). However, due to wide data scatter, particularly in bryophytes, differences in halving distance were not statistically significant (all $P > 0.15$). Interestingly, a relatively large number of site pairs (11% of all pairwise comparisons) shared no bryophyte taxa. The corresponding number was 2% for macroinvertebrates, whereas in diatoms all site pairs shared at least some taxa (i.e. community similarity was always > 0) (Fig. 1).

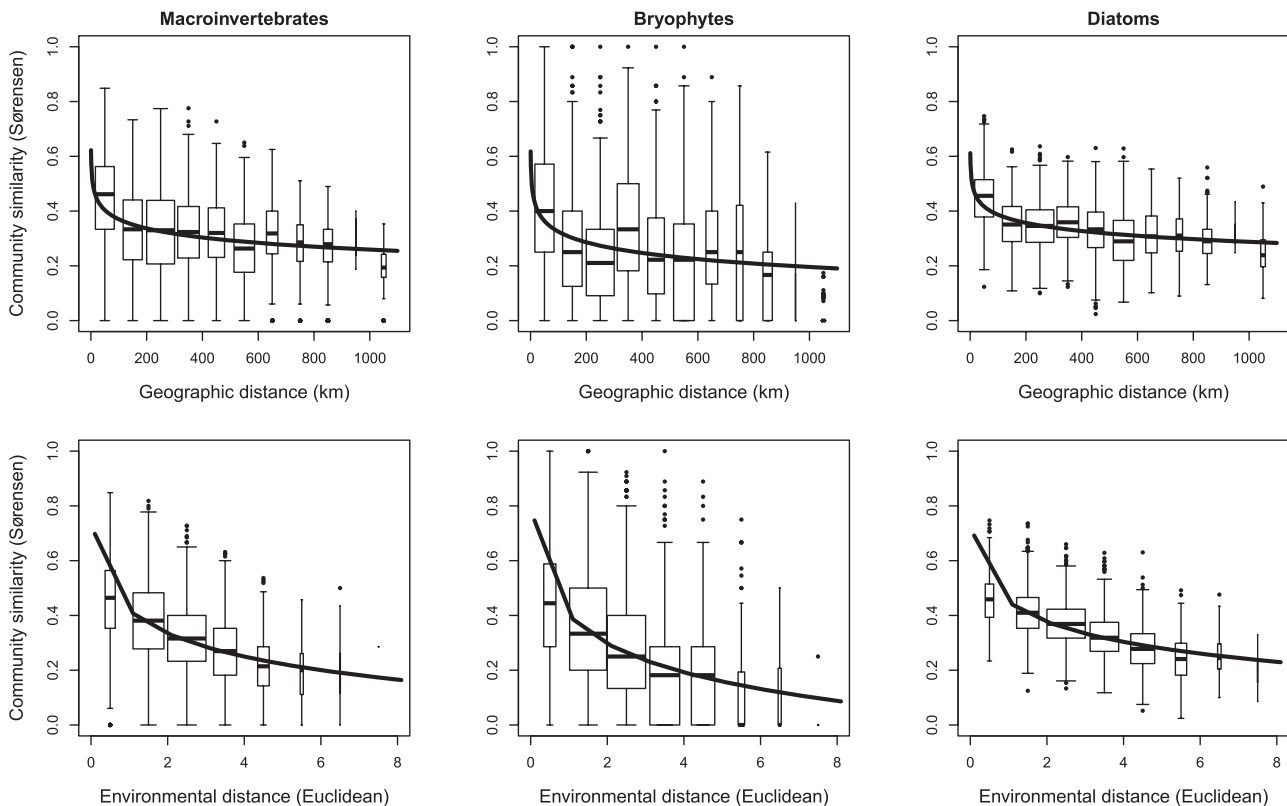


Figure 1 Relationship between community similarity and geographic and environmental distance for stream macroinvertebrates, bryophytes and diatoms. The relationship was best approximated by a logarithmic model in each case (for regression equations, see Table 2). Community similarity observations were grouped into 100-km bins (geographic distance) and 1.0-unit Euclidean distance (environmental distance). The width of each boxplot represents the number of observations in each bin, the wider boxplots having more observations.

Distance	Taxa	R^2	P	IS	Slope	HD (km)
Geographic	Macroinvertebrates	0.114	< 0.001	0.408	-0.048	467
	Bryophytes	0.079	< 0.001	0.417	-0.056	181
	Diatoms	0.206	< 0.001	0.415	-0.043	902
Environmental	Macroinvertebrates	0.207	< 0.001	-	-	-
	Bryophytes	0.180	< 0.001	-	-	-
	Diatoms	0.347	< 0.001	-	-	-

P indicates the significance of the model. IS, initial similarity; HD, halving distance.

Table 3 Partial Mantel correlations between community similarity and environmental distance controlling for geographic distance, and vice versa, for each taxonomic group. Statistical significance for each partial Mantel correlation value is given in parentheses.

Taxa	Environmental distance	Geographic distance
Macroinvertebrates	-0.366 (< 0.001)	-0.069 (0.06)
Bryophytes	-0.346 (< 0.001)	-0.087 (0.007)
Diatoms	-0.499 (< 0.001)	-0.024 (0.226)

The level of community similarity was to some degree affected by sampling intensity, with similarity generally decreasing as sample size decreased ($P = 0.001$), but the pattern in relation to distance remained similar (Fig. S2). Initial similarity and halving distance were unaffected by sampling intensity (both $P > 0.18$), whereas slope became slightly though non-significantly lower as sample size decreased ($P = 0.082$).

Community similarity was generally more closely related to environmental distance when geographical distance was controlled for than vice versa (Table 3). Geographic distance was only significant for bryophytes, but even then the relationship was rather weak. Environmental distance, by contrast, was always significant, accounting for a relatively high proportion (0.35–0.50) of variation in community similarity (Table 3).

We further tested whether the differences in the observed species number among the studied taxonomic groups could have caused the observed differences between groups by randomly resampling 83 species (with replacement, $n = 100$ iterations) for diatoms and macroinvertebrates. A distance matrix was then calculated for each randomization, and Mantel correlation was calculated between each resampled and original distance matrix. Correlations were high overall (benthic invertebrates, mean 0.86, range 0.77–0.95; diatoms, mean 0.65, range 0.58–0.77), and although this procedure created more scatter in the distance-decay relationship, the overall patterns remained consistent when testing the importance of environmental and geographic distance with partial Mantel test: environmental distance always remained significant for both diatoms and invertebrates, whereas geographic distance showed weak and mostly non-significant correlations (highest Mantel $r = 0.136$, $P < 0.001$) to community similarity. These analyses thus confirmed that the observed patterns were largely independent of the variable species numbers among the study groups.

Table 2 Regression parameters for the relationship between community similarity and distance (geographic and environmental) for each taxonomic group.

Table 4 Partial Mantel correlations between community similarity and environmental distance controlling for geographic distance, and vice versa, for each macroinvertebrate dispersal group. Statistical significance for each partial Mantel correlation value is given in parentheses.

Dispersal group	Environmental distance	Geographic distance
Low (41 species)	-0.169 (< 0.001)	-0.146 (< 0.001)
Intermediate (37 species)	-0.261 (< 0.001)	-0.140 (< 0.001)
High (12 species)	-0.346 (< 0.001)	0.009 (0.592)

In closely similar environments, macroinvertebrate and diatom communities showed strong distance decay (both $P < 0.001$; Fig. 2, left panels). In diatoms, however, the distance-decay pattern levelled off fairly rapidly (cf. lowess versus linear regressions in Fig. 2), suggesting that spatial autocorrelation in community similarity was only present at small (within watershed) spatial scales. Bryophyte community similarity in closely similar environments was unrelated to geographical distance ($P = 0.249$; Fig. 2, left panels). In strongly contrasting environments (sites in the fourth environmental distance quartile), macroinvertebrate and diatom communities did not exhibit any relationship between compositional and geographic distance (both $P > 0.40$; Fig. 2, right panels). For bryophytes, community similarity was to some degree related to distance in these environments, but the relationship was quite weak ($P = 0.015$) (Fig. 2, right panels).

Benthic macroinvertebrates with the strongest dispersal capacity bore no relationship to geographic distance, whereas the poorest dispersers showed a distinct decay of community similarity with geographic distance (Fig. 3). In partial Mantel tests, the relationship with geographic distance, with environmental distance partialled out, was significant for both poor and intermediate but not for efficient dispersers (Table 4). Conversely, the partial relationship with environmental distance increased from poor to good dispersers (Table 4).

DISCUSSION

There is an ongoing debate as to whether microorganisms exhibit biogeographical patterns similar to those of macroorganisms, or whether such patterns are obscured by the huge

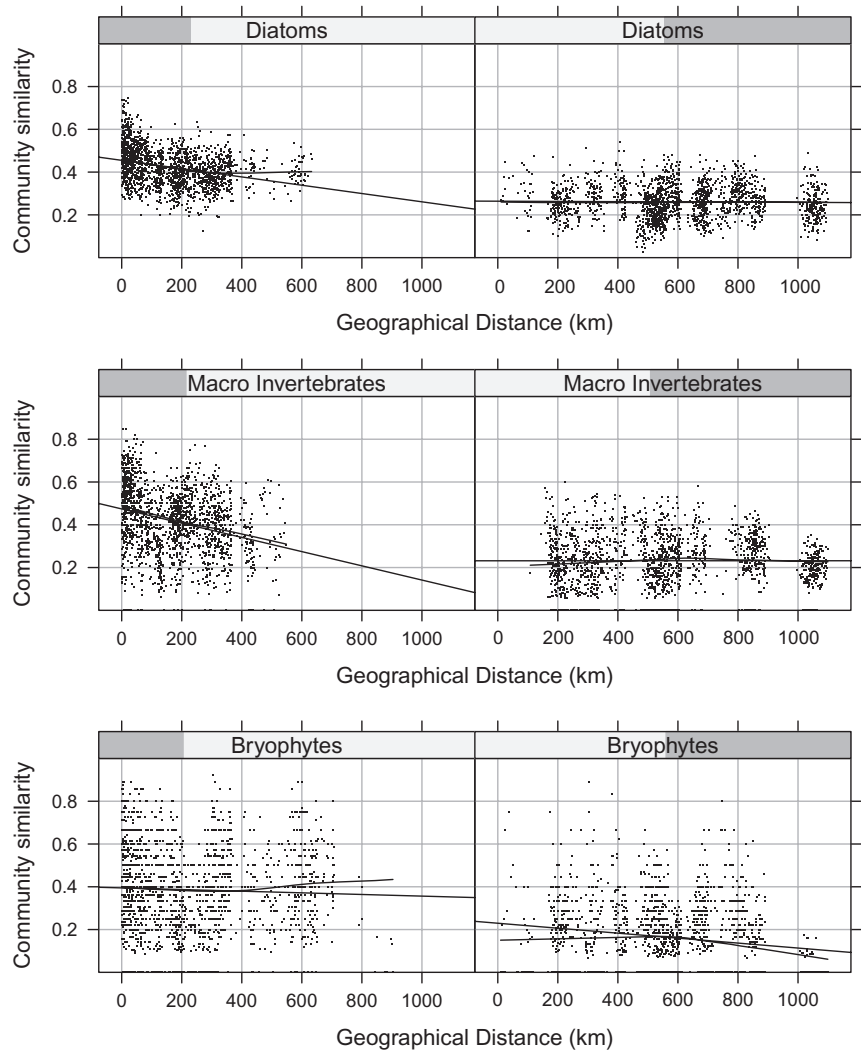


Figure 2 Relationship between community similarity and geographic distance in closely similar environments (the lowest quartile of site pairs based on the environmental distance between pairs; left panels), and in highly dissimilar environments (the highest quartile of site pairs; right panels), for diatoms, macroinvertebrates and bryophytes. The darkly shaded areas in the bars above each plot indicate the environmentally most homogeneous (dark grey bar on the left) and the most heterogeneous quartile (dark grey bar on the right) along geographic distances, the width of the bar being proportional to the level of environmental heterogeneity in each subset. The lines fitted are linear regressions and locally fitted weighted sums of squares (lowess, span = 0.75) regressions.

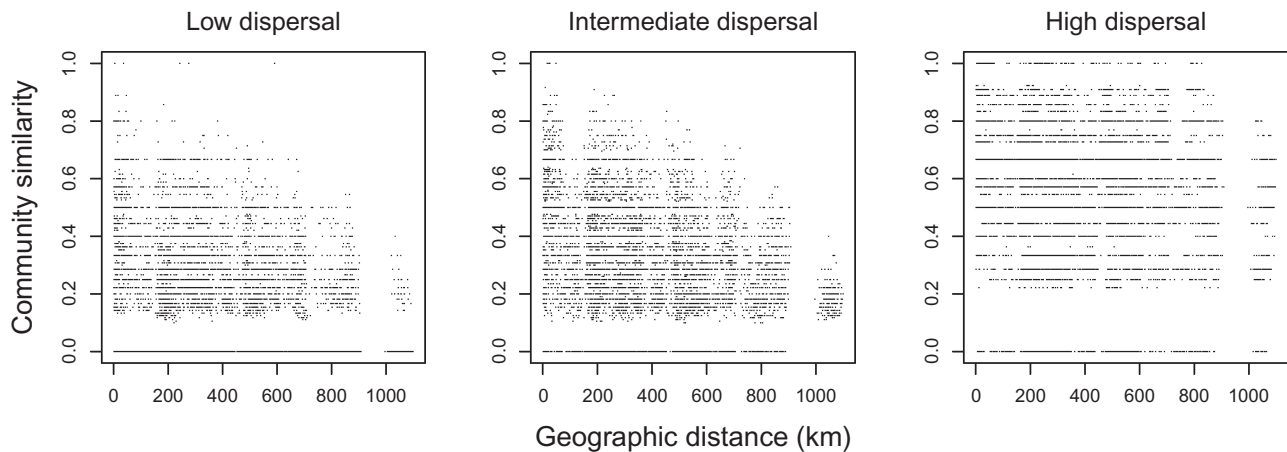


Figure 3 Relationship between community similarity and geographic distance for three macroinvertebrate dispersal categories (for macroinvertebrate dispersal categories, see Table S2).

population sizes and dispersal capacities of microorganisms (Finlay *et al.*, 1996; Finlay, 2002; but see Martiny *et al.*, 2006). This is one of the few studies where the spatial variation in community similarity of micro- and macroorganisms has been

directly compared across the same set of sites, and across a large environmental and geographic gradient. Our results suggest that all three groups (stream diatoms, bryophytes and benthic macroinvertebrates) are more controlled by niche-related factors

than by spatially limited dispersal. The level of environmental control was by far the strongest for diatoms, whereas some groups of benthic macroinvertebrates exhibited relatively strong dispersal limitation. Geographic distance was important only at relatively small spatial extents and in taxa with poor dispersal ability. Thus, there is no discrete boundary between microbes and macroorganisms in their level of spatial structuring, but the difference is gradual. This observation is in line with Hillebrand *et al.* (2001), who showed that while the community similarity of benthic diatoms and ciliates generally decreases with increasing geographic distance, indicating restricted dispersal, the distance effect is much weaker for microorganisms than for metazoans.

At the spatial extent of our study, stream diatoms did not seem to be strongly dispersal limited, but were instead controlled by a few environmental factors, particularly those related to stream water chemistry. Conductivity has been suggested as a universal regulator of lotic diatom community structure (Biggs, 1995), and pH has also been repeatedly identified as a key determinant of diatom (Soininen *et al.*, 2004; Telford *et al.*, 2006) and other microbial communities (Fierer & Jackson, 2006; Fierer *et al.*, 2007), at scales ranging from individual streams and lakes to continents (Bennett *et al.*, 2010). When environmental distance was controlled for, geographic distance was not significantly related to diatom community similarity, suggesting that dispersal limitation was of minor importance to diatoms. However, geographic distance was significantly related to diatom community similarity at small spatial extents (< 200 km), but only in closely similar environments, indicating scale dependency of the distance-decay pattern. This result provides partial support for neutral theory, which predicts that the decay of community similarity should be particularly strong at short spatial distances (Hubbell, 2001; Jones *et al.*, 2006). This observation may not, however, reflect dispersal limitation per se but rather a mass effect ('overly dispersed systems', Leibold & Norberg, 2004) whereby dispersal from nearby sites swamps any local effects of species sorting, resulting in strong small-scale spatial autocorrelation. The absence of such a signal at environmentally distinct sites (right-hand panels of Fig. 2) could be due to a lack of mass effect or the scarcity of sites that are closely located yet environmentally dissimilar. It is thus possible that the compositional similarity of diatom communities in nearby sites may be attributable to a combination of niche control and dispersal-related factors.

Nevertheless, stream diatoms showed more spatially structured distribution (a stronger distance-decay relationship) than most other unicellular organisms, e.g. terrestrial fungi (Green *et al.*, 2004) or bacteria (Horner-Devine *et al.*, 2004). Our findings thus suggest that: (1) unicellular organisms may show distribution patterns not profoundly different from those of higher organisms, and (2) any generalizations across all unicellular organisms are unwarranted, as they exhibit wide variability in patterns of species turnover. Our findings are also in line with Martiny *et al.*'s (2006) notion that the relative importance of environment and history (i.e. distance effects) in microorganisms varies with scale: at continental scales, distance effects may

overwhelm any effects of local environmental factors (see also Vyverman *et al.*, 2007; Bennett *et al.*, 2010), whereas at intermediate scales (100–3000 km), historical effects and contemporary ecological factors both influence biotic communities. However, due to differences in the species concept used, our results may not be directly comparable to other studies on the spatial ecology of microorganisms. Microbial studies typically use molecular-based taxonomy, which may represent a different level of taxonomic resolution than more traditional techniques relying on morphological characteristics. Adoption of molecular techniques may also modify the species concept in diatom taxonomy (Ryner & Armbrust, 2000), with potential implications for the level of biogeographical differentiation detected in freshwater diatoms.

The rate of distance decay of compositional similarity was highest for bryophytes. Both environmental and geographic distances were related to bryophyte community similarity, although the partial correlation coefficient was much lower for geographic distance. Community similarity was related to the same water chemistry variables as for diatoms, i.e. conductivity, TP and water colour. Elevation range and CV of monthly precipitation, a variable reflecting hydrological stability (Vinson & Hawkins, 2003), were also important determinants of bryophyte community similarity. Although we measured a suite of environmental factors known, or assumed, to be important in regulating lotic communities, we were unable to provide a direct measure of disturbance rate in terms of substrate movement, a factor known to be of profound importance to sessile organisms such as stream bryophytes (Muotka & Virtanen, 1995). Thus, our study may in fact have somewhat underestimated the degree of environmental control in stream bryophytes.

Bryophyte community similarity showed a wide scatter, and even nearby sites with closely similar environments often shared no bryophyte taxa. We assume that this pattern is related to randomness of dispersal in bryophytes. Not much is known about the among-stream dispersal of bryophytes (Stream Bryophyte Group, 1999), but many terrestrial studies have documented that long-distance dispersal of bryophytes takes place mainly by wind, aided by thermal uplift of spores and vegetative propagules (Muñoz *et al.*, 2004; Sundberg, 2005). A great majority of spores will land very close (within a few metres) to the parent plant, and therefore bryophytes are often considered strongly dispersal limited (Zartman & Nascimento, 2006). However, the presence of a 'spore cloud' that contains spores from an undefined but apparently very wide area has also been suggested (Hylander, 2009). Propagule rain from these clouds may be relatively random in terms of species composition (Hutsemekers *et al.*, 2008), thus causing a wide scatter in the distance-decay pattern of bryophytes. Because of the nature of dispersal in bryophytes, it is possible that the distance-decay relationship manifests at the scale of tens or hundreds of metres (e.g. among riffles within a stream), having simply gone undetected in this regionally oriented study.

For benthic macroinvertebrates, community similarity was best explained by pH, water colour, temperature and elevation range. When environmental distance was controlled for, geo-

graphic distance was unrelated to macroinvertebrate community similarity. Nonetheless, a more detailed analysis revealed that geographic distance was strongly related to community similarity in closely similar environments, suggesting at least some degree of dispersal limitation in this group. When macroinvertebrates were further divided into dispersal categories, these responded differently to environmental and geographic distance. As expected, species with low dispersal ability were significantly related to geographic distance, while more effective dispersers showed no relationship to geography but were instead strongly related to environmental distance. Indeed, Thompson & Townsend (2006) documented a corresponding pattern at a much smaller, within-watershed, spatial extent, emphasizing the key role of dispersal and local environmental control in explaining the spatial turnover of benthic macroinvertebrate community composition.

In a recent meta-analysis, Shurin *et al.* (2009) found a positive relationship between the level of spatial autocorrelation (indicative of dispersal limitation) and body size of freshwater organisms, but only in lakes and only when fish were included. In their study, macroinvertebrates did not differ from diatoms, phytoplankton or zooplankton, suggesting that they all exhibit similar dispersal ability. They therefore suggested that the size threshold between highly dispersing microorganisms and more dispersal-limited metazoans is larger than the 1-mm limit proposed by Finlay (2002). Our results are at odds with their observation, and a partial explanation may be that their comparison, as in any meta-analysis, was based on data combined across multiple surveys, often conducted by different researchers in different continents, whereas our data were derived from the same set of sites, using strictly standardized sampling protocols. While fishes are undoubtedly more dispersal limited than, for example, freshwater insects, our results show that stream macroinvertebrates cannot be regarded as a unified group of organisms but consist of species that vary widely in terms of dispersal ability and, consequently, level of beta diversity. In fact, Shurin *et al.* (2009) also cautioned against using broad taxonomic groupings when relating the level of spatial autocorrelation to dispersal capacity, because such broad generalizations may hide biologically important detail.

Overall, our study supports the view that niche control interacts with dispersal limitation in determining patterns of spatial turnover in multiple groups of stream organisms, but the relative importance of these processes varies both within and among broad taxonomic groupings. Van der Gucht *et al.* (2007) suggested recently that bacterial communities display strong species sorting over a much broader spatial scale than most macroorganisms because in bacteria sufficient (though not necessarily vast) long-distance dispersal is coupled with extremely fast population growth rates, allowing establishment whenever environmental conditions are suitable. We believe that the same mechanism may also work for freshwater diatoms, at least at the spatial extent of our study, and probably on much wider extents as well. It appears that the relative importance of niche versus dispersal processes is not simply a function of organism size (see also Martiny *et al.*, 2006) but that other species traits (e.g. life-

history type, dispersal type) may obscure this deceptively simple relationship in many taxonomic groups. Finally, information about the scale and constraints of dispersal in each organism group studied may have a bearing on how these organisms are used in stream bioassessment and conservation. Bennett *et al.* (2010) suggested recently that diatoms and other unicellular organisms should be excellent candidates for detecting environmental impacts because of their very effective dispersal and strong environment–community relationship across broad geographic extents (niche conservatism). Our results, however, suggest that stream managers using diatoms in regional-scale biomonitoring protocols should carefully consider the possibility that the environment–organism signal might be partly obscured by mass effects. At the scale where most impact assessments are conducted, dispersal limitation may not be a crucial factor, and macroorganisms whose ecological requirements and functionally important traits are well known (e.g. benthic macroinvertebrates) will in most cases be a highly appropriate target group for regional-scale stream bioassessment.

ACKNOWLEDGEMENTS

This study was funded by the Academy of Finland and University of Oulu (Thule Institute). We acknowledge the constructive comments made by three referees and G. Mittelbach on a previous draft of the manuscript.

REFERENCES

- Astorga, A., Heino, J., Luoto, M. & Muotka, T. (2011) Freshwater biodiversity at a regional scale: determinants of macroinvertebrate species richness in headwater streams. *Ecography*, doi: 10.1111/j.1600-0587.2010.06427.x.
- Beisner, B.E., Peres-Neto, P.R., Lindström, E.S., Barnett, A. & Longhi, M.L. (2006) The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology*, **87**, 2985–2991.
- Bennett, J.R., Cumming, B.F., Ginn, B.K. & Smol, J.P. (2010) Broad-scale environmental response and niche conservatism in lacustrine diatom communities. *Global Ecology and Biogeography*, **19**, 724–732.
- Biggs, B.J.F. (1995) The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. *Freshwater Biology*, **33**, 419–438.
- Briers, R.A., Gee, J.H.R., Cariss, H.M. & Geoghegan, R. (2004) Inter-population dispersal by adult stoneflies detected by stable isotope enrichment. *Freshwater Biology*, **49**, 425–431.
- Brown, B.L. & Swan, C.M. (2010) Dendritic network structure constrains metacommunity properties in riverine ecosystems. *Journal of Animal Ecology*, **79**, 571–580.
- Bryant, J.A., Lamanna, C., Morlon, H., Kerkhoff, A.J., Enquist, B.J. & Greens, J.L. (2008) Microbes on mountainsides: contrasting patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences USA*, **105**, 11505–11511.

- Cao, Y. & Epifanio, J. (2010) Quantifying the responses of macroinvertebrate assemblages to simulated stress: are more accurate similarity indices less useful? *Methods in Ecology and Evolution*, **1**, 380–388.
- Clarke, K.R. & Ainsworth, M. (1993) A method of linking multivariate community structure to environmental variables. *Marine Ecology Progress Series*, **92**, 205–219.
- Fierer, N. & Jackson, R.B. (2006) The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences USA*, **103**, 626–631.
- Fierer, N., Morse, J.L., Berthrong, S.T., Bernhardt, E.S. & Jackson, R.B. (2007) Environmental controls on the landscape-scale biogeography of stream bacterial communities. *Ecology*, **88**, 2162–2173.
- Finlay, B.J. (2002) Global dispersal of free-living microbial eukaryote species. *Science*, **296**, 1061–1063.
- Finlay, B.J., Esteban, G.F. & Fenchel, T. (1996) Global diversity and body size. *Nature*, **383**, 132–133.
- Finn, D.S., Theobald, D.M., Black, W.C. & Poff, N.L. (2006) Spatial population genetic structure and limited dispersal in a Rocky Mountain alpine stream insect. *Molecular Ecology*, **15**, 3553–3566.
- Foissner, W. (2006) Biogeography and dispersal of microorganisms: a review emphasizing protists. *Acta Protozoologica*, **45**, 111–136.
- Gilbert, B. & Lechowicz, M.J. (2004) Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Sciences USA*, **101**, 7651–7656.
- Godfray, H.C.J. & Lawton, J.H. (2001) Scale and species numbers. *Trends in Ecology and Evolution*, **16**, 400–404.
- Gravel, D., Canham, C.D., Beaudet, M. & Messier, C. (2006) Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters*, **9**, 399–409.
- Green, J.L., Holmes, A.J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., Gillings, M. & Beattie, A. (2004) Spatial scaling of microbial eukaryote diversity. *Nature*, **432**, 747–750.
- Hillebrand, H., Watermann, F., Karez, R. & Berninger, U.G. (2001) Differences in species richness patterns between unicellular and multicellular organisms. *Oecologia*, **126**, 114–124.
- Horner-Devine, C.M., Lage, M., Hughes, J.B. & Bohannan, B.J.M. (2004) A taxa–area relationship for bacteria. *Nature*, **432**, 750–753.
- Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.
- Hutsemekers, V., Dopagne, C. & Vanderpoorten, A. (2008) How far and how fast do bryophytes travel at the landscape scale? *Diversity and Distributions*, **14**, 483–492.
- Hylander, K. (2009) No increase in colonization rate of boreal bryophytes close to propagule sources. *Ecology*, **90**, 160–169.
- Jones, M.M., Tuomisto, H., Clark, D.B. & Olivas, P. (2006) Effects of mesoscale environmental heterogeneity and dispersal limitation on floristic variation rain forest ferns. *Journal of Ecology*, **94**, 181–195.
- Krammer, K. & Lange-Bertalot, H. (1986–1991) *Bacillariophyceae. Süßwasserflora von Mitteleuropa 2 (1–4)*. Gustav Fischer Verlag, Stuttgart.
- Lange-Bertalot, H. & Metzeltin, D. (1996) *Iconographia diatomologica, Volume 2. Indicators of oligotrophy. 800 taxa representative of three ecologically distinct lake types: carbonate buffered, oligodystrophic, weakly buffered soft water*. Koeltz Scientific Books, Koenigstein.
- Leibold, M.A. & Norberg, J. (2004) Biodiversity in metacommunities: plankton as complex adaptive systems? *Limnology and Oceanography*, **49**, 1278–1289.
- Macneale, K.H., Peckarsky, B.L. & Likens, G.E. (2005) Stable isotopes identify dispersal patterns of stonefly populations living along stream corridors. *Freshwater Biology*, **50**, 1117–1130.
- Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Övreås, L., Reysenbach, A.-L., Smith, V.H. & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology*, **4**, 102–112.
- Muñoz, J., Felicísimo, A.M., Cabezas, F., Burgaz, A.R. & Martínez, I. (2004) Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, **304**, 1144–1147.
- Muotka, T. & Virtanen, R. (1995) The stream as a habitat template for bryophytes: species' distributions along gradients in disturbance and substratum heterogeneity. *Freshwater Biology*, **33**, 141–160.
- National Board of Waters and the Environment (1981) *Vesihallinnon analyysimenetelmät*. National Board of Waters, Finland, Report 213.
- Nekola, J.C. & White, P.S. (1999) The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, **26**, 867–878.
- Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G., Stevens, M.H. & Wagner, H. (2008) *Vegan: community ecology package*. Version 1.13-1. Available at: <http://vegan.r-forge.r-project.org> (accessed 28 April 2011).
- Qian, H. (2009) Beta diversity in relation to dispersal ability for vascular plants in North America. *Global Ecology and Biogeography*, **18**, 327–332.
- Rydin, H. (2009) *Population and community ecology of bryophytes*. *Bryophyte biology* (ed. by B. Goffinet and A.J. Shaw), pp. 393–444. Cambridge University Press, Cambridge.
- Ryneason, T.A. & Armburst, E.V. (2000) DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom *Ditylum brightwellii*. *Limnology and Oceanography*, **45**, 1329–1340.
- Shurin, J.R., Cottenie, K. & Hillebrand, H. (2009) Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia*, **159**, 151–159.
- Soininen, J., Paavola, R. & Muotka, T. (2004) Benthic diatom communities in boreal streams: community structure in relation to environmental and spatial gradients. *Ecography*, **27**, 330–342.
- Soininen, J., McDonald, R. & Hillebrand, H. (2007) The distance decay of similarity in ecological communities. *Ecography*, **30**, 3–12.

- Stream Bryophyte Group (1999) Roles of bryophytes in stream ecosystems. *Journal of the North American Benthological Society*, **18**, 151–184.
- Sundberg, S. (2005) Larger capsules enhance short-range spore dispersal in *Sphagnum*, but what happens further away? *Oikos*, **108**, 115–124.
- Telford, R.J., Vandvik, V. & Birks, H.J.B. (2006) Dispersal limitations matter for microbial morphospecies. *Science*, **312**, 1015.
- Thompson, R. & Townsend, C. (2006) A truce with neutral theory: local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *Journal of Animal Ecology*, **75**, 476–484.
- Tilman, D. (1982) *Resource competition and community structure*. Princeton University Press, Princeton, NJ.
- Tuomisto, H. (2010) A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, **33**, 23–45.
- Tuomisto, H., Ruokolainen, K. & Yli-Halla, M. (2003) Dispersal, environment, and floristic variation of western Amazonian forests. *Science*, **299**, 241–244.
- Urban, M.C., Skelly, D.K., Burchsted, D., Price, W. & Lowry, S. (2006) Stream communities across a rural–urban landscape gradient. *Diversity and Distribution*, **12**, 337–350.
- Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., Jeppesen, E., Conde-Porcuna, J.-M., Schwenk, K., Zwart, G., Degans, H., Vyverman, W. & De Meester, L. (2007) The power of species sorting: local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences USA*, **104**, 20404–20409.
- Van der Linden, S. & Christensen, J.H. (2003) Improved hydrological modeling for remote regions using a combination of observed and simulated precipitation data. *Journal of Geophysical Research*, **108**, 4072–4083.
- Venäläinen, A. & Heikinheimo, H. (2002) Meteorological data for agricultural applications. *Physics and Chemistry of Earth*, **27**, 1045–1050.
- Vinson, M.A. & Hawkins, C.P. (2003) Broad-scale geographical patterns in local stream insect genera richness. *Ecography*, **26**, 751–767.
- Vyverman, W., Verleyen, E., Sabbe, K., Vanhoutte, K., Sterken, M., Hodgson, D.A., Mann, D.G., Juggins, S., Van de Vijver, B., Jones, V., Flower, R., Roberts, D., Chepurinov, V.A., Kilroy, C.,

- Vanormelingen, P. & De Wever, A. (2007) Historical processes constrain patterns in global diatom diversity. *Ecology*, **88**, 1924–1931.
- Zartman, C.E. & Nascimento, H.E.M. (2006) Are habitat-tracking metacommunities dispersal limited? Inferences from abundance–occupancy patterns of epiphylls in Amazonian forest fragments. *Biological Conservation*, **127**, 46–54.

SUPPORTING INFORMATION

Additional Supporting information may be found in the online version of this article:

Figure S1 Relationship between geographic and environmental distance between all site pairs (Mantel correlation = 0.40).

Figure S2 Decay of community similarity against geographic distance in the original sample and random subsamples half of the size of the original community data.

Table S1 Substratum particle size was measured in ten 50 cm × 50 cm quadrats placed randomly in each riffle, using a modified Wentworth scale. Category, class and diameter range for each category are given.

Table S2 Classification of benthic macroinvertebrate species into three groups based on potential dispersal capacity.

Table S3 Coefficients of determination (R^2) of all the regression models explored.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Anna Astorga's research interest is the ecology and biodiversity of freshwater organisms.

Author contributions: A.A., T.M., J.S. and J.O. conceived the ideas; T.M., J.S., M.L. and R.V. provided the data; A.A. and J.O. analysed the data; and A.A. and T.M. led the writing.

Editor: Gary Mittelbach