

Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States

LUIS MAURICIO BINI,^{1,6} VÍCTOR LEMES LANDEIRO,² ANDRÉ ANDRIAN PADIAL,³ TADEU SIQUEIRA,⁴ AND JANI HEINO⁵

¹*Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás, CP 131, CEP 74001-970, Goiânia, Goiás, Brasil*

²*Departamento de Botânica e Ecologia, IB, Universidade Federal de Mato Grosso, CEP 78060-900, Cuiabá, MT, Brasil*

³*Departamento de Botânica, Universidade Federal do Paraná, cp 19031, CEP 81531-990, Curitiba, PR, Brasil*

⁴*Departamento de Ecologia, Universidade Estadual Paulista, CEP 13506-900, Rio Claro, SP, Brasil*

⁵*Finnish Environment Institute, Natural Environment Centre, Ecosystem Change Unit, P.O. Box 413, FI-90014, University of Oulu, Oulu, Finland*

Abstract. Beta diversity, the spatial or temporal variability of species composition, is a key concept in community ecology. However, our ability to predict the relative importance of the main drivers of beta diversity (e.g., environmental heterogeneity, dispersal limitation, and environmental productivity) remains limited. Using a comprehensive data set on stream invertebrate assemblages across the continental United States, we found a hump-shaped relationship between beta diversity and within-ecoregion nutrient concentrations. Within-ecoregion compositional dissimilarity matrices were mainly related to environmental distances in most of the 30 ecoregions analyzed, suggesting a stronger role for species-sorting than for spatial processes. The strength of these relationships varied considerably among ecoregions, but they were unrelated to within-ecoregion environmental heterogeneity or spatial extent. Instead, we detected a negative correlation between the strength of species sorting and nutrient concentrations. We suggest that eutrophication is a major mechanism disassembling invertebrate assemblages in streams at a continental scale.

Key words: *assemblage structure; beta diversity; continental United States; ecoregions; environmental heterogeneity; eutrophication; metacommunity; nutrient concentration; spatial extent; species sorting; stream invertebrate; Wadeable Streams Assessment.*

INTRODUCTION

Spatial variation in biodiversity has received much interest in the last century, yet our knowledge of the alpha, beta, and gamma components of species diversity remains limited for many ecological systems, organisms, and regions. Although the term “beta diversity” was coined decades ago (Whittaker 1960), it has not been a topic of rigorous investigation until the last two decades (Anderson et al. 2011). In the present study we investigated two facets of beta diversity: (1) the variation in assemblage structure among a set of sites (henceforward, “beta diversity”) and (2) the strength of relationships between compositional dissimilarities and environmental or spatial distances (Anderson et al. 2011). We examined relationships between these two facets of beta diversity and environmental heterogeneity, spatial extent, and average environmental conditions.

In an analysis encompassing multiple regions (each one with a number of sites), variation in beta diversity among these regions (see Fig. 4 of Anderson et al. 2011) is often thought to be caused by environmental

heterogeneity (e.g., Veech and Crist 2007). This relationship is expected because more heterogeneous conditions in a region provide more niche opportunities, leading to greater variation in species composition between sites (Leibold et al. 2004). Also, spatial extent is likely to be correlated with beta diversity because the likelihood of a species dispersing from one site to another should decrease with increasing region size (Hubbell 2001, Shurin et al. 2009).

Nutrient concentration has also been shown to be an important factor modulating beta-diversity patterns (Van der Gucht et al. 2007). Nutrients can be regarded as proxies for productivity (Langenheder et al. 2012) and, in this context, a positive relationship between beta diversity and productivity has been detected in experimental (Chase 2010) and observational studies (e.g., Langenheder et al. 2012) conducted at different spatial scales. However, using gridded data (cells of ~676 km²) over a large spatial extent, Bonn et al. (2004) found a hump-shaped relationship between productivity (normalized-difference vegetation index, NDVI) and beta diversity in southern African birds. Thus, similar to the uncertainties regarding the species richness–productivity relationship (see the review by Whittaker [2010]), even the form of the relationship between beta diversity and productivity is open to debate.

Manuscript received 16 April 2013; revised 15 October 2013; accepted 3 December 2013. Corresponding Editor: W. V. Sobczak.

⁶ E-mail: bini@ufg.br

Empirical studies have typically found that either the environment alone or both the environmental and spatial distances between sites are related to variation in compositional dissimilarity (e.g., Thompson and Townsend 2006). However, the conditions under which the environment should be important and those when spatial effects should increase in importance remain largely unknown. In this context, variation in the strength of the relationships between community dissimilarity and geographical or environmental distances (see Fig. 3 in Anderson et al. 2011) can be associated with differences in the spatial extent of the study regions. Greater spatial distances between sites within a region are likely to lead to higher strengths of relationship between community dissimilarity and geographical distances because dispersal limitation should play an increasingly important role with increasing geographic distance (Hubbell 2001, Shurin et al. 2009). Similarly, environmental heterogeneity is thought to be positively related to the strength of relationships between community dissimilarity and environmental distances because species sorting should be increasingly more important over wider environmental gradients (Leibold et al. 2004).

The importance of stochastic assembly relative to species-sorting processes is expected to increase with productivity (Chase 2010). The idea behind a strong role of stochastic assembly in more productive environments may be understood if one considers that a large fraction of the regional species pool can maintain viable populations in these environments. Conversely, Chase (2010) found that only a subset of the total species persisted in environments with low productivity, suggesting a stronger role for species-sorting processes. However, other studies suggest a positive association between the efficiency of species sorting and productivity assessed as the degree of eutrophication (Van der Gucht et al. 2007, Langenheder et al. 2012).

Streams are suitable model systems for examining the interacting effects of spatial extent and environmental heterogeneity on beta diversity because streams harbor taxa that vary widely in their dispersal modes (e.g., taxa with or without flying adults) and range from relatively environmentally homogeneous to extremely heterogeneous systems. Our study benefits from a comprehensive data set gathered at a fine spatial resolution (i.e., individual stream sites) and covering a large spatial extent (Hawkins et al. 2008, Paulsen et al. 2008a), a rare combination in macroecological studies (Beck et al. 2012). Specifically, we examined 689 local assemblages of stream macroinvertebrates across 30 ecoregions of the conterminous United States. We predicted that within-ecoregion beta diversity would be positively correlated with within-ecoregion (1) environmental heterogeneity, (2) spatial extent, and (3) nutrient concentration. We also predicted that the strength of the relationships between the compositional dissimilarity and environmental or geographical distances would be positively

related to environmental heterogeneity and spatial extent of the ecoregions, respectively. To our knowledge, these relationships have not been tested explicitly in the field, most likely due to the paucity of available data sets (i.e., covering multiple regions, each with different sampling sites). Finally, following Chase (2010), we predicted that the strength of relationships between community dissimilarity and explanatory distance matrices (spatial or environmental) would be weakest in regions with high nutrient concentrations, suggesting an increased role of deterministic processes in community assembly at nutrient-poor sites.

METHODS

The stream data set

We used a large data set collected by the United States Environmental Protection Agency (US-EPA, *available online*⁷) to assess the biological condition of wadeable streams in the United States (Hawkins et al. 2008, Paulsen et al. 2008a). After removing missing values, the data set included 30 level-three ecoregions (sensu Omernik 1995) with ≥ 10 stream sites, for a total of 689 stream sites. The maximum number of stream sites within an ecoregion was 62.

Macroinvertebrate data set

The complete U.S. EPA data set comprised 918 taxonomic units, of which 739 were identified to the genus level; the remaining 179 taxonomic units were identified at the family, order, class, or phylum level and were not included in our analyses. After excluding the ecoregions with fewer than 10 streams and the sites with fewer than 300 individuals (see next paragraph), the data contained 606 genera. The use of data sets that contain taxa above the species level to represent ecological patterns in communities has been debated, at least with regard to aquatic systems (Lenat and Resh 2001). Although many authors defend the use of species-level data, aquatic studies have often found that the species-level patterns of macroinvertebrate assemblages are well reproduced through a multivariate analysis of genus- or family-level data (Heino 2008). Moreover, genus is thought to represent the main taxonomic unit with respect to ecological variation in aquatic invertebrates (Wiggins and MacKay 1978). We therefore consider the U.S. EPA's genus-level data to adequately portray the variation in assemblage structure within and between ecoregions in our multivariate analyses.

The Wadeable Streams Assessment (WSA) used a standard field collection method (USEPA 2004), but the number of individuals reported in EPA website are based on nonuniform subsampling of the entire samples (Charles P. Hawkins, *personal communication*). Thus, we wrote an R code to randomly resample individuals from

⁷ <http://water.epa.gov/type/rsl/monitoring/streamsurvey/index.cfm>

each site to a fixed count of 300 individuals (for an evaluation of the effects of sampling effort on sensitivities to stress, see Cao and Hawkins 2005). All sites with <300 individuals and all taxa registered as NA (“not applicable”) were excluded prior to the analyses.

Environmental variables

We included a total of 25 environmental variables that could be retrieved from the EPA website in our analysis: watershed area (km²); elevation (m); canopy cover (quantified using a spherical densitometer, counts from 0 to 17); temperature (°C); instantaneous discharge (m³/s); bankfull width (m); depth (cm); wetted width (m); the proportion (in area) of the stream bed covered by algae, macrophytes, wood debris, small debris, overhanging vegetation, undercut banks, boulders, and artificial structures; potassium (µeq/L); NH₄ (µeq/L); color (PCU); conductivity (µS/cm at 25°C); total nitrogen (µg/L); total phosphorus (µg/L); suspended solids (mg/L); turbidity (NTU; nephelometric turbidity units); and alkalinity (µeq/L). A description of the field measurements and laboratory methods used to determine these variables can be found in USEPA (2004). In addition, long-term climatic data (annual temperature, temperature seasonality, annual precipitation and precipitation seasonality) were obtained from WorldClim (Hijmans et al. 2005). These or related variables have previously been found to be associated with variation in the assemblage structure of stream macroinvertebrates in large-scale studies (Carlisle and Hawkins 2008, Carlisle et al. 2008, Hawkins et al. 2010, Chinnayakanahalli et al. 2011). Also, these variables were used in an attempt to describe environmental heterogeneity in a comprehensive way, as environmental heterogeneity is thought to be a key predictor of beta diversity.

Resemblance matrices, ordination, and average differences in assemblage structure and environmental conditions

We constructed biological resemblance matrices from the sites-by-genera data set based on the Bray-Curtis (abundance data), Sørensen (presence-absence), and Simpson (presence-absence) coefficients. We used the Bray-Curtis coefficient for the abundance data, $\log(y + 1)$ -transformed, to emphasize the abundant species and the Sørensen coefficient for the presence-absence data to give equal weight to both rare and abundant species. The Sørensen coefficient incorporates both taxon replacement and taxon richness differences. In contrast, the Simpson coefficient is independent of taxon richness differences and more directly portrays the true compositional differences between sites (Koleff et al. 2003). The standardized Euclidean distance was employed to calculate an environmental dissimilarity matrix (**E**) using all environmental variables.

Using the distance matrices derived from the biological and environmental data sets, we performed principal coordinate analyses (PCoA; Legendre and Legendre

1998) to reduce the dimensionality of the data and ordinate the sites. The first two axes of the PCoA applied to the environmental data set were used as explanatory variables of our main response variables (i.e., beta diversity and strengths of relationship between biological and environmental or geographic distance matrices). We then performed a permutational multivariate analysis of variance (PERMANOVA; Anderson 2001) using the distance matrices to test for between-ecoregion differences in assemblage composition and environmental conditions.

Differences in the heterogeneity of the assemblage structure and environmental conditions

We also used the biological and environmental dissimilarity matrices described above to evaluate whether ecoregions differ in terms of beta diversity and environmental heterogeneity. For this task, we used the permutational analysis of multivariate dispersions (PERMDISP; Anderson et al. 2006). This method is similar to the Levene’s test of homogeneity of variances in the univariate case.

Measures of beta diversity, environmental heterogeneity, and spatial extent

Following Anderson et al. (2006), we estimated beta diversity as the average compositional dissimilarity (d_B) from individual streams to the centroid formed by all streams within an ecoregion. Thus, the higher the value of d_B in an ecoregion, the higher is beta diversity of this ecoregion. According to our predictions, d_B should be positively correlated with environmental heterogeneity, spatial extent and average environmental conditions (see *Modeling beta diversity*).

By applying the method proposed by Anderson et al. (2006) to matrix **E** and to the matrix of geographic distances between streams, we also estimated the environmental heterogeneity (d_{EH}) and spatial extent (d_S) of each ecoregion, respectively. However, we reasoned a priori that d_{EH} , as a composite variable estimated using all environmental variables, could mask the effects of key environmental factors on our response variables. We then calculated more-specific metrics of environmental heterogeneity to be used as explanatory variables in the regression models. We did this by also applying the method proposed by Anderson et al. (2006), separately, on four subsets of variables. The first subset included general physiographical characteristics and stream attributes (d_{strm} : watershed area, elevation, canopy cover, instantaneous discharge, bankfull width, depth, wetted width, and the areal proportion of overhanging vegetation). The second subset encompassed substrate characteristics (d_{subs} : the areal proportion of algae, macrophytes, debris, undercut banks, boulders, and artificial structures). The third subset included only water-quality characteristics (d_{WQ} : temperature, potassium, NH₄, color, conductivity, total nitrogen, total phosphorus, suspended solids, turbidity,

and alkalinity). The fourth subset included the climatic variables from WorldClim (d_{WC} : annual temperature, temperature seasonality, annual precipitation, precipitation seasonality). These quantities (i.e., d_B , d_{EH} , d_{strm} , d_{subs} , d_{WQ} , d_{WC} , and d_S) were then used to formulate our models (see below and Appendix A: Fig. A1).

Modeling beta diversity

We tested the hypothesis that the average distance from individual streams to their ecoregion centroids (i.e., the variation in invertebrate assemblage structure or beta diversity according to Anderson et al. 2006, 2011) defined by the biological data (d_B) was positively correlated with our measures of environmental heterogeneity (d_{EH} or d_{strm} , d_{subs} , d_{WQ} , and d_{WC}) and spatial extent (d_S).

The first PCoA axis derived from the environmental data was positively (considering absolute loadings ≥ 0.70) correlated with nutrient concentrations (K, N, and P), conductivity, temperature, suspended solids and turbidity (Appendix B: Table B1). As expected, these results were consistent with those obtained by Herlihy et al. (2008) using largely the same data set. Therefore, we used this axis as a new predictor (i.e., in addition to d_S and d_{EH}) to test the responsiveness of beta diversity (d_B) to this set of variables (i.e., the ones correlated with the first PCoA axis). The second axis, which was also included as a predictor of d_B , was positively correlated with wetted width (Appendix B: Table B1), a measure of stream size. Significant partial regression coefficients associated with these axes would indicate responsiveness of the variation in assemblage structure (d_B) to the average environmental conditions.

Modeling the strength of relationships between compositional dissimilarities and environmental or spatial distances

We used Mantel and partial Mantel tests to estimate the strength of relationships between the compositional dissimilarity matrices and the explanatory matrices (environmental or geographical distances) within each ecoregion. The r values of the partial Mantel tests pertaining to pure spatial and nonspatially structured environmental effects were further employed in regression analyses in which an r value was considered the effect size of the strength of the relationships within each ecoregion and was used as a response variable. As explanatory variables in these across-ecoregion regressions ($n = 30$), we used the first two axes derived from the PCoA applied to the environmental data, and the average distance from individual streams to their within-ecoregion centroids in geographical (d_S) and environmental (d_{EH}) spaces. Therefore, the across-ecoregion regressions tested the relationship between the effect size and the four ecoregion characteristics. We also used generalized linear models (GLMs) with a log-link function to quantify the relationship between assemblage dissimilarity and the explanatory matrices

(geographic and environmental distances). According to Millar et al. (2011), GLMs are the best approach to cope with dissimilarity values of 1 (or 0, if similarity coefficients are employed). These values are likely to occur when the environmental or geographic distances are so large that no species are shared between the sites. Thus, the GLM-estimated slopes were also used as response variables in the regression models, with the PCoA axes scores, spatial extent (d_S), and environmental heterogeneity (d_{EH}) being used as the explanatory variables.

Spatial autocorrelation analysis

We used Moran's I coefficient to test for spatial autocorrelation in the original variables and in the residuals of all of the models described above based on 10 distance classes. The number of distance classes was defined using the Sturges' rule (Legendre and Legendre 1998), and the global significance of the spatial correlograms was tested using Bonferroni's correction (as suggested by Oden 1984), at $P < 0.05/10$ (where 10 is the number of distance classes). The spatial autocorrelation analysis was conducted to test for significant spatial patterns and to verify whether spatial autocorrelation was an issue that we should address in the analyses described above.

We used 9999 permutations to assess the significance levels in PERMDISP, PERMANOVA, and Mantel tests using the package *vegan* (Oksanen et al. 2013) for the R language and environment for statistical computing (version 2.15.1; R Core Development Team 2012). Multiple regression models (including GLMs) and Moran's I coefficients were estimated using the R packages *stats* and *spdep* (Bivand et al. 2013), respectively.

RESULTS

We found that the ecoregions differed significantly among each other in terms of their average environmental conditions (PERMANOVA results, $F = 12.3$, $P < 0.001$; Fig. 1a) and assemblage structure for both abundance and presence-absence data (Appendix B: Table B2; Fig. 1b). The spatial autocorrelation analyses applied to the ecoregions' mean scores on the first principal coordinate axis (PCoA) retrieved from each data set further indicated that the environmental conditions and assemblage structure (Fig. 1c) were spatially patterned ($P < 0.05$ for 10 distance classes [$P < 0.05/10$], for both correlograms). Indeed, according to both data sets, there was a clear longitudinal pattern, with the highest PCoA scores derived from the environmental data being predominantly found in the Central region of the United States. Nutrients (K, TN, TP, and NH_4), suspended solids, conductivity, turbidity, water temperature, and alkalinity were the variables that were most strongly and positively correlated with the first PCoA axis (Appendix B: Table B1); accordingly, the values of these variables increased toward the

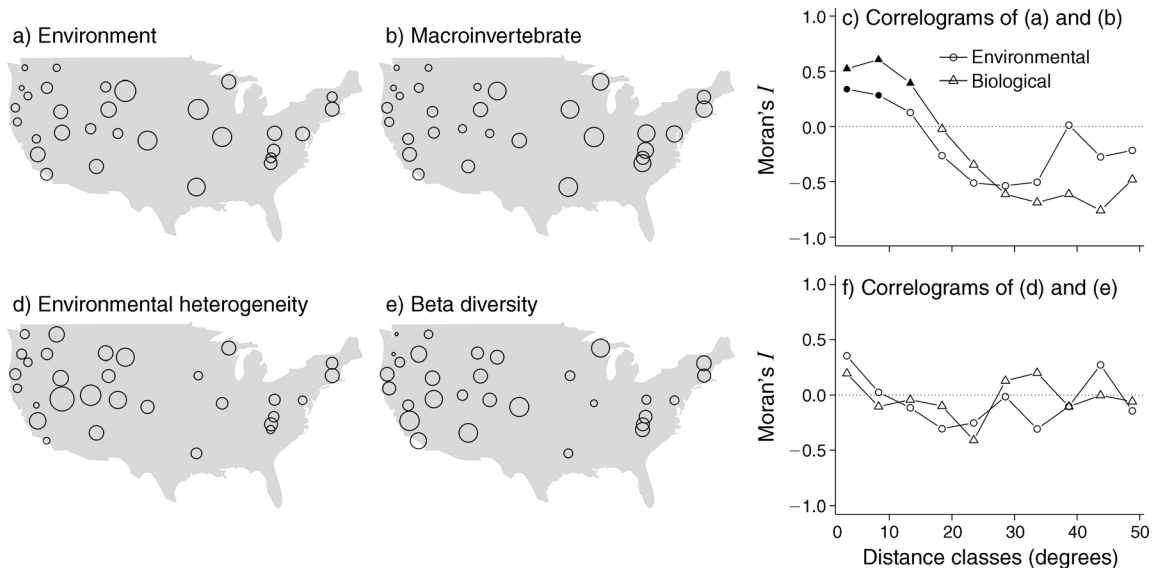


FIG. 1. Spatial variation of the average continental U.S. ecoregion scores along the first axis of a principal coordinate analysis (PCoA) applied to the (a) environmental and (b) macroinvertebrate data sets. The latitudinal and longitudinal midpoints of Omernik's (1995) ecoregions (with more than 10 streams) are presented. The size of the symbols is proportional to the average axis scores. (c) Moran's I correlograms of the axis scores shown in panels (a) and (b). The average distance of each stream to its ecoregion centroid (from PERMDISP analysis), indicating the (d) environmental heterogeneity and (e) beta diversity within ecoregions, and (f) their respective correlograms are also provided. Significant Moran's I coefficients ($P < 0.05$) are indicated by solid symbols.

Central ecoregions (with the highest axis 1 scores). Similarly, the highest scores from the PCoA applied to the biological data were mainly found in the Central regions of the United States (Fig. 1b). After adjusting for spatial autocorrelation, the average scores for the ecoregions on the first axis of the PCoA applied to the biological data were significantly correlated with the scores derived from the PCoA applied to the environmental data (Appendix B: Fig. B1).

The magnitude of environmental heterogeneity (PERMDISP results, $F = 4.71$; $P < 0.001$; Fig. 1d) and beta diversity (Appendix B: Table B2; Fig. 1e) differed significantly among ecoregions, but these variables were not spatially patterned (Fig. 1f; $P > 0.05$ for both correlograms).

We did not detect significant relationships between our measure of beta diversity (d_B) and spatial extent (d_S) or the different measures of environmental heterogeneity (d_{EH} , d_{strm} , d_{subs} , d_{WQ} , d_{WC} ; Fig. 2; results for d_{strm} , d_{subs} , d_{WQ} , and d_{WC} were omitted for brevity). However, we detected a quadratic relationship (hump shaped) between d_B and the first environmental axis derived from the PCoA (Fig. 2). Thus, higher values of d_B were found in the middle of the environmental gradient represented by this axis (mainly associated with nutrient concentrations; see Appendix B: Table B1). The second environmental axis was not significantly correlated with d_B . The results derived from the application of the Sørensen and Simpson distances were similar to those described above (Fig. 2).

The matrix correlation tests (Appendix B: Fig. B2) and generalized linear models (Appendix B: Fig. B3)

indicated that the biological distance matrices were mainly correlated with environmental distances rather than with geographical distances. The strengths of the correlations between biological and environmental distances were independent of spatial extent (d_S) and environmental heterogeneity (d_{EH}) in a specific ecoregion; however, they were significantly and negatively correlated with the first-axis scores from the PCoA (Fig. 3 and Appendix B: Fig. B4). The strengths of the correlations between biological and geographical distances were unrelated to all explanatory variables (Fig. 3). These results were obtained independently of the measure of the effect size used (i.e., matrix correlations or regression coefficients derived from GLM based on the Bray-Curtis, Sørensen, and Simpson coefficients) and how heterogeneity was calculated (results not shown).

Our results are unlikely to be influenced by a violation of ordinary-least-squares assumptions, as the residuals were independent (Appendix C: Fig. C1) and normally distributed (Appendix C: Table C1). Additionally, the variance inflation factors (VIF) were always lower than 1.7, which rules out multicollinearity problems (Appendix C: Table C2). Finally, after removing a few outliers, the patterns described above were even clearer.

DISCUSSION

Predicting variation in beta diversity

Our results indicate that ecoregions with significantly lower or higher environmental and biological variability,

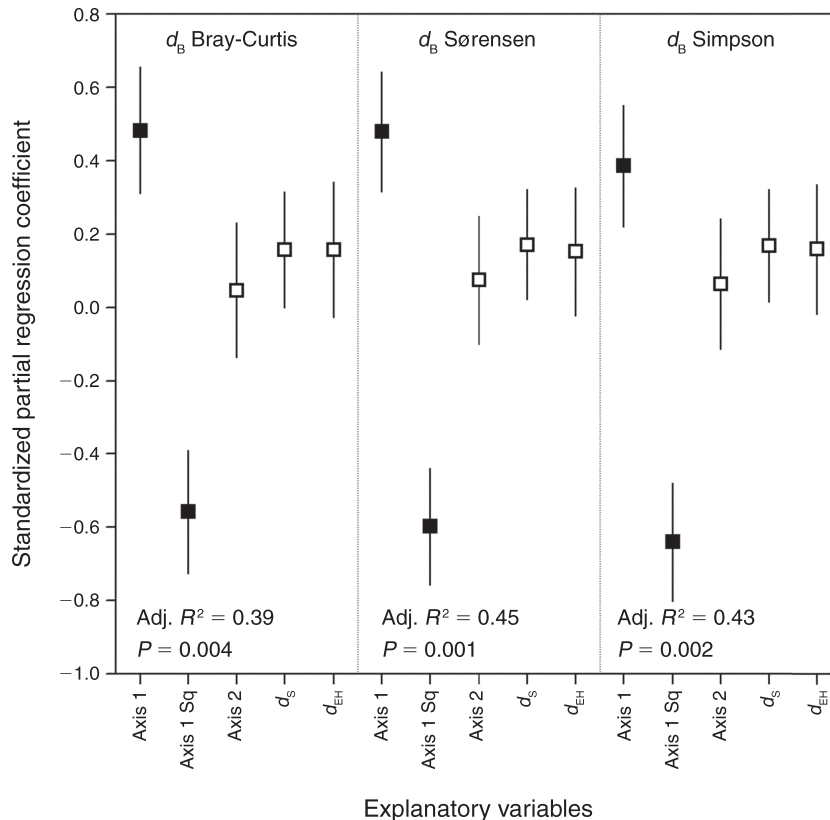


FIG. 2. Standardized partial regression coefficients (\pm SE) from models regressing beta diversity (d_b ; average distance from a stream to its ecoregion centroid) against the first two principal coordinate analysis (PCoA) axis scores (Axis 1, Axis 2; “Axis 1 Sq” means that the square of the first PCoA axis was also included as an explanatory variable), spatial extent (d_s), and environmental heterogeneity (d_{EH}). The adjusted R^2 and P values of the overall regression models, including all five explanatory variables, are provided; the solid squares indicate significant coefficients ($P < 0.05$).

in terms of the variables used here, can be found anywhere in the continental United States. However, the results did not support our hypothesis of a positive relationship between within-ecoregion beta diversity and environmental heterogeneity or spatial extent. Significant results would support the idea that environmental heterogeneity, by creating niche opportunities, would increase species coexistence (Cardinale 2011) and variation in species composition within a given region (Ellingsen and Gray 2002). A relationship with spatial extent, after accounting for environmental heterogeneity, would be consistent with the hypothesis of a decreasing rate of dispersal with geographic distance (Thompson and Townsend 2006). In this context, there are studies showing that beta diversity was strongly related to components of environmental heterogeneity. For instance, beta diversity was found to increase with increasing environmental heterogeneity in kelp holdfast assemblages in New Zealand (Anderson et al. 2006). Likewise, McKnight et al. (2007) found a positive relationship between beta diversity and topographic heterogeneity. Our results display, therefore, some striking differences with these studies that were also able to map beta diversity.

There are also reports describing a lack of or a weak relationship between beta diversity and environmental heterogeneity (Heino et al. 2013), indicating that other factors and processes may also control beta diversity. For example, Chase (2010) showed a positive relationship between productivity and beta diversity and attributed this relationship to a strong role of stochastic processes of community assembly in more-productive ecosystems. Similarly, Langenheder et al. (2012) found that productivity influenced beta diversity more strongly than did environmental heterogeneity in a bacterial metacommunity. Thus, despite confounding factors, our results support the idea that potential productivity (as suggested by nutrient concentrations and temperature) is an important driver of beta diversity. More importantly, however, our findings also suggest that the prevalence of stochastic assembly processes in warmer and more nutrient-rich environments, leading to an increase in beta diversity (Chase 2010), may prevail even at a continental scale. As detailed below (*Predicting the strength of relationships*), this inference is strongly supported by the fact that the strengths of the relationships between compositional dissimilarities and environmental distances (demonstrating deterministic

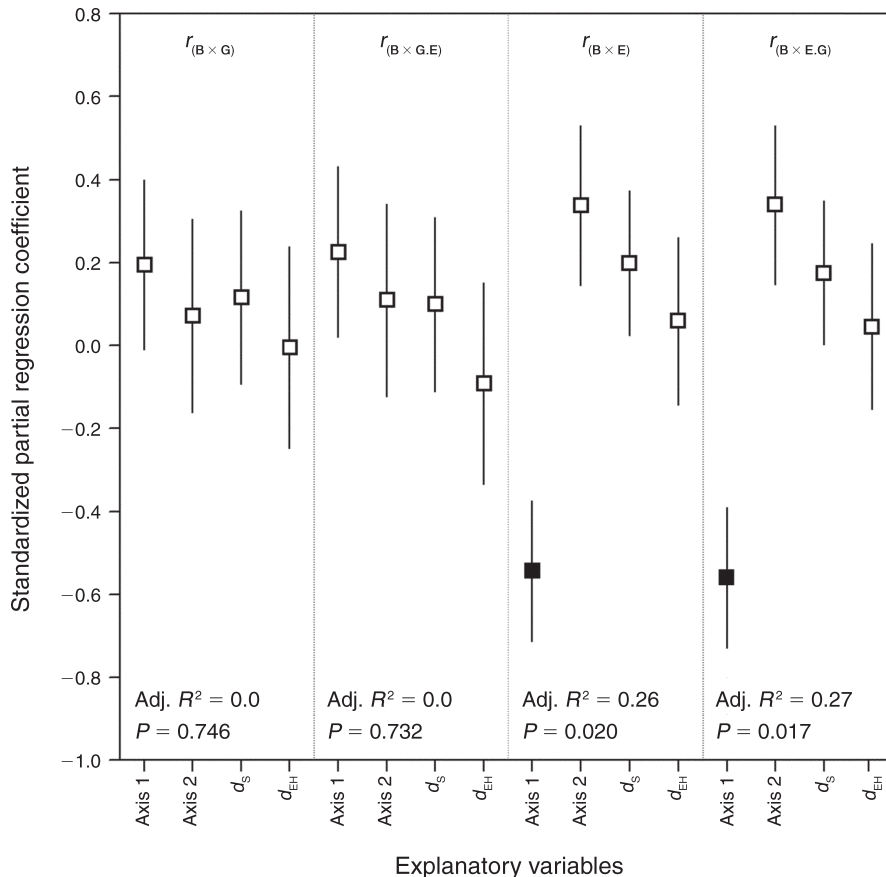


FIG. 3. Standardized partial regression coefficients (\pm SE) from models of effect sizes regressed against the first two PCoA axis scores, spatial extent (d_s), and environmental heterogeneity (d_{EH}). The effect sizes, as response variables, include the simple matrix correlations (r) between community dissimilarities (Bray-Curtis coefficient, **B**) and geographical distances (**G**) or environmental distances (standardized Euclidean, **E**): $r_{(B \times G)}$ and $r_{(B \times E)}$, respectively. The partial matrix correlations for which either **E** or **G** was held constant are also shown: $r_{(B \times G, E)}$ and $r_{(B \times E, G)}$, respectively. The solid squares indicate significant coefficients ($P < 0.05$).

processes of community assembly) are consistently lower in more nutrient-rich and warmer ecoregions.

The relative importance of environmental and spatial predictors

We detected a strong relationship between assemblage structure and environmental conditions (Appendix B: Fig. B1). This pattern was mainly driven by the highest nutrient concentrations and sediment loads found in the Central U.S. ecoregions (see Paulsen et al. [2008b] for a comprehensive evaluation of the effects of these stressors on stream macroinvertebrates). We also found that compositional dissimilarities were mainly correlated with environmental distances rather than with geographic distances, independent of the analytical approach (i.e., simple and partial Mantel's tests or slopes obtained from GLMs). In addition to reinforcing the validity of stream invertebrates as reliable bioindicators of stream health at a broad scale, our results, using either a raw-data approach (Appendix B: Fig. B1) or a distance-based approach (Appendix B: Figs. B2 and B3),

strongly emphasize “the power of species sorting” (Van der Gucht et al. 2007). It is interesting to note that results from previous analyses of the comprehensive US EPA data set, which focused on different response variables (e.g., O/E index from RIVPACS-type models), also support species-sorting mechanisms as the primary drivers of macroinvertebrate assemblage structure (e.g., Carlisle and Hawkins 2008, Ode et al. 2008, Hawkins et al. 2010).

Despite the higher predictive power of environmental variables compared with geographical ones, as commonly found in aquatic community studies (Cottenie 2005, Van der Gucht et al. 2007), the strengths of the relationships between the response and predictor matrices were highly variable. For instance, the partial Mantel correlations between the Bray-Curtis and environmental distances ranged from -0.04 to 0.45 (Appendix B: Fig. B4). With this substantial variability, our focus should shift from defining a generally applicable metacommunity model to understanding the cause of this variability itself, which we address below.

Predicting the strength of relationships

Heterogeneous conditions provide suitable environments for different species (Leibold et al. 2004). Thus, assuming a predominant role of species-sorting mechanisms driving beta diversity variation, the strength of the relationship between compositional dissimilarities and environmental distances is predicted to increase with increasing environmental heterogeneity. On the other hand, beta diversity can also be driven by dispersal limitation. In this case, in a cross-region analysis, spatial extent would be correlated with beta diversity because increasing distances among sites would increase dispersal limitation that, in turn, should increase the differences in species composition between distant sites (Hubbell 2001, Shurin et al. 2009). This is equivalent to stating that the relative importance of species-sorting processes and spatial processes should increase in more environmentally heterogeneous ecoregions and in ecoregions with larger spatial extents, respectively. However, we obtained no support for these predictions, despite the fact that there was a significant relationship between the biological and environmental distances within most of the ecoregions. It is important to emphasize that we observed no significant relationship between environmental heterogeneity and the strength of correlation between compositional dissimilarities and environmental distances, even after allowing for more specific measurements of environmental heterogeneity (i.e., d_{strm} , d_{subs} , d_{WQ} , and d_{WC} instead of d_{EH}). Thus, it is unlikely that the lack of significant relationships between the strength of the matrix correlation and environmental heterogeneity can be explained by the composite nature of the measurements of environmental heterogeneity.

Previous studies attempting to identify correlates of the relative importance of species-sorting effects and dispersal limitation were based on heterogeneous data sets (e.g., examining different groups of organisms and regions; e.g., Cottenie 2005, Soininen et al. 2007, De Bie et al. 2012). Thus, comparisons with previous results are not straightforward. Even so, contrary to the results of a meta-analysis (Soininen et al. 2007), we did not detect a significant and positive relationship between spatial extent and the effect sizes indicating the strengths of species-sorting or spatial processes. Instead, we found that nutrient concentration was a significant correlate of the strength of the relationship between biological and environmental distances. In general, the matrix correlations and slopes derived from GLMs decreased with increases in nutrient concentrations. This result strongly suggests a large role for stochastic community assembly processes in more nutrient-rich environments and that species-sorting mechanisms prevail in nutrient-poor environments (Chase 2010; but see next paragraph). However, in marked contrast with our results, Langenheder et al. (2012) found that species sorting was more frequent during periods of high productivity, environmental heterogeneity, and only when beta-diversity was relatively high.

Chase (2010) also observed a strong role of species sorting in his lowest productivity treatments. However, assemblages with the lowest species richness were found in these treatments, whereas richer assemblages, where stochastic assembly processes prevailed, were found in the most highly productive treatments. In contrast, we detected a strong role of species sorting in ecoregions with richer macroinvertebrate assemblages. This difference can most likely be accounted for by the fact that the nutrient gradients in running waters in the United States result partially from eutrophication caused by agriculture and urbanization (Paulsen et al. 2008b, Van Sickle and Paulsen 2008), and this eutrophication gradient is positively associated with the effects of other stressors (e.g., sediment loading) that can severely impair macroinvertebrate assemblages (Carlisle and Hawkins 2008; but see also Olson and Hawkins [2013] who reported large natural variation in nutrient concentrations). Indeed, we found a negative and significant correlation between average species richness and our first PCoA axis summarizing the environmental data ($r = -0.73$; $P = 0.009$; after accounting for spatial autocorrelation according to Clifford et al. [1989]). Thus, our results suggest that eutrophication is a disassembly mechanism, decoupling the relationship between compositional dissimilarities and environmental distances. In a related context, Gutiérrez-Cánovas et al. (2013), analyzing five benthic macroinvertebrates data sets from Great Britain, Iberian Peninsula, and Himalayan Mountains, found that the nestedness component of dissimilarities in community composition was mainly correlated with environmental distance matrices based on anthropogenic stressors (acidity, metals, land use). On the other hand, the turnover component was correlated with environmental distance matrices based on natural gradients (elevation and salinity). It is not possible, however, to disentangle the effects of an increase in nutrient concentrations per se and other stressors accompanying eutrophication on this mechanism of disassembly. In any case, our results clearly indicate a hitherto-unrecognized consequence of eutrophication pertaining to community disassembly.

Conclusions

Using a unique broad-scale data set of stream invertebrate assemblages across the United States, we showed that within-ecoregion compositional dissimilarity matrices were mainly related to environmental distances in most of the ecoregions studied. This finding strongly suggests a more important role for species-sorting than dispersal-driven processes in determining the composition of stream invertebrate assemblages. The strengths of these compositional dissimilarity–environmental distance relationships varied considerably among ecoregions. They were unrelated to within-ecoregion environmental heterogeneity or spatial extent. Instead, we detected a negative correlation between the strength of species sorting and nutrient levels across the

ecoregions, which we interpret as a disassembling mechanism. Further studies should (1) aim to determine breakpoints in nutrient levels at which beta diversity begins to decrease due to a human-caused eutrophication process and (2) disentangle the effects of eutrophication from other anthropogenic stressors on biodiversity. The US EPA data set can be further examined to accomplish these goals.

ACKNOWLEDGMENTS

The Academy of Finland (J. Heino), CNPq and CAPES (L. M. Bini, V. L. Landeiro, A. A. Padial, and T. Siqueira) have continuously supported our research group in community ecology. We acknowledge US EPA for providing access to the Wadeable Stream Assessment data set and Bradford A. Hawkins (University of California) for valuable suggestions and discussions. Comments and suggestions from Charles P. Hawkins (Utah State University) and an anonymous reviewer were pivotal to improve the quality of a previous version of this paper.

LITERATURE CITED

- Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26:32–46.
- Anderson, M. J., K. E. Ellingsen, and B. H. McCaule. 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9:683–693.
- Anderson, M. J., et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters* 14:19–28.
- Beck, J., et al. 2012. What's on the horizon for macroecology? *Ecography* 35:673–683.
- Bivand, R. 2013. *spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.5-65. <http://CRAN.R-project.org/package=spdep>
- Bonn, A., D. Storch, and K. J. Gaston. 2004. Structure of the species-energy relationship. *Proceedings of the Royal Society of London B*: 271:1685–1691.
- Cao, Y., and C. P. Hawkins. 2005. Simulating biological impairment to evaluate the accuracy of ecological indicators. *Journal of Applied Ecology* 42:954–965.
- Cardinale, B. J. 2011. Biodiversity improves water quality through niche partitioning. *Nature* 472:86–89.
- Carlisle, D. M., and C. P. Hawkins. 2008. Land use and the structure of western US stream invertebrate assemblages: predictive models and ecological traits. *Journal of the North American Benthological Society* 27:986–999.
- Carlisle, D. M., C. P. Hawkins, M. R. Meador, M. Potapova, and J. Falcone. 2008. Biological assessments of Appalachian streams based on predictive models for fish, macroinvertebrate, and diatom assemblages. *Journal of the North American Benthological Society* 27:16–37.
- Chase, J. M. 2010. Stochastic community assembly causes higher biodiversity in more productive environments. *Science* 328:1388–1391.
- Chinnayakanahalli, K. J., C. P. Hawkins, D. G. Tarboton, and R. A. Hill. 2011. Natural flow regime, temperature and the composition and richness of invertebrate assemblages in streams of the western United States. *Freshwater Biology* 56:1248–1265.
- Clifford, P., S. Richardson, and D. Hémon. 1989. Assessing the significance of the correlation between two spatial processes. *Biometrics* 45:123–134.
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* 8:1175–1182.
- De Bie, T., et al. 2012. Body size and dispersal mode as key traits determining metacommunity structure of aquatic organisms. *Ecology Letters* 15:740–747.
- Ellingsen, K., and J. S. Gray. 2002. Spatial patterns of benthic diversity: Is there a latitudinal gradient along the Norwegian continental shelf? *Journal of Animal Ecology* 71:373–389.
- Gutiérrez-Cánovas, C., A. Millán, J. Velasco, I. P. Vaughan, and S. J. Ormerod. 2013. Contrasting effects of natural and anthropogenic stressors on beta diversity in river organisms. *Global Ecology and Biogeography* 22:796–805.
- Hawkins, C. P., Y. Cao, and B. Roper. 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. *Freshwater Biology* 55:1066–1085.
- Hawkins, C. P., S. G. Paulsen, J. Van Sickle, and L. L. Yuan. 2008. Regional assessments of stream ecological condition: scientific challenges associated with the USA's national Wadeable Stream Assessment. *Journal of the North American Benthological Society* 27:805–807.
- Heino, J. 2008. Influence of taxonomic resolution and data transformation on biotic matrix concordance and assemblage-environment relationships in stream macroinvertebrates. *Boreal Environmental Research* 13:359–369.
- Heino, J., M. Grönroos, J. Ilmonen, T. Karhu, M. Niva, and L. Paasivirta. 2013. Environmental heterogeneity and β diversity of stream macroinvertebrate communities at intermediate spatial scales. *Freshwater Science* 32:142–154.
- Herlihy, A. T., S. G. Paulsen, J. Van Sickle, J. L. Stoddard, C. P. Hawkins, and L. L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. *Journal of the North American Benthological Society* 27:860–877.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, New Jersey, USA.
- Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology* 72:367–382.
- Langenheder, S., M. Berga, Ö. Östman, and A. J. Székely. 2012. Temporal variation of β -diversity and assembly mechanisms in a bacterial metacommunity. *The ISME Journal* 6:1107–1114.
- Legendre, P., and L. Legendre. 1998. *Numerical ecology*. Second edition. Elsevier, Amsterdam, The Netherlands.
- Leibold, M. A., et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7:601–613.
- Lenat, D. R., and V. H. Resh. 2001. Taxonomy and stream ecology: the benefits of genus-and species-level identifications. *Journal of the North American Benthological Society* 20:287–298.
- McKnight, M. W., P. S. White, R. I. McDonald, J. F. Lamoreux, W. Sechrest, R. S. Ridgely, and S. N. Stuart. 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biology* 5:e272.
- Millar, R., M. Anderson, and N. Tolimieri. 2011. Much ado about nothings: using zero similarity points in distance-decay curves. *Ecology* 92:1717–1722.
- Ode, P. R., C. P. Hawkins, and R. D. Mazor. 2008. Comparability of biological assessments derived from predictive models and multimetric indices of increasing geographic scope. *Journal of North American Benthological Society* 27:967–985.
- Oden, N. L. 1984. Assessing the significance of a spatial correlogram. *Geographical Analysis* 16:1–16.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. *vegan: Community*

- Ecology Package. R package version 2.0-8. <http://CRAN.R-project.org/package=vegan>
- Olson, J. R., and C. P. Hawkins. 2013. Developing site-specific nutrient criteria from empirical models. *Freshwater Science* 32:719–740.
- Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49–65 in W. S. Davis and T. P. Simon, editors. *Biological assessment and criteria*. Lewis Publishers, London, UK.
- Paulsen, S. G., C. P. Hawkins, J. Van Sickle, L. L. Yuan, and S. M. Holdsworth. 2008a. An invitation to apply national survey data to ecological research. *Journal of the North American Benthological Society* 27:1017–1018.
- Paulsen, S. G., et al. 2008b. Condition of stream ecosystems in the US: an overview of the first national assessment. *Journal of the North American Benthological Society* 27:812–821.
- R Core Development Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Shurin, J. B., K. Cottenie, and H. Hillebrand. 2009. Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia* 159:151–159.
- Soininen, J., R. McDonald, and H. Hillebrand. 2007. The distance decay of similarity in ecological communities. *Ecography* 30:3–12.
- Thompson, R., and C. Townsend. 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.
- USEPA. 2004. Wadeable stream assessment: field operations manual. EPA841-B-04-004. U.S. Environmental Protection Agency, Office of Water and Office of Research and Development, Washington, D.C., USA.
- Van der Gucht, K., et al. 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 Sickle, J., and S. G. Paulsen. 2008. Assessing the attributable risks, relative risks, and regional extents of aquatic stressors. *Journal of the North American Benthological Society* 27:920–931.
- Veech, J. A., and T. O. Crist. 2007. Habitat and climate heterogeneity maintain beta-diversity of birds among landscapes within ecoregions. *Global Ecology and Biogeography* 16:650–656.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs* 30:279–338.
- Whittaker, R. J. 2010. Meta-analyses and mega-mistakes: calling time on meta-analysis of the species richness–productivity relationship. *Ecology* 91:2522–2533.
- Wiggins, G. B., and R. J. Mackay. 1978. Some relationships between systematics and trophic ecology in Nearctic aquatic insects, with special reference to Trichoptera. *Ecology* 59:1211–1220.

SUPPLEMENTAL MATERIAL

Appendix A

Schematic representation of the main analyses used in this study ([Ecological Archives E095-138-A1](#)).

Appendix B

Loadings from PCA, PERMANOVA, and PERMDISP results, scatterplots, and results from GLM and matrix comparisons ([Ecological Archives E095-138-A2](#)).

Appendix C

Tests of models assumptions ([Ecological Archives E095-138-A3](#)).