

## A comparative analysis reveals weak relationships between ecological factors and beta diversity of stream insect metacommunities at two spatial levels

Jani Heino<sup>1,\*</sup>, Adriano S. Melo<sup>2,\*</sup>, Luis Mauricio Bini<sup>2,\*</sup>, Florian Altermatt<sup>3,4</sup>, Salman A. Al-Shami<sup>5,6</sup>, David G. Angeler<sup>7</sup>, Núria Bonada<sup>8</sup>, Cecilia Brand<sup>9</sup>, Marcos Callisto<sup>10</sup>, Karl Cottenie<sup>11</sup>, Olivier Dangles<sup>12,13</sup>, David Dudgeon<sup>14</sup>, Andrea Encalada<sup>15</sup>, Emma Göthe<sup>16</sup>, Mira Grönroos<sup>1</sup>, Neusa Hamada<sup>17</sup>, Dean Jacobsen<sup>18</sup>, Victor L. Landeiro<sup>19</sup>, Raphael Ligeiro<sup>10</sup>, Renato T. Martins<sup>17</sup>, María Laura Miserendino<sup>9</sup>, Che Salmah Md Rawi<sup>5</sup>, Marciel E. Rodrigues<sup>20</sup>, Fabio de Oliveira Roque<sup>20</sup>, Leonard Sandin<sup>7</sup>, Denes Schmera<sup>21,22</sup>, Luciano F. Sgarbi<sup>2</sup>, John P. Simaika<sup>23</sup>, Tadeu Siqueira<sup>24</sup>, Ross M. Thompson<sup>25</sup> & Colin R. Townsend<sup>26</sup>

<sup>1</sup>Finnish Environment Institute, Natural Environment Centre, Biodiversity, Oulu, Finland

<sup>2</sup>Departamento de Ecologia, Universidade Federal de Goiás, Goiânia, GO, Brazil

<sup>3</sup>Department of Aquatic Ecology, Eawag: Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

<sup>4</sup>Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Zürich, Switzerland

<sup>5</sup>School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia

<sup>6</sup>Biology Department, Faculty of Science, University of Tabuk, Tabuk, Saudi Arabia

<sup>7</sup>Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden

<sup>8</sup>Departament d'Ecologia, Grup de Recerca Freshwater Ecology and Management (FEM), Universitat de Barcelona, Barcelona, Catalonia, Spain

<sup>9</sup>LIESA-CONICET-Universidad Nacional de la Patagonia SJB, Chubut, Argentina

<sup>10</sup>Departamento de Biologia Geral, Instituto de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil

<sup>11</sup>Department of Integrative Biology, University of Guelph, Guelph, ON, Canada

<sup>12</sup>Laboratory of Entomology, School of Biological Sciences, Pontifical Catholic University of Ecuador, Quito, Ecuador

<sup>13</sup>IRD, Institut de Recherche pour le Développement, Laboratoire Evolution, Génomes et Spéciation, Gif-sur-Yvette, France

<sup>14</sup>School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China

<sup>15</sup>Laboratorio de Ecología Acuática Colegio de Ciencias Biológicas y Ambientales Universidad San Francisco de Quito, Quito, Ecuador

<sup>16</sup>Department of Bioscience, Aarhus University, Silkeborg, Denmark

<sup>17</sup>Instituto Nacional de Pesquisas da Amazônia, Coordenação de Biodiversidade, Manaus, AM, Brazil

<sup>18</sup>Department of Biology, University of Copenhagen, Copenhagen, Denmark

<sup>19</sup>Departamento de Botânica e Ecologia, Universidade Federal do Mato Grosso, Cuiabá, Brazil

<sup>20</sup>Departamento de Ciências Biológicas e da Saúde, Universidade Federal de Mato Grosso do Sul, Campo Grande, Mato Grosso do Sul, Brazil

<sup>21</sup>Section of Conservation Biology, Department of Environmental Sciences, University of Basel, Basel, Switzerland

<sup>22</sup>Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences, Tihany, Hungary

<sup>23</sup>Department of Conservation Ecology and Entomology, Stellenbosch University, Stellenbosch, South Africa

<sup>24</sup>Instituto de Biotecnologia, UNESP - Universidade Estadual Paulista, Rio Claro, São Paulo, Brazil

<sup>25</sup>Institute for Applied Ecology, University of Canberra, Canberra, ACT, Australia

<sup>26</sup>Department of Zoology, University of Otago, Dunedin, New Zealand

### Keywords

Altitude range, comparative analysis, environmental filtering, insects, latitude, spatial extent, variance partitioning.

### Correspondence

Jani Heino, Finnish Environment Institute, Natural Environment Centre, Biodiversity P.O. Box 413, FI-90014 Oulu, Finland.  
Tel: +358 295 251 157;  
E-mail: jani.heino@environment.fi

### Funding Information

Our research on community ecology and running waters has been continuously supported by the Academy of Finland and

### Abstract

The hypotheses that beta diversity should increase with decreasing latitude and increase with spatial extent of a region have rarely been tested based on a comparative analysis of multiple datasets, and no such study has focused on stream insects. We first assessed how well variability in beta diversity of stream insect metacommunities is predicted by insect group, latitude, spatial extent, altitudinal range, and dataset properties *across multiple drainage basins* throughout the world. Second, we assessed the relative roles of environmental and spatial factors in driving variation in assemblage composition *within each drainage basin*. Our analyses were based on a dataset of 95 stream insect metacommunities from 31 drainage basins distributed around the world. We used dissimilarity-based indices to quantify beta diversity for each metacommunity and, subsequently, regressed beta diversity on insect group, latitude, spatial extent, altitudinal range, and dataset properties (e.g., number of sites and percentage of

the Brazilian Council of Science and Technology CNPq.

Received: 3 October 2014; Revised: 26 January 2015; Accepted: 27 January 2015

doi: 10.1002/ece3.1439

\*The first three authors contributed equally to this study.

presences). Within each metacommunity, we used a combination of spatial eigenfunction analyses and partial redundancy analysis to partition variation in assemblage structure into environmental, shared, spatial, and unexplained fractions. We found that dataset properties were more important predictors of beta diversity than ecological and geographical factors across multiple drainage basins. In the within-basin analyses, environmental and spatial variables were generally poor predictors of variation in assemblage composition. Our results revealed deviation from general biodiversity patterns because beta diversity did not show the expected decreasing trend with latitude. Our results also call for reconsideration of just how predictable stream assemblages are along ecological gradients, with implications for environmental assessment and conservation decisions. Our findings may also be applicable to other dynamic systems where predictability is low.

## Introduction

The importance of understanding broad-scale patterns of biodiversity is becoming ever more urgent as landscapes, ecosystems, and communities are increasingly transformed by processes such as habitat loss, invasion by exotic species, eutrophication, and climate change (Dudgeon *et al.* 2006; Vörösmarty *et al.* 2010). Increased understanding of broad-scale biodiversity patterns is a prerequisite for the advancement of both basic and applied ecology. From the perspective of basic ecology and for the sake of generality, more information on biodiversity patterns of a variety of under-studied taxa is needed (Leather 2009). From the perspective of applied ecology, knowledge of patterns in biodiversity is essential to guide both conservation actions and environmental assessment (Caro 2010). Accordingly, better understanding of patterns in regional species richness ( $\gamma$ -diversity), local species richness ( $\alpha$ -diversity), and variation in species composition between sites ( $\beta$ -diversity) at the global, continental, and regional scales would be valuable (Heino 2013; Bini *et al.* 2014).

Hypotheses about the processes driving beta diversity are closely intertwined with recent developments in metacommunity theory (Leibold *et al.* 2004; Logue *et al.* 2011). While variation in local community composition is thought to be typically driven by species sorting along environmental gradients (Cottenie 2005), spatial processes (e.g., dispersal between sites) also have the potential to affect local community composition (Brown and Swan 2010; Altermatt *et al.* 2013; Heino and Peckarsky 2014; Heino *et al.* 2015a). The relative importance of species sorting versus spatial processes may be contingent on the lengths of environmental gradients (e.g., range in stream temperature or nutrient concentrations within a drainage basin) and the spatial extent of the study (Jackson *et al.* 2001; Cottenie 2005; Bini *et al.* 2014; Heino *et al.* 2015a). One might expect species sorting to increase with increas-

ing environmental gradient length (Jackson *et al.* 2001; Grönroos *et al.* 2013) and spatial factors to gain importance with increasing spatial extent of the region studied (Cottenie 2005; Heino 2011). Very few studies have explicitly tested these hypotheses using multiple datasets, which is surprising given that environmental heterogeneity and spatial scale are two key ideas behind metacommunity theory (Leibold *et al.* 2004). While the importance of species sorting and spatial factors for metacommunity organization can be addressed in single case studies, such studies do not allow the identification of robust patterns and cannot lead to broad generalizations. Instead, the relative influence of environmental heterogeneity (cf. species sorting) and spatial extent (cf. spatial factors) can only be investigated in a general sense across multiple metacommunities in different geographic regions using a comparative approach. Such broad-scale analyses have seldom involved stream organisms, and the analyses that have been undertaken have either focused on a small component of benthic insect assemblages (Boyer *et al.* 2012) or on specific stream types (Jacobsen and Dangles 2012).

Running waters offer an ideal model system to disentangle the relative influences of species sorting and spatial processes on metacommunity organization (Brown *et al.* 2011; Heino *et al.* 2015b). Stream ecosystems show high geomorphological heterogeneity (Allan and Castillo 2007), are structured in dendritic networks (Altermatt 2013), and harbor exceptional levels of biodiversity relative to their limited spatial extent (Dudgeon *et al.* 2006). Aquatic insects are prominent organisms in streams, playing key roles in food webs and ecosystem processes (Allan and Castillo 2007) and showing high diversity in terms of phylogenetic origins, dispersal traits, species richness, and endemism (Lancaster and Downes 2013). However, they are assumed to be strongly impacted by various anthropogenic factors, including pollution (Rosenberg and Resh 1993), habitat modification (Allan and Castillo 2007), and

climate change (Jacobsen *et al.* 2012). This sensitivity potentially makes stream insects valuable as biological indicators (Rosenberg and Resh 1993) and, hence, understanding patterns and scales of variability in their metacommunities is a priority (Heino 2013).

Current knowledge on stream insect metacommunities mainly relies on case studies conducted in individual regions (e.g., Thompson and Townsend 2006; Heino and Mykrä 2008; Brown and Swan 2010; Landeiro *et al.* 2012; Siqueira *et al.* 2012; Al-Shami *et al.* 2013a). Most studies have shown that species sorting is prevalent in shaping these metacommunities, although the strength of dispersal limitation may increase with increasing spatial scale (Heino and Peckarsky 2014). The mechanisms structuring metacommunities may also be contingent on system-specific factors (e.g., high dispersal rates may be important in mainstem rivers; Brown and Swan 2010) or be related to dispersal traits (e.g., the importance of dispersal limitation increases with decreasing dispersal ability; Thompson and Townsend 2006; Heino *et al.* 2015b). These and other hypotheses associated with broad-scale patterns are difficult to test with a single or a few datasets, and more general perspectives may be obtained through a comparative analysis of multiple datasets. Using a large number of stream insect datasets from different parts of the world, we attempted to reveal the main factors structuring stream insect metacommunities by answering the following questions: (Q1) How well can variability in beta diversity *across* multiple stream insect metacommunities (i.e., spatial level 1; Appendix S1) be accounted for by insect group, latitude, altitudinal range, and spatial extent? (Q2) What are the relative roles of species sorting

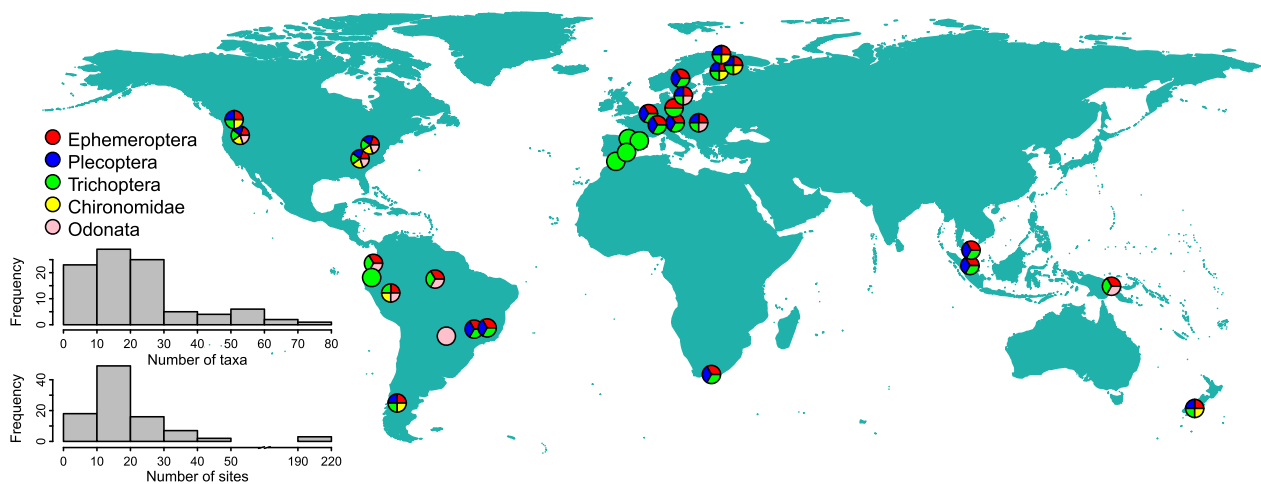
and spatial processes in predicting assemblage composition *within* stream insect metacommunities (i.e., spatial level 2; Appendix S2)?

We assembled a comprehensive dataset comprising 95 stream insect metacommunities from different regions, ranging from tropical to Arctic latitudes and lowland to montane habitats. First, we analyzed spatial variation in beta diversity, with the expectation of higher beta diversity in tropical regions than nearer the poles (Soininen *et al.* 2007; but see Boyero *et al.* 2011). Second, we tested whether the importance of environmental factors in accounting for variation in assemblage composition tended to be greater within drainage basins that had higher environmental variability (Heino 2011; Grönroos *et al.* 2013), and whether spatial factors increased in importance with increasing regional extent (Cottenie 2005; Bini *et al.* 2014). We showed that beta diversity patterns may be poorly predictable in highly dynamic stream systems, which calls for reconsideration of the predictability of ecological communities.

## Methods

### Study areas and datasets

We assembled data for five major stream insect taxa: mayflies (Ephemeroptera), stoneflies (Plecoptera), caddisflies (Trichoptera), nonbiting midges (Diptera: Chironomidae), and dragonflies (Odonata) from 31 drainage basins distributed around the world, but predominantly from the Americas and Europe (Fig. 1). These five insect groups often dominate in stream invertebrate communities



**Figure 1.** Geographical locations of the 31 drainage basins in this study. Analyses were carried out for each insect taxon separately. Thereby, 95 datasets were available for analyses of biological data only and 61 datasets for analyses using environmental predictors. In some cases, symbols have been shifted slightly to avoid overlap. The inset histograms show the frequency distribution of number of species (upper histogram) and number of sites per metacommunity (lower histogram).

(Lancaster and Downes 2013; Heino and Peckarsky 2014), and the family Chironomidae may alone account for a very large share of species in a dataset (e.g., Raunio *et al.* 2011). However, given the lack of data for other important taxa, such as beetles (Coleoptera), our results are only applicable to the five taxonomic groups analyzed. Both insect and environmental data were available for 23 basins, while insect data only were available for the remaining eight basins. We initially intended to use datasets consisting of at least 15 stream sites per basin, but included a few basins with a lower number of sites to increase geographic coverage for across-basins beta diversity analyses (see below). The average number of sites per drainage basin was 24.5 (range: 4 to 213), and most drainage basins included data for at least three insect taxonomic groups (see below) (Fig. 1). Each insect dataset was also checked to ensure that surveys were conducted within a single year and that sampling did not span multiple seasons (i.e., samples within a drainage basin were taken within a short period of time). This was performed to guarantee that we were studying sets of potentially interacting species (see Leibold *et al.* 2004). In addition, each insect dataset was based on standardized sampling methods (e.g., kick-netting or Surber samplers), and between-site differences in species composition were thus comparable within each dataset. We focused on datasets from streams subject to little anthropogenic impact, although in a few basins, some streams were affected by adjacent human land use (e.g., some streams drained managed forests).

Insect data for each drainage basin were separated by taxonomic group for analyses, yielding a total of 95 datasets, of which 75 included both insect and environmental data. However, in the within-basin analyses described below, we used only 61 datasets that included at least 15 sites. We used this approach because each taxonomic group may show distinct latitudinal patterns in diversity (Vinson and Hawkins 2003; Pearson and Boyero 2009), exhibit different dispersal abilities (Bonada *et al.* 2012), or respond differently to environmental factors (Heino and Mykrä 2008). Furthermore, different taxonomic groups may exhibit different patterns of metacommunity organization (Heino and Mykrä 2008; Altermatt *et al.* 2013).

Taxonomic resolution varied among the insect datasets due to differences in taxonomic knowledge among regions; however, all datasets involved identification to at least the level of genus. We also controlled for regional differences in taxonomic knowledge by including the variable “proportion of taxa identified to species level” in analyses comparing variability in beta diversity *across* metacommunities. We had to rely on incidence (i.e., presence–absence) data because we did not have strictly

comparable abundance data from all drainage basins. Although numerical resolution may affect patterns in beta diversity (Anderson *et al.* 2011; Heino *et al.* 2015a), previous studies on stream invertebrates have shown that the main patterns can be reproduced using either presence–absence or abundance data (Al-Shami *et al.* 2013b; Heino *et al.* 2013; Bini *et al.* 2014).

## Environmental variables

The physical and chemical variables that were measured in each drainage basin varied greatly (Appendix S3). We removed some categorical variables from drainage basin datasets where there were fewer sites than explanatory variables. We also pooled fine and coarse estimates of benthic organic matter when both were present in the same dataset. Similarly, for datasets that included proportional coverage of substrate size classes, the finer categories (i.e., silt, sand, gravel, and pebble) were pooled. We also removed variables that provide similar information (e.g., we retained only total nitrogen when several forms of nitrogen were reported), as well as variables only measured at a few sites within a drainage basin and some variables related to riparian vegetation (e.g., width of the riparian vegetation) that affect stream insects indirectly.

After this initial data screening, we finalized environmental datasets for each insect group in each drainage basin. This involved deletion of sites from which a particular insect group was absent and deletion of variables for which observations were only available for a few sites. Finally, we removed entire datasets including <15 sites. The 61 final environmental datasets were derived from 20 basins and included 4–21 variables, and a total of 67 variables were used. Conductivity, total phosphorus, and depth were the most common variables, available for 54, 41, and 40 datasets, respectively (Appendix S3).

## Beta diversity analyses

Multiple approaches are necessary to adequately describe patterns in beta diversity, because each approach may provide distinct information about this facet of biodiversity (Anderson *et al.* 2011). We used three approaches to estimate beta diversity for each of the 95 metacommunities. The first was the average of pairwise dissimilarities between sites within a drainage basin. The second was the multiple-site version of the same metric (Diserud and Ødegaard 2007), which is a generalization of the 2-samples formula to handle more than two samples. The third was the average biological distance of sites within a single metacommunity to the metacommunity centroid (Anderson *et al.* 2006). Each of these three approaches was based on three different dissimilarity coefficients: (i) Sørensen

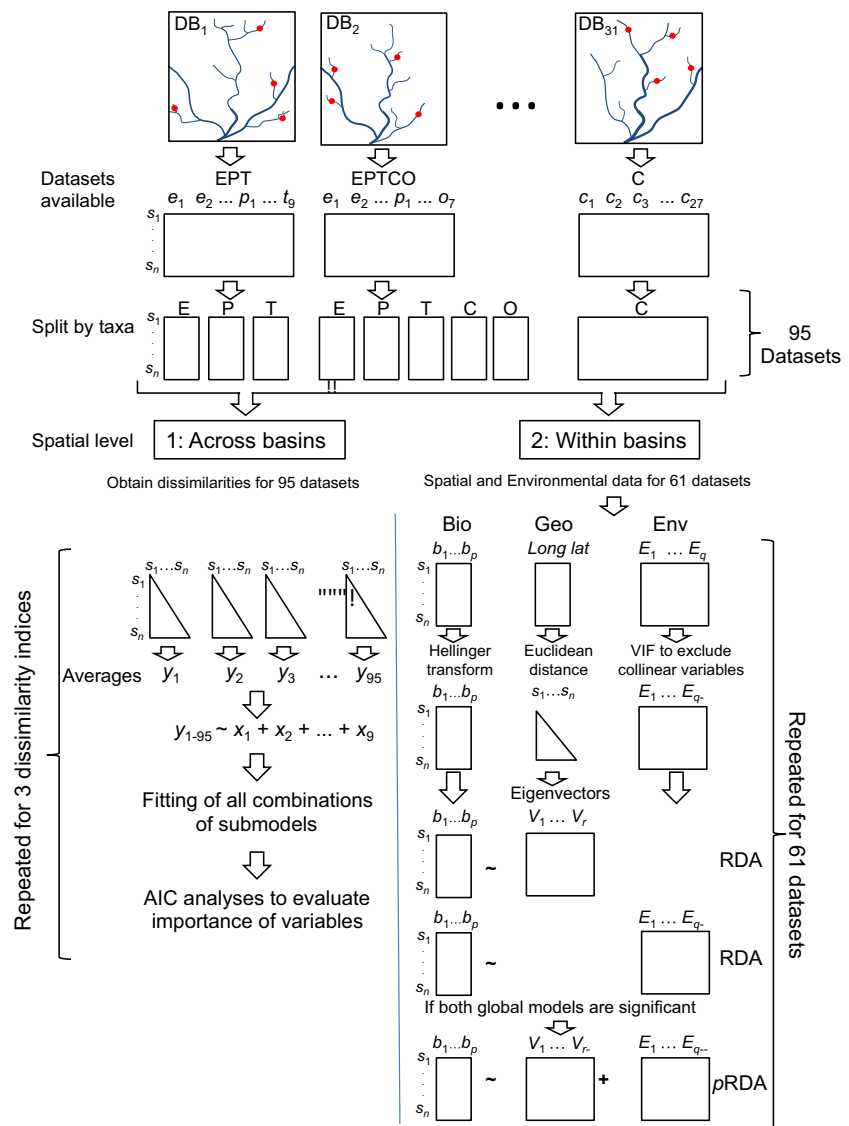
coefficient (i.e., a measure of overall beta diversity), (ii) Simpson coefficient (i.e., a measure of turnover immune to nestedness resulting from species richness differences), and (iii) a coefficient measuring nestedness resulting from species richness differences (Baselga 2010).

## Comparative analyses of beta diversity

We regressed dissimilarity-based beta diversities obtained for each drainage basin ( $n = 95$ ) against (i) taxonomic group (as a dummy variable), (ii) latitude (absolute values), (iii) range in altitude, and (iv) spatial extent of a drainage basin. We also included (v) an interaction term between range in altitude and latitude (i.e., the product of range in altitude and latitude) because the effects of range in altitude may depend on latitude

(Hawkins and Diniz-Filho 2006). In addition to these factors, estimates of beta diversity may be influenced by the properties of the datasets (Podani and Schmera 2011). Thus, we also included in the regression models the following predictor variables: (vi) number of sites in the dataset, (vii) number of species, (viii) matrix fill (percentage of presences), and (ix) percentage of taxa identified to species level. Three datasets included many more stream sites than others (Fig. 1), so we log-transformed this variable to improve the distribution of residuals. See Figure 2 for the main steps of our statistical approach.

We hypothesized that beta diversity should be associated with environmental heterogeneity (Anderson et al. 2006; Heino et al. 2013). However, a variable directly related to environmental heterogeneity could not be



**Figure 2.** Flow chart of the statistical analyses used. Different analyses were employed at (1) the across-basins level and (2) the within-basin level. See main manuscript text for details. DB, Drainage basin; E, Ephemeroptera; P, Plecoptera; T, Trichoptera; C, Chironomidae; O, Odonata; Bio, biological data, Env, environmental data; Geo, geographic coordinates.  $s$ , sites;  $n$ , number of sites;  $y$ , averages of pairwise dissimilarities (from the first to the 95th dataset);  $x_{1-9}$ , explanatory variables for the across-basins analysis;  $p$ , number of taxa in a given biological data matrix;  $q$ ,  $q_+$ , and  $q_-$ , number of environmental variables ( $E$ ) before, after using VIF and after using forward selection, respectively;  $r$  and  $r_-$ , number of eigenvectors variables ( $V$ ) before and after using forward selection.



estimated for all datasets because only biological data were available for some of them. Also, the number and type of environmental variables measured in each basin varied. We therefore assessed whether range in altitude could serve as a proxy for environmental variation among stream sites within a drainage basin. We obtained altitudinal range for each dataset from the Shuttle Radar Topography Mission (Jarvis et al. 2008). We then calculated the average coefficient of variation for stream environmental variables within each dataset and correlated it with altitudinal range, revealing a positive correlation ( $r = 0.55$ ). We thus opted to use altitudinal range as a surrogate variable for environmental variability or heterogeneity because this metric allowed us to include all 95 datasets in the comparative analyses rather than only the 75 (i.e., also including datasets with <15 sites) for which environmental data were available. Altitudinal range is likely to be related to variation in multiple environmental factors affecting the distribution of stream organisms, including temperature, oxygen concentration, and differences in habitat conditions between lowland and mountain streams (Ward 1998; Jacobsen and Dangles 2012). Hence, the larger the altitudinal range in a drainage basin, the larger the variation in stream environmental conditions organisms have to cope with.

Two metrics of spatial extent were obtained initially. The first was the area of a minimum convex polygon including all sites in the drainage basin. Although straightforward, this metric is expected to be biased in cases where a few sites are located far from most of the others. Accordingly, we also obtained the average distance of sites to the geographic centroid of all sites within a drainage basin. These two metrics were strongly correlated ( $r = 0.84$ ), and we opted to use the average distance of sites to the geographic centroid.

## Model evaluations

We calculated the second-order Akaike's information criterion (AIC) to rank the best approximating models explaining variation in our measures of beta diversity (Fig. 2). AIC differences ( $AIC_i - AIC_{min}$ , where  $AIC_{min}$  is the AIC for the best model in the set, given the data) were also calculated. AIC differences were then used to estimate the Akaike weight of each model ( $w_i$ ), which "may be interpreted as the probability that model  $i$  is the actual expected Kullback–Leibler distance best model for the sampling situation considered" (Burnham and Anderson 2002). The sum of Akaike weights over all models that included an explanatory variable ( $w_+(j)$ ) was calculated to estimate the relative importance of explanatory variables. We present the results for all models with AIC differences <2.0.

## Explaining variation in the biological datasets within each drainage basin

In each drainage basin, spatial variables were generated from the Euclidean distance matrix through Moran eigenvector maps (MEM), formerly called principal coordinates of neighbor matrices (Borcard and Legendre 2002). The spatial variables (eigenvectors) derived from MEM represent orthogonal patterns of relationships among sampling sites and are used as spatial predictor variables (Fig. 2). Eigenvectors with high eigenvalues represent broad-scale patterns, whereas those associated with low eigenvalues represent fine-scale patterns. MEM thus produces multiple spatial variables that are efficient in capturing complex spatial patterns in the response data. As watercourse distances between sites were available in only a few datasets, we had to rely on using overland distances in MEM. This is likely to be appropriate because all of our taxa have flying adult stages. Although some studies have suggested that watercourse distances are more meaningful for stream organisms than overland distances (Altermatt et al. 2013), others have found that these two measures of distance are often strongly correlated, providing virtually the same information about the spatial structuring of a metacommunity (Thompson and Townsend 2006; Landeiro et al. 2012; Grönroos et al. 2013).

We used redundancy analysis (RDA; Legendre & Legendre, 2012) to examine the relative contributions of environmental and spatial variables (from MEM) to variation in assemblage composition (using abundance data when available) among sites. Partial RDA (pRDA) was employed to estimate fractions of the total variation of site-by-taxon matrices explained by environmental and spatial variables. The insect matrices were subjected to the Hellinger transformation that is suitable for both presence-absence and abundance data (Legendre et al. 2005), making assemblage data analyzable by linear ordination methods (Peres-Neto et al. 2006). We excluded environmental variables with variance inflation factors higher than 10 (Kutner et al. 2004). Then, we ran two global RDA models, one using all spatial predictors (i.e., eigenvectors) and the second using all remaining environmental predictors. If no global model was significant, the analysis was terminated, and the environmental and spatial fractions were assumed to be zero. When there was a significant global model, we ran a forward selection procedure to retain only the most important variables (Blanchet et al. 2008). In the forward selection, each variable retained should be significant at an alpha level of 0.05 and the adjusted  $R^2$  of the final RDA model should not be greater than its respective global model. When both spatial and environmental global models were significant, we ran a pRDA using the selected

spatial and environmental variables to evaluate how much of the total biological variance each set of explanatory variables could explain. The variance partitioning relative to this pRDA was based on the adjusted  $R^2$  (Peres-Neto et al. 2006).

All analyses were undertaken in R (The R Core Team 2012). Dissimilarities and RDAs were obtained using functions from the *vegan* (Oksanen et al. 2012) and *betapart* (Baselga et al. 2013) packages. Candidate models were compared and model averaging performed using functions of the *MuMIn* package (Bartoń, 2014).

## Results

### Explaining dissimilarity-based beta diversity across drainage basins

Based on the Sørensen coefficient, beta diversity values showed high variation across the 95 datasets, ranging from 0.06 to 0.94 for the average of pairwise dissimilarities and from 0.14 to 0.99 for its multiple-sites version. Beta diversity quantified as the average of pairwise dissimilarities was strongly correlated to the average of distances to the metacommunity centroid ( $r = 0.99$ ). The average of pairwise dissimilarities was also correlated to the multiple-sites version, although to a lesser degree ( $r = 0.75$ ).

The correlation structure among the three beta diversity metrics based on the Simpson coefficient and among the three nestedness-resultant beta diversity metrics was similar to that found for the Sørensen coefficient. The average of pairwise dissimilarities of the Simpson coefficient was strongly correlated with the average of distances to the metacommunity centroid ( $r = 0.95$ ) and to a lesser extent to the multiple-sites version ( $r = 0.76$ ). For the nestedness-resultant beta diversity metrics, the average of pairwise dissimilarities was strongly correlated with the average of distances to the metacommunity centroid ( $r = 0.89$ ) and to a lesser extent with the multiple-sites version ( $r = 0.67$ ).

As all three dissimilarity-based beta diversity approaches were strongly correlated, regression results are shown for the average of pairwise dissimilarities only. The best models for the Sørensen-based beta diversity did not provide support for the hypothesized relationships with altitudinal range, insect group, or latitude (Appendix S4). The best models explained around 68% (adjusted  $R^2$ ) of the variance and included variables related to matrix properties, namely matrix fill, number of sites, and number of taxa (Tables 1 and 2). In decreasing order, the six most important variables were matrix fill, number of sites, number of taxa, altitudinal range, latitude, and spatial extent (Table 2). Similar results were found for the Simpson-based beta diversity, except for the inclusion of altitu-

dinal range and proportion of taxa identified to species in some of the best models (Tables 1 and 2). Regression results for nestedness-resultant beta diversity differed from those for Sørensen and Simpson indices (Tables 1 and 2). A smaller set of models showed AIC<sub>c</sub> differences  $>2.0$ , and they usually tended to fit the data poorly (adjusted  $R^2 < 0.18$ ).

Regardless of the measure of beta diversity used, the most important explanatory variables were more strongly related to matrix dimensions (number of sites and number of species) and dataset characteristics (especially matrix fill) than to biological, ecological, or geographical factors (Table 2). There was also substantial uncertainty regarding the best models, as indicated by their Akaike weights and the evidence ratio between the best models (Table 1). However, the weighted average of the estimates based on model uncertainty also suggested a greater importance of data properties as compared to the other explanatory variables (Table 2).

### Explaining metacommunity variation within drainage basins: variance partitioning

RDAs indicated that about half of the insect datasets were associated with environmental or spatial predictors. Global RDA models for environmental predictors were significant ( $P < 0.05$ ) for 28 of the 61 metacommunities. For spatial predictors, global RDA models were significant for 13 metacommunities.

In only nine cases was metacommunity structure related to both environmental and spatial predictors (Fig. 3). Among these, pRDA models obtained by forward selection indicated that the percentages of variation explained by the environmental variables (fraction [a]) and spatial variables (fraction [c]) were on average (mean  $\pm$  SD)  $13.0 \pm 6.8\%$  and  $6.1 \pm 3.1\%$ , respectively. The average amount of variance shared by environmental and spatial variables [b] was  $10.6 \pm 6.2\%$ . Total explained variance ([a] + [b] + [c]) by the forward-selected pRDA models was  $29.7 \pm 11.5\%$  (Fig. 3).

The structures of 19 metacommunities ( $28 - 9 = 19$ ) were associated exclusively with environmental predictors, and four datasets ( $13 - 9 = 4$ ) exclusively with spatial predictors (Fig. 3). Among the former, environment explained on average  $22.6 \pm 10.0\%$ . For the four datasets exclusively related to spatial predictors, variation explained was  $16.7 \pm 7.9\%$ .

After exclusion of environmental variables with variance inflation factors  $>10$ , the numbers of variables per dataset used in the RDA analyses were reduced to an average of  $10.9 \pm 3.0$  (range: 4–16) with a total of 61 variables. The final environmental RDA models included  $2.6 \pm 0.9$  predictor variables per dataset and a total of 30

**Table 1.** Summary of best models to explain variation in beta diversity quantified as average pairwise dissimilarities for each metacommunity. Models were obtained for Sørensen dissimilarity (total beta diversity), Simpson dissimilarity (beta diversity due to turnover), and nestedness dissimilarity resulting from richness differences. Full models included ecological variables hypothesized to have effects on dissimilarities, including (i) taxonomic group (group), (ii) latitude (absolute values), (iii) range in altitude (alt.rng), and (iv) spatial extent (spt.ext). We included (v) an interaction term because effects of range in altitude may depend on latitude. In addition to ecological factors, we included covariates related to matrix properties (vi) number of sites (log-transformed) (n.sites), (vii) number of species (n.spp), (viii) matrix fill (percentage of presences; fill), and (ix) percentage of taxa identified to species. Best models were selected according to the AICc statistics.

	AICc	df	Delta	Weight	Adj. $R^2$
<b>Sørensen</b>					
fill+n.sites	-178.8	4	0.00	0.373	0.684
fill+n.sites+n.spp	-177.8	5	1.01	0.225	0.685
fill+n.sites+spt.ext	-176.9	5	1.93	0.142	0.682
<b>Simpson</b>					
fill+n.sites+n.spp	-138.1	5	0.00	0.430	0.635
fill+n.sites+n.spp+prop.sp	-136.9	6	1.24	0.231	0.635
fill+n.sites+n.spp+alt.rng	-136.5	6	1.62	0.191	0.634
<b>Richness-resultant</b>					
n.sites+n.spp+prop.sp	-258.7	5	0.00	0.110	0.159
n.sites+n.spp	-258.7	4	0.04	0.108	0.148
n.sites+n.spp+alt.rng+spt.ext	-258.5	6	0.26	0.097	0.168
n.sites+n.spp+alt.rng+spt.ext+prop.sp	-258.1	7	0.63	0.081	0.176
n.sites+n.spp+spt.ext+prop.sp	-258.0	6	0.71	0.077	0.164
n.sites+n.spp+spt.ext	-257.7	5	1.05	0.065	0.150
n.sites+n.spp+alt.rng	-257.6	5	1.16	0.062	0.149
n.sites+n.spp+alt.rng+spt.ext+lat	-257.6	7	1.20	0.061	0.159
n.sites+n.spp+prop.sp+fill	-257.3	6	1.40	0.055	0.158
n.sites+n.spp+fill	-257.3	5	1.43	0.054	0.147
n.sites+n.spp+alt.rng+spt.ext+fill	-257.2	7	1.55	0.051	0.168
n.sites+n.spp+alt.rng+prop.sp	-257.1	6	1.60	0.050	0.156
n.sites+n.spp+spt.ext+prop.sp+fill	-257.0	7	1.70	0.047	0.166
n.sites+n.spp+alt.rng+spt.ext+prop.sp+fill	-256.9	8	1.88	0.043	0.176

df, degrees of freedom; Delta, AIC difference regarding the best model; Weight, Akaike weight; adj  $R^2$ , ordinary adjusted coefficient of determination.

variables. Site elevation and stream width were usually important in explaining variation within metacommunities. However, some variables available in many datasets were never (e.g.,  $\text{SO}_4$  and dead wood material) or seldom selected in the models (e.g., total phosphorus and conductivity; Appendix S5). On the other hand, multivariate models selected a few variables more often than would have been expected from their proportional availability (e.g., bank modification, clogging, and stream order). These variables, however, were available for few datasets (<4), making it difficult to determine their general importance in structuring the metacommunities.

## Discussion

Our findings showed that variation in the beta diversity of stream insects *across* the 95 metacommunities (i.e., spatial level 1) was not well explained by ecological predictors, such as insect group, spatial extent, latitude, and altitudinal range. In addition, environmental, and spatial

variables were poor predictors of assemblage composition *within* most metacommunities (i.e., spatial level 2), although environmental variables tended to be better predictors than spatial ones. Because our findings were consistent across a large number of datasets from around the world, we believe that they represent worldwide patterns in stream insect metacommunities. Below, we will discuss our finding in the context of spatial extent, environmental heterogeneity, latitude, and predictability.

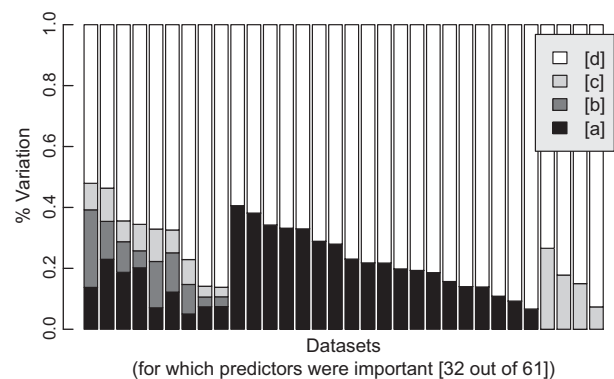
Beta diversity is expected to increase with increasing spatial extent for four reasons (Bini *et al.* 2014; Heino *et al.* 2015b). First, larger areas encompass higher environmental heterogeneity than small areas. Therefore, an increase in environmental heterogeneity is hypothesized to be positively related to the strength of species sorting processes, although evidence for such a relationship is scant (Landeiro *et al.* 2012; Grönroos *et al.* 2013). Second, the effect of dispersal limitation, promoting differences in species composition among sites, is expected to increase with spatial extent (Cottenie 2005; Heino 2011).



**Table 2.** Relative importance of predictor variables for pairwise Sørensen, Simpson and richness-resultant dissimilarities and standardized (beta) coefficients obtained from model averaging over all combinations of model terms. The insect taxon was a categorical variable with five levels coded as a dummy variable. The coefficient for Chironomidae was set to zero.

	Sørensen	Simpson	Richness-resultant
Matrix fill	1.00	1.00	0.32
Relative importance of predictor variables			
Number of sites	0.99	0.99	0.99
Number of taxa	0.38	0.93	0.99
Altitudinal range	0.27	0.36	0.49
Latitude	0.26	0.32	0.34
Spatial extent	0.26	0.27	0.52
Proportion identified species	0.24	0.33	0.47
Insect group	0.05	0.17	0.11
Altitudinal range × Latitude	0.02	0.07	0.07
Model averaging			
Matrix fill	−0.927	−0.826	0.112
Number of sites	−0.268	−0.390	0.416
Number of taxa	0.081	0.219	−0.430
Altitudinal range	0.017	0.029	−0.131
Latitude	0.007	0.018	−0.042
Spatial extent	0.028	−0.039	0.175
Proportion identified species	0.008	0.068	−0.154
Insect taxon: Ephemeroptera	0.078	0.122	−0.050
Odonata	0.144	0.233	−0.229
Plecoptera	0.041	0.127	−0.149
Trichoptera	0.062	0.162	−0.223
Altitudinal range × Latitude	0.095	0.199	−0.234

Third, a positive relationship between beta diversity and spatial extent may arise from sampling different regional species pools (Heino et al. 2015a). Fourth, the relationship between beta diversity and spatial extent is also expected due to a negative relationship between pairwise similarity in assemblage composition and geographic distance (i.e., the distance decay of similarity; Nekola and White 1999). However, our findings did not support any of these mechanisms underlying variability in beta diversity, despite the relatively large variation in spatial extent among the metacommunities studied (measured as average distance to group geographical centroid: min = 0.39 km, max = 125.06 km). Similarly, altitudinal range, an important predictor of biodiversity patterns (Bini et al. 2004; Melo et al. 2009), was a poor predictor of beta diversity in our study. This was surprising because the drainage basins we studied ranged from lowland to montane areas (range of altitudinal range: min = 11 m a.s.l., max = 2136 m a.s.l.), potentially involving considerable variation in environmental factors that affect stream insect distributions between streams located at low and high altitudes (e.g., temperature, flow, and substratum; see Ward 1998; Jacobsen and Dangles 2012).



**Figure 3.** Variation partitioning of the 32 metacommunities for which environmental or spatial predictors were important. In the remaining 29 metacommunities, neither environmental nor spatial predictors significantly explained observed variation. Fractions [a], [b], [c], and [d] correspond to those due to environment, shared environment-space, space, and unexplained variation, respectively. Environmental and spatial predictors were both important for nine metacommunities and all four fractions obtained from partial redundancy analysis (pRDA) models are presented. Environmental or spatial predictors were important for 19 and 4 metacommunities, respectively, and their explained fractions obtained from RDA models. In all cases, a forward selection procedure was used to select predictor variables.

According to Rapoport's rule, species ranges tend to be larger at high latitudes (Stevens 1989), and this phenomenon may give rise to a latitudinal gradient in beta diversity (Soininen et al. 2007). However, latitude was a poor predictor of beta diversity across the metacommunities. It seems that patterns in alpha, beta, and gamma diversity of aquatic insects do not necessarily match those of mammals, birds, or vascular plants (Boyero et al. 2011; Heino 2011). While many terrestrial taxa show pronounced latitudinal gradients (Stevens 1989), gamma diversities of mayflies and stoneflies both peak at mid- to high latitudes, that of caddisflies show no significant relationship with latitude, and that of dragonflies show a latitudinal cline typical of those observed in most terrestrial taxa (Pearson and Boyero 2009). Furthermore, alpha diversities of mayflies, stoneflies, and caddisflies do not follow typical latitudinal clines (Vinson and Hawkins 2003). Although knowledge of global variation in beta diversity of stream insects is limited, information about their gamma and alpha diversities suggests that they are unlikely to exhibit clear latitudinal gradients in beta diversity (Pearson and Boyero 2009; Boyero et al. 2011). Moreover, even when alpha and gamma diversity show a latitudinal cline, such a cline may not necessarily be expected in beta diversity, because different diversity components may vary independent of each other spatially and temporally (Angeler and Drakare 2013).

A salient finding was that dataset characteristics were the main predictors of variation in beta diversity across the 95 metacommunities. Although it seems obvious that matrix fill should be negatively related to beta diversity, our results clearly demonstrate the importance of taking this data property into account when attempting to find correlates of beta diversity. This issue has rarely been considered and we suggest that taking dataset characteristics into account as covariables, before conjecturing about ecological explanations, should be standard practice in future comparative studies.

A combination of low matrix fill (i.e., numerous absences in the site-by-species matrix) and large numbers of rare species typically results in high beta diversity values (Podani and Schmera 2011). Rare species are a typical property of stream insect datasets (Malmqvist *et al.* 1999; Heino 2005). This may be a result of the high variability in stream hydrology. Such variability makes the distributions of individual rare species more difficult to model than the distributions of individual common species (Soininen *et al.* 2013) and, hence, variation in species composition may likewise be highly challenging to model at the assemblage level.

Streams are dynamic systems due to recurrent high or low flows (Resh *et al.* 1988; Townsend 1989; McGarvey 2014), and this variability in flow affects habitat conditions. Such disturbance may temporarily eliminate species from a site, leading to unexpected absences at sites that are otherwise environmentally suitable (Heino and Pekarisky 2014). If such disturbances are frequent, one is likely to find a large number of rare species in a metacommunity at a given point in time. Streams are also highly heterogeneous at various spatial scales (Leung and Dudgeon 2011; Heino *et al.* 2013), promoting ecological specialization and leading to a high number of rare species in stream insect metacommunities (Heino 2005; Allan and Castillo 2007). Thus, high frequencies of rare species in stream insect metacommunities may not only stem from high variability in stream hydrology but also be enhanced by high environmental heterogeneity. Although the relationships between flow variability and different response variables (e.g., abundance and diversity) have been extensively studied (see review in Poff and Zimmerman 2010), we are not aware of any studies that have examined the effects of flow variability on beta diversity. We believe that testing the relationship between beta diversity and flow variability would be a fruitful idea for future research.

Even in cases where predictor variables accounted for statistically significant amounts of variation in assemblage composition *within* metacommunities, the proportions of variation they explained were rather low. A low proportion of explained variation has been a typical finding in

recent studies of stream and other metacommunities using modern analytical methods comparable to those used in our study (Beisner *et al.* 2006; Nabout *et al.* 2009; Landeiro *et al.* 2012; Alahuhta and Heino 2013; Göthe *et al.* 2013; Grönroos *et al.* 2013). The robustness of this finding across the world suggests that the structure of stream insect metacommunities shows low predictability. At the very least, the environmental variables that researchers typically measure in ecological studies may not always account for a large proportion of variation in the assemblage composition of stream insects. The low explanatory power of environmental predictors may stem from the possibility that single snapshot sampling of biological assemblages and environmental variables fails to reveal strong assemblage–environment relationships (Beisner *et al.* 2006; Erős *et al.* 2012). If assemblage–environment relationships generally vary in time (Heino and Mykrä 2008; Erős *et al.* 2012), then this may have major consequences for applied ecology.

Although our predictor variables were important in about half of the datasets and the proportion of variation explained in all cases was low (<50%), environmental variables tended to have a more important influence on stream insect metacommunity structure than spatial variables. This result agrees with previous studies restricted to a single or a few drainage basins (Landeiro *et al.* 2012; Al-Shami *et al.* 2013a; Göthe *et al.* 2013; Grönroos *et al.* 2013). Furthermore, among the many environmental variables included in our study, we found that some easily obtainable variables had particular explanatory power, including stream elevation and stream width. On the other hand, some frequently measured variables (Appendix S3 and S5) were never selected in the models and thus only seem relevant in very specific regional contexts.

Environmental assessment and biodiversity conservation are often based on indicator taxon groups (Caro 2010). In running waters, mayflies, stoneflies, and caddisflies have typically been considered as sensitive indicators of environmental degradation (Rosenberg and Resh 1993), whereas dragonflies have been proposed as a candidate indicator group of overall biodiversity, especially in the tropics (Simaika and Samways 2011). To be efficient in either task, different taxa should be indicators of changes in environmental conditions and variation in their biodiversity should show congruence with other groups (Caro 2010). Although we did not aim to directly examine cross-taxon congruence in biodiversity patterns, it is worth noting that patterns exhibited by single taxonomic groups of the five studied did not necessarily match those of other taxa, which may limit their use as biodiversity indicators at broad spatial scales.

There are three limitations in our current study. First, although our insect data were resolved to the lowest

possible level of identification in each drainage basin, taxonomic resolution varied to some extent among the metacommunities sampled. We controlled for this by a variable describing the percentage of taxa identified to the species level in the comparative analysis, but this variable did not show a significant relationship with variation in beta diversity *across* the metacommunities. Second, we cannot rule out the possibility that unmeasured ecological variables are the main drivers of variation in species composition *within* each metacommunity. Potentially important variables include fish predation, primary production, and proxies of stream-bed disturbance, which have been shown to be important drivers in single case studies (e.g., Townsend *et al.* 2000). Obtaining such site-level variables that are strictly comparable across datasets, however, was impossible given our broad-scale approach. Third, we were limited by data availability to considering only five taxonomic groups, all of which shared some characteristics (e.g., all had flying adult stages). We cannot predict how patterns of beta diversity may play out for groups that are restricted to in-stream dispersal (e.g., fish, mollusks, crayfish, and shrimps).

An important consideration related to the second and third limitations of our study is that spatial variables derived from watercourse and overland distances might have different predictive power in explaining variation in species composition within each drainage basin (Jacobson and Peres-Neto 2010; Altermatt 2013; Altermatt *et al.* 2013). Stream insects disperse from one stream to another by following watercourses or crossing land between headwater streams (Malmqvist 2002; Lancaster and Downes 2013). Thus, it is not clear which dispersal route is more strongly associated with variation in insect species composition among streams. Previous studies have suggested that there is little difference between the predictive power of spatial variables derived from watercourse or overland distances if one uses robust spatial analysis methods, such as Moran eigenvector maps (Landeiro *et al.* 2012; Göthe *et al.* 2013; Grönroos *et al.* 2013). We believe that the Moran's eigenvector map approach used would have modeled spatial structuring of stream insect assemblage composition adequately if, indeed, such structuring existed. Furthermore, using multiple potentially efficient spatial variables did not lead to additional predictive power in our analyses within drainage basins.

In general, the low explanatory power we found in analyses of beta diversity is in line with the comprehensive study of Low-Décarie *et al.* (2014). They evaluated 18 000 ecological articles and found a temporal decline in the coefficient of determination in research studies. We favor one of the three hypothesis proposed by Low-Décarie *et al.* (2014) to explain this trend: "The low hanging

fruit hypothesis proposes that simple discoveries are made early in the development of a discipline and what remains to be explained, at the margins, is increasingly complicated and difficult to reach. In ecology, there appears to be a trend away from single species studies toward more complex community studies, as well as less emphasis on topics that are more observational and arguably less dependent on statistics, such as behavior and physiology, with concurrent increases in statistically complex topics such as biodiversity." Alternatively, the low explanatory power in our analyses may indicate that we should increase our effort in measuring more difficult, albeit potentially more relevant, predictor variables during field surveys.

Our main finding is that variation in the beta diversity of stream insects is difficult to predict *across* metacommunities and *within* each metacommunity. This finding may be related to frequent disturbances in stream systems, resulting in large numbers of rare species in local communities. The weak relationships between beta diversity and latitude, altitudinal range and spatial extent also suggest that stream insects do not follow the geographical patterns observed in alpha, beta, and gamma diversity of various terrestrial taxa. Our comparative analysis of stream insect metacommunities thus (i) reveals a deviance from the general distribution of biodiversity across the world and (ii) calls for reconsideration of the predictability of the responses of stream insect assemblages to ecological gradients.

## Acknowledgments

The writing of this manuscript was supported by grants from the Academy of Finland to JH and the Brazilian Council of Science and Technology (CNPq) to ASM and LMB. SAAS and CSMR were funded by FRGS of Ministry of Higher Education, Malaysia. NB acknowledges the GUADALMED project (HID98-0323-C05 and REN2001-3438-C07) funded by the Spanish Ministry of Science and Technology and the EU-funded project BioFresh (7th FWP contract No 226874). CB and LMM were supported by PIP 5733 CONICET. MC and RL were funded by Companhia Energética de Minas Gerais (Peixe Vivo Program) & P&D ANEEL/CEMIG GT-487 and by Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG PPM-00077/13). NH and RTM received financial support from PRONEX-CNPq-FAPEAM, CNPq and INPA. FOR received productivity grant from CNPq. LFS receives a PhD fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). DS would like to thank Zsuzsa Steindl and Andrea Zagyva (Ministry of Environment and Water, Hungary) for allowing access to the ECOSURV database, and the people participating in

the ECOSURV project. JS was funded by a Claude Leon Postdoctoral Fellowship. TS was supported by grant #2013/50424-1 from São Paulo Research Foundation (FAPESP). RTM received Programa de Apoio à Fixação de Doutores no Amazonas – FIXAM/AM fellowship, and CT-Hidro/Climatic Changes/Water Resources/CNPq (Proc. 403949/2013-0) supported the field collections in Manaus. RT was funded by an Australian Research Council Future Fellowship (FT110100957). We also thank Daniel Hering for data from Germany and comments on the first manuscript draft. Funding to FA is from the Swiss National Science Foundation Grant, and the Swiss Federal Office for the Environment provided the BDM data from Switzerland.

## Conflict of Interest

None declared.

## References

- Alahuhta, J., and J. Heino. 2013. Spatial extent, regional specificity and metacommunity structuring in lake macrophytes. *J. Biogeogr.* 40:1572–1582.
- Allan, J. D., and M. M. Castillo. 2007. *Stream ecology. Structure and function of running waters.* Springer, New York, NY.
- Al-Shami, S. A., M. R. Che Salmah, A. H. Ahmad, and M. R. Madrus. 2013a. Biodiversity of stream insects in the Malaysian Peninsula: spatial patterns and environmental constraints. *Ecol. Entomol.*, 38, 238–249.
- Al-Shami, S. A., J. Heino, M. R. Che Salmah, A. H. Ahmad, S. A. Hamid, and M. R. Madrus. 2013b. Drivers of beta diversity of macroinvertebrate communities in tropical forest streams. *Freshw. Biol.*, 58:1126–1137.
- Altermatt, F. 2013. Diversity in riverine metacommunities: a network perspective. *Aquat. Ecol.* 47:365–377.
- Altermatt, F., M. Seymour, and N. Martinez. 2013. River network properties shape  $\alpha$ -diversity and community similarity of aquatic insect communities across major drainage basins. *J. Biogeogr.* 12:2249–2260.
- Anderson, M. J., K. E. Ellingsen, and B. H. McCaule. 2006. Multivariate dispersion as a measure of beta diversity. *Ecol. Lett.* 9:683–693.
- Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, et al. 2011. Navigating the multiple meanings of  $\beta$  diversity: a roadmap for the practicing ecologist. *Ecol. Lett.* 14:19–28.
- Angeler, D. G., and S. Drakare. 2013. Tracing alpha, beta and gamma diversity responses to environmental change in boreal lakes. *Oecologia* 172:1191–1202.
- Bartoń, K. 2014. MuMIn: multi-model inference. R package version 1.12.1. <http://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf>.
- Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.* 19:134–143.
- Baselga, A., D. Orme, S. Villeger, J. D. Bortoli, and F. Leprieur. 2013. betapart: Partitioning beta diversity into turnover and nestedness components. R package version 1.3. <http://CRAN.R-project.org/package=betapart>
- Beisner, B. E., P. R. Peres-Neto, E. Lindström, A. Barnett, and M. L. Longhi. 2006. The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology* 87:2985–2991.
- Bini, L. M., J. A. F. Diniz-Filho, and B. A. Hawkins. 2004. Macroecological explanations for differences in species richness gradients: a canonical analysis of South American birds. *J. Biogeogr.* 31:1819–1827.
- Bini, L. M., V. L. Landeiro, A. A. Padial, T. Siqueira, and J. Heino. 2014. Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States. *Ecology* 95:1569–1578.
- Blanchet, F. G., P. Legendre, and D. Borcard. 2008. Forward selection of explanatory variables. *Ecology* 89:2623–2632.
- Bonada, N., S. Doledec, and B. Statzner. 2012. Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *J. Biogeogr.* 39:56–68.
- Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecol. Model.* 153:51–68.
- Boyer, L., R. Pearson, D. Dudgeon, M. Graça, M. Gessner, R. J. Albariño, et al. 2011. Global distribution of a key trophic guild contrasts with common latitudinal diversity patterns. *Ecology* 92:1839–1848.
- Boyer, L., R. Pearson, D. Dudgeon, V. Ferreira, M. Graça, M. Gessner, et al. 2012. Global patterns of stream detritivore distribution: implications for biodiversity loss in changing climates. *Glob. Ecol. Biogeogr.* 21:134–141.
- Brown, B. L., and C. M. Swan. 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. *J. Anim. Ecol.* 79:571–580.
- Brown, B. L., C. M. Swan, D. A. Auerbach, E. H. C. Grant, N. P. Hitt, K. O. Maloney, et al. 2011. Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *J. North Am. Benthol. Soc.* 30:310–327.
- Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd Edition. Springer-Verlag, New York, NY.
- Caro, T. 2010. *Conservation by proxy: indicator, umbrella, keystone, flagship, and other surrogate species.* Island Press, Washington, DC.
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecol. Lett.* 8:1175–1182.
- Diserud, O. H., and F. Ødegaard. 2007. A multiple-site similarity measure. *Biol. Lett.* 3:20–22.

- Dudgeon, D., A. H. Arthington, M. O. Gessner, Z.-I. Kawabata, D. J. Knowler, C. Lévêque, et al. 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol. Rev.* 81:163–182.
- Erős, T., P. Saly, P. Takacs, A. Specziar, and P. Biro. 2012. Temporal variability in the spatial and environmental determinants of functional metacommunity organization – stream fish in a human-modified landscape. *Freshw. Biol.* 57:1914–1928.
- Göthe, E., D. G. Angeler, and L. Sandin. 2013. Metacommunity structure in a small boreal stream network. *J. Anim. Ecol.* 82:449–458.
- Grönroos, M., J. Heino, T. Siqueira, V. L. Landeiro, J. Kotanen, and L. M. Bini. 2013. Metacommunity structuring in stream networks: roles of dispersal mode, distance type and regional environmental context. *Ecol. Evol.* 3:4473–4487.
- Hawkins, B. A., and J. A. F. Diniz-Filho. 2006. Beyond Rapoport's rule: evaluating range size patterns of New World birds in a two-dimensional framework. *Glob. Ecol. Biogeogr.* 15:461–469.
- Heino, J. 2005. Positive relationship between regional distribution and local abundance in stream insects: a consequence of niche breadth or niche position? *Ecography* 28:345–354.
- Heino, J. 2011. A macroecological perspective of diversity patterns in the freshwater realm. *Freshw. Biol.* 56:1703–1722.
- Heino, J. 2013. The importance of metacommunity ecology for environmental assessment research in the freshwater realm. *Biol. Rev.* 88:166–178.
- Heino, J., and H. Mykrä. 2008. Control of stream insect assemblages: roles of spatial configuration and local environmental variables. *Ecol. Entomol.* 33:614–622.
- Heino, J., and B. L. Peckarsky. 2014. Integrating behavioral, population and large-scale approaches for understanding stream insect communities. *Curr. Opin. Insect Sci.* 2:7–13.
- Heino, J., M. Grönroos, J. Ilmonen, T. Karhu, M. Niva, and L. Paasivirta. 2013. Environmental heterogeneity and beta diversity of stream macroinvertebrate communities at intermediate spatial scales. *Freshw. Sci.* 32:142–154.
- Heino, J., A. S. Melo, and L. M. Bini. 2015a. Reconceptualising the beta diversity-environmental heterogeneity relationship in running water systems. *Freshw. Biol.* 60:223–235.
- Heino, J., A. S. Melo, T. Siqueira, J. Soininen, S. Valanko, and L. M. Bini. 2015b. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshw. Biol.*, doi:10.1111/fwbi.12533, *in press*.
- Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities – the roles of biotic, abiotic and spatial factors. *Can. J. Fish Aquat. Sci.* 58:157–170.
- Jacobsen, D., and O. Dangles. 2012. Environmental harshness and global richness patterns in glacier-fed streams. *Glob. Ecol. Biogeogr.* 21:647–656.
- Jacobsen, D., A. M. Milner, L. E. Brown, and O. Dangles. 2012. Biodiversity under threat in glacier-fed river systems. *Nat. Clim. Chang.* 2:361–364.
- Jacobson, B., and P. R. Peres-Neto. 2010. Quantifying and disentangling dispersal in metacommunities: how close have we come? How far is there to go? *Landscape Ecol.* 25:495–507.
- Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara. 2008. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), available from <http://srtm.csi.cgiar.org>.
- Kutner, M. H., C. J. Nachtsheim, and J. Neter. 2004. Applied linear regression models, 4th Edition. McGraw-Hill Irwin, New York, NY.
- Lancaster, J., and B. J. Downes. 2013. Aquatic entomology. Oxford Univ. Press, Oxford.
- Landeiro, V. L., L. M. Bini, A. S. Melo, A. M. O. Pes, and W. E. Magnusson. 2012. The roles of dispersal limitation and environmental conditions in controlling caddisfly (Trichoptera) assemblages. *Freshw. Biol.* 57:1554–1564.
- Leather, S. R. 2009. Taxonomic chauvinism threatens the future of entomology. *Biologist* 56:10–13.
- Legendre, P., and L. F. Legendre. 2012. Numerical ecology. Elsevier, Amsterdam.
- Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecol. Monogr.* 75:435–450.
- Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.* 7:601–613.
- Leung, A. S. L., and D. Dudgeon. 2011. Scales of spatiotemporal variability in macroinvertebrate abundance and diversity in monsoonal streams: detecting environmental change. *Freshw. Biol.* 56:1193–1208.
- Logue, J.B., N. Mouquet, H. Peter, H. Hillebrand, and The Metacommunity Working Group. 2011. Empirical approaches to metacommunities: a review and comparison with theory. *Trends Ecol. Evol.* 26:482–491.
- Low-Décarie, L., C. Chivers, and M. Granados. 2014. Rising complexity and falling explanatory power in ecology. *Front. Ecol. Environ.* 12:412–418.
- Malmqvist, B. 2002. Aquatic invertebrates in riverine landscapes. *Freshw. Biol.* 47:679–694.
- Malmqvist, B., Y. Zhang, and P. H. Adler. 1999. Diversity, distribution and larval habitats of North Swedish blackflies (Diptera: Simuliidae). *Freshw. Biol.* 42:301–314.
- McGarvey, D. J. 2014. Moving beyond species-discharge relationships to a flow-mediated, macroecological theory of fish species richness. *Freshw. Sci.* 33:18–31.



- Melo, A. S., T. F. L. V. B. Rangel, and J. A. F. Diniz-Filho. 2009. Environmental drivers of beta-diversity patterns in New-World birds and mammals. *Ecography* 32:226–236.
- Nabou, J. C., T. Siqueira, L. M. Bini, and I. S. de Nogueira. 2009. No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta Oecol.*, 35:720–726.
- Nekola, J. C., and P. S. White. 1999. The distance decay of community similarity in biogeography and ecology. *J. Biogeogr.* 26:867–878.
- Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, et al. (2012) *vegan: Community Ecology Package*. R package version 2.0-5. <http://CRAN.R-project.org/package=vegan>.
- Pearson, R. G., and L. Boyero. 2009. Gradients in regional diversity of freshwater taxa. *J. North Am. Benthol. Soc.* 28:504–514.
- Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87: 2614–2625.
- Podani, J., and M. Schmera. 2011. A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data. *Oikos* 120: 1625–1638.
- Poff, N. L., and J. H. K. Zimmerman. 2010. Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshw. Biol.* 55:194–205.
- Raunio, J., J. Heino, and L. Paasivirta. 2011. Non-biting midges in biodiversity conservation and environmental assessment: findings from boreal freshwater ecosystems. *Ecol. Ind.* 11:1057–1064.
- Resh, V. H., A. Brown, A. P. Covich, M. E. Gurtz, H. W. Li, W. Minshall, et al. 1988. The role of disturbance theory in stream ecology. *J. North Am. Benthol. Soc.* 7:433–455.
- Rosenberg, D. M., and V. H. Resh. 1993. *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York, NY.
- Simaika, J. P., and M. J. Samways. 2011. Comparative assessment of indices of freshwater habitat conditions using different invertebrate taxon sets. *Ecol. Ind.* 11:30–378.
- Siqueira, T., L. M. Bini, F. O. Roque, M. Pepinelli, R. C. Ramos, S. R. M. Couceiro, et al. 2012. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. *Ecography* 35:183–192.
- Soininen, J., R. MacDonald, and H. Hillebrand. 2007. The distance decay of similarity in ecological communities. *Ecography* 30:3–12.
- Soininen, J., J. J. Korhonen, and M. Luoto. 2013. Stochastic species distributions are driven by organism size. *Ecology* 94:660–670.
- Stevens, G. C. 1989. The latitudinal gradients in geographical range: how so many species co-exist in the tropics. *Am. Nat.* 133:240–256.
- The R Core Team. 2012. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Thompson, R. M., and C. R. Townsend. 2006. A truce with neutral theory: local deterministic factors, species traits and dispersal limitation together determine patterns of diversity in stream invertebrates. *J. Anim. Ecol.* 75: 476–484.
- Townsend, C. R. 1989. The patch dynamics concept of stream community ecology. *J. North Am. Benthol. Soc.* 8:36–50.
- Townsend, C. R., R. M. Thompson, A. R. McIntosh, C. Kilroy, E. D. Edwards, and M. R. Scarsbrook. 2000. Disturbance, resource supply and food-web architecture in streams. *Ecol. Lett.* 1:200–209.
- Vinson, M. A., and C. P. Hawkins. 2003. Broad-scale geographical patterns in local stream insect genera richness. *Ecography* 26:751–767.
- Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, et al. 2010. Global threats to human water security and river biodiversity. *Nature*, 467:555–561.
- Ward, J. V. 1998. Riverine landscapes: biodiversity patterns, disturbance regimes, and aquatic conservation. *Biol. Conserv.* 83:269–278.

## Supporting Information

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

**Appendix S1.** A schematic figure showing the spatial level 1 of our analyses: across multiple metacommunities.

**Appendix S2.** A schematic figure showing the spatial level 2 of our analyses: within each metacommunity.

**Appendix S3.** Environmental variables available and the frequency of datasets in which they appeared.

**Appendix S4.** Relationship between Sørensen beta diversity and latitude for the five insect taxa.

**Appendix S5.** Frequency of environmental variables in datasets and frequency at which they were selected in RDA or pRDA models.