ECOGRAPHY

Research

A comprehensive examination of the network position hypothesis across multiple river metacommunities

Renato Henriques-Silva, Maxime Logez, Nathalie Reynaud, Pablo A. Tedesco, Sebastien Brosse, Stephanie R. Januchowski-Hartley, Thierry Oberdorff and Christine Argillier

R. Henriques-Silva (http://orcid.org/0000-0003-2731-2023) (renato.henriques@irstea.fr), M. Logez, N. Reynaud and C. Argillier, Irstea, UR RECOVER, Aix-en-Provence, France. – P. A. Tedesco, S. Brosse and T. Oberdorff, UMR EDB, IRD 253, CNRS 5174, UPS, ENFA, Univ. Paul Sabatier, Toulouse, France. – S. R. Januchowski-Hartley, Swansea Univ., Dept of Biosciences, Swansea, UK.

The hierarchical branching nature of river networks can have a strong influence on the assembly of freshwater communities. This unique structure has spurred the development of the network position hypothesis (NPH), which states that the strength of different assembly processes depends on the community position in the river network. Specifically, it predicts that 1) headwater communities should be exclusively controlled by the local environment given that they are more isolated and environmentally heterogeneous relative to downstream reaches. In contrast, 2) downstream communities should be regulated by both environmental and dispersal processes due to increased connectivity given their central position in the riverscape. Although intuitive, the NPH has only been evaluated on a few catchments and it is not yet clear whether its predictions are generalizable. To fill this gap, we tested the NPH on river dwelling fishes using an extensive dataset from 28 French catchments. Stream and climatic variables were assembled to characterize environmental conditions and graph theory was applied on river networks to create spatial variables. We tested both predictions using variation partitioning analyses separately for headwater and downstream sites in each catchment. Only 10 catchments supported both predictions, 11 failed to support at least one of them, while in 7 the NPH was partially supported given that spatial variables were also significant for headwater communities. We then assembled a dataset at the catchment scale (e.g. topography, environmental heterogeneity, network connectivity) and applied a classification tree analysis (CTA) to determine which regional property could explain these results. The CTA showed that the NPH was not supported in catchments with high heterogeneity in connectivity among sites. In more homogeneously connected catchments, the NPH was only supported when headwaters were more environmentally heterogeneous than downstream sites. We conclude that the NPH is context dependent even for taxa dispersing exclusively within streams.

Keywords: context-dependent, river network, environmental heterogeneity

www.ecography.org

Introduction

The establishment of metacommunity theory (Leibold et al. 2004) revolutionized our understanding of community assembly (Cottenie 2005, Logue et al. 2011, Heino et al. 2015). While ecologists have been aware of dispersal-driven ecological processes since the seminal work of MacArthur and Wilson (1963), it was long assumed that the local environment was the primary factor selecting which species could persist in any given community. However, it is now increasingly acknowledged that community composition is also substantially influenced by the dispersal of organisms across habitat patches (Cottenie et al. 2003). When colonization rates exceed extinction rates, species are able to occupy all habitat patches in which the environment is tolerable for them (i.e. species sorting; Chase and Leibold 2003). In contrast, when colonization rates are non-existent or low (i.e. dispersal limitation) species may be absent from otherwise environmentally suitable habitats (Leibold et al. 2004), and when colonization rates are very high (i.e. mass effects), species can temporarily occupy habitat patches that are not suitable for them (Pulliam 1988).

More recently, ecologists have focused on determining how ecological factors such as ecosystem type, taxonomic group, natural disturbance and landscape connectivity influence the relative importance of environmental and dispersal factors on metacommunities (Cottenie 2005, Henriques-Silva et al. 2013, Fernandes et al. 2014). Particular attention has been given to landscape connectivity in river networks because of their dendritic nature that creates preferential (or obligatory, considering strictly aquatic taxa) connections between habitat patches (Fagan 2002, Brown and Swan 2010, Altermatt et al. 2013, Zhao et al. 2017, Schmera et al. 2018). Hence, the spatial positioning of habitats within a stream network can play a crucial role in determining how communities are structured by dispersal and environmental filtering processes (Altermatt et al. 2013). Based on previous studies, Brown and Swan (2010) developed two predictions about community assemblages in river networks that were recently formalized into the network position hypothesis (NPH) by Schmera et al. (2018). The NPH first predicts that 1) communities in headwater reaches are exclusively regulated by environmental processes because they are relatively more isolated and environmentally heterogeneous than downstream reaches (Brown and Swan 2010). Secondly, the NPH predicts that 2) community assemblages in downstream reaches are increasingly influenced by mass effects due to a surplus of dispersal because of their central position in the river network. Hence, both environmental conditions and dispersal should account for compositional variation in downstream communities (Brown and Swan 2010). While the NPH has been highly influential in river ecology, few studies have tested its predictions explicitly and those that have were rather inconclusive (Tonkin et al. 2016, Erős et al. 2017, Göthe et al. 2017, Schmera et al. 2018). Examples of inconsistencies include a lack of support for the NPH predictions across taxonomic groups (Schmera et al. 2018), and the relative importance of environmental and dispersal processes being dependent on the systems being studied rather than network position (Datry et al. 2016a, b, Tonkin et al. 2016). With previous findings in mind, an opportunity exists to determine whether the NPH predictions are general (i.e. applicable across all river networks), or context-dependent (Schmera et al. 2018).

There are also several challenges to evaluating the NPH predictions, and their influences on community assembly in river metacommunities. First, the NPH is founded on the structure of river networks, meaning that it can only be validated on taxa that disperse mostly through river corridors. The assumption that headwater reaches are more isolated than those further downstream does not necessarily hold for taxa that can disperse overland (e.g. aquatic insects that can fly in their adult stage) (Schmera et al. 2018). Second, the NPH explicitly invokes river network position in determining community assemblages, but the network itself has often been disregarded or not explicitly considered in previous studies. Most, if not all, studies evaluating the NPH used spatial eigenfunction analysis or watercourse and geographical distance to model spatial predictors in river networks. However, spatial eigenvectors are known to also represent unmeasured spatially-structured environmental factors (Smith and Lundholm 2010, Fernandes et al. 2014) and simple spatial distances may not be adequate to account for the complex structure of river networks (i.e. they do not account for variation in connectivity across sites). However, river networks are ideal systems to be represented by graphs – mathematical constructs representing nodes (points) connected by a set of edges (links) (Urban and Keitt 2001) that can be used to estimate the complex connectivity patterns of riverine ecosystems (Erős et al. 2011, 2012). Third, studies evaluating the NPH predictions have not been sufficiently replicated; the majority of studies were based on a few catchments (but see Göthe et al. 2017). There is a need to expand previous studies to evaluate the NPH across multiple river metacommunities, and expand our understanding on the applicability and generality of its predictions (Tonkin et al. 2016, 2018). A large geographical region encompassing multiple, diverse, river networks is an exceptional venue to evaluate the NPH and to search for general processes that regulate community assembly in river networks.

In the present study we evaluate the NPH on an extensive dataset of riverine fish assemblages, containing thousands of sampling records in multiple French catchments. France's heterogeneous climate and topography make it an excellent area to test the NPH, and to evaluate the processes that regulate community assembly in river networks. The objectives of our study were two-fold: first 1), to determine if the strength of environmental and dispersal processes varies between headwater and downstream fish communities. For this goal, we applied graph theory instead of traditional methods to account for the river network configuration in the spatial variables. Second 2), to verify whether catchment characteristics may explain the variation in the importance of environmental

and dispersal processes across river networks. For instance, environmental heterogeneity (Heino et al. 2012), topographic heterogeneity (Tonkin et al. 2017), climate (e.g. arid vs humid) (Datry et al. 2016a) as well as variation in river network configuration among drainage basins (i.e. spatial arrangement of sites within the river network; Tonkin et al. 2017, 2018) may influence the effects of environmental and dispersal processes in riverine metacommunities.

Methods

Catchment, river networks and fish community data

Catchments were delineated according to the pan-European catchment characterization and modelling (CCM2; Vogt et al. 2007) database, and represented river networks in each of the respective catchments using the French digital river network (RHT; Pella et al. 2012) GIS database. The RHT was derived from a 50 m digital elevation model and is considered a hydrological reference in France (Pella et al. 2012). Drawing on these data we defined headwater reaches as Strahler orders 1 and 2, and downstream reaches as those with Strahler order > 2. Downstream reaches were significantly more distant from the source ($t=-24.8$; p < 0.0001) and larger than headwater reaches both across ($t=-17.702$; p < 0.0001) and within catchments.

Fish community data was represented by species abundances (i.e. number of individuals) and was obtained from national fishing surveys conducted by the French Biodiversity Agency (AFB) on French rivers (Poulet et al. 2011) between 1995 and 2012. All fish sampling was carried out using electrofishing both by wading smaller streams (85% of sampled areas) and using boat for higher order reaches. Given that the surface area of sampling operations varied across sites due to variation in river width, species abundances were converted to densities by dividing them by log-transformed sampling surface area. Most sites were sampled with one or two pass removal but, to homogenize the data, we only included individuals sampled in the first pass in our dataset. We also eliminated all sites sampled before May to avoid recruitment timing bias and excluded brackish and marine species (i.e. *Alosa fallax, Atherina presbyter, Chelon labrosus, Liza aurata, Liza ramada, Mugil cephalus, Platichthys flesus* and *Pleuronectes platessa*). Given that recently introduced and heavily stocked species may not be regulated by local environmental and dispersal processes, we identified those using Keith et al. (2011) and excluded them from the dataset (i.e. *Ctenopharyngodon idella, Hypophthalmichthys molitrix, Neogobius melanostomus, Oncorhynchus mykiss, Pimephales promelas* and *Ponticola kessleri*). Moreover, some river reaches contained more than one sampling site (i.e. fish community). To reduce spatial bias, we performed spatial thinning and kept only one sampling site per reach. For each catchment, we chose the site that was sampled in (or near to) the year that represented the average sampling year across all sites in the catchment to avoid temporal bias in our analyses. We included only the data from

catchments that had at least 15 sampled sites in both headwater and downstream reaches. Our final dataset included 57 species sampled in 3353 sites across the 28 catchments (Fig. 1). Descriptive data for each catchment and each species are presented in the Supplementary material Appendix 1 Table A1 and Table A2, respectively. The data treatment and statistical analyses described below were conducted with the R software (R Core Team). Note that the terms 'catchment' and 'metacommunity', and the terms 'reach' and 'site', are used interchangeably.

Environmental variables

For each reach, we included environmental variables [E] available from the RHT, and previously shown to be strongly related to stream fishes distributions (Oberdorff et al. 2001, Buisson et al. 2008), including: mean stream width (WID, m), elevation (ELEV, m), river slope (SLO, ‰), mean discharge (MD, m^3 s^{-1}) and river bed surface grain size (RBSGS; Snelder et al. 2011). We also extracted temperature and precipitation variables from SAFRAN (8 km grid-data; Vidal et al. 2010) between the periods of 1990 and 2015, which encompasses the total period of sampling of the fish dataset. We then used these data to compute bioclimatic variables following the Worldclim methodology (Hijmans et al. 2005). For details on the computation of these variables and their units we refer to <www.worldclim.org>. Descriptive statistics for all these variables computed at the site scale are presented in Supplementary material Appendix 1 Table A3. Prior to statistical analysis, we verified and reduced multicollinearity in environmental variables in each catchment by following Naimi et al. (2014). Variables with variance inflation factor larger than 10 were excluded using a stepwise procedure.

Spatial variables

The spatial variables [S] were selected in order to represent various aspects of river network configuration, and included: distance to the source (DIS, km), upstream drainage area (UDA, km^2) and three network metrics. These first two variables were log-transformed and they both are typically used as proxies for the reach's position in the upstream-downstream longitudinal gradient (Oberdorff et al. 2001). We characterized the river network of each catchment as graphs using the network analysis extension in ArcGIS software ver. 10.3 (ESRI 2014), in which sampling sites were defined as nodes and the stream corridors connecting them as edges (Erős et al. 2012). For simplicity, we developed undirected graphs because fishes may disperse either through passive downstream drift or active upstream swimming depending on their life stages (Olden et al. 2001). The edges connecting nodes were weighted using the following function: $w_{ij}=1-(d_{ij}/d_{MAX})^3$, where d_{ij} is the watercourse distance between site i and j and d_{MAX} is the maximum distance found between any two nodes in the catchment. The weights were used to represent the ease of dispersal among sites and the
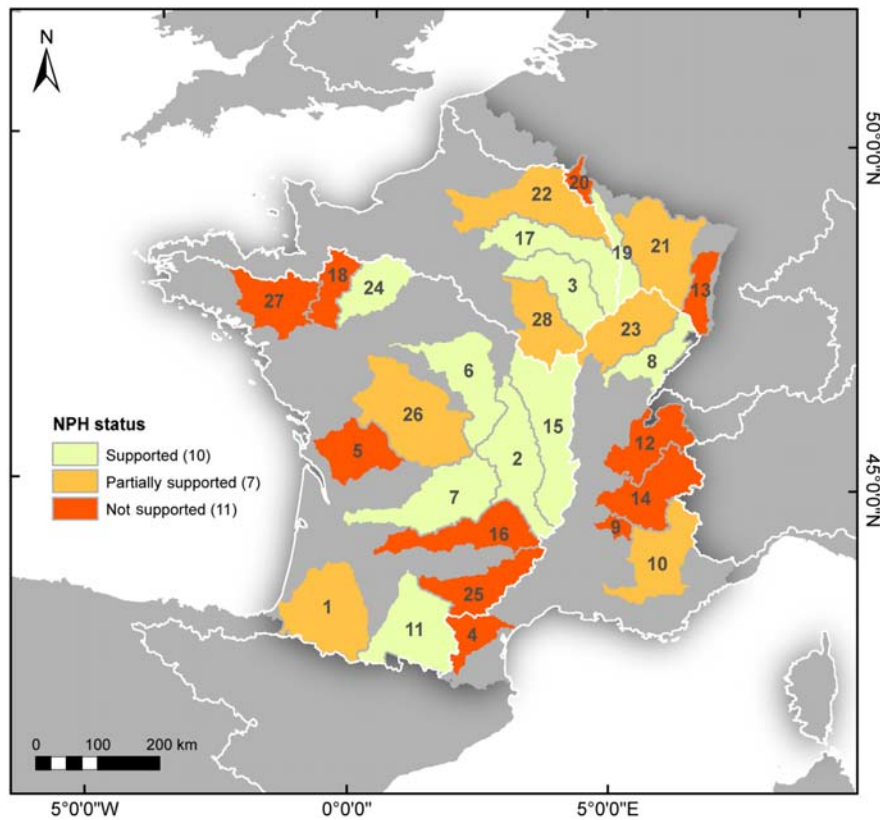
Figure 1. Map of France with selected catchments used in the analysis. The coloring pattern shows whether we found support (*Su*, pale green), partial support (*PSu*, orange) or no support (*NSu*, red) for the NPH predictions in the catchment under the VP$_{NETWORK}$ framework. The numbers in the legend refers to the number of catchments that were assigned each of these classifications. Refer to the Methods section to understand how these three classifications were assigned. The numbers inside the catchments represent their codes (see tables in Supplementary material Appendix 1–9 for catchment names/codes correspondences).

exponent decrease the importance of sites that are very distant from each other (Borcard et al. 2011, Zhao et al. 2017). Hence, these weighted graphs provided information on both watercourse distance and variation in connectivity among sites. We then computed three node metrics for each catchment's graph: degree (DE), the number of nodes connected to the focal node; betweenness centrality (BC), the number of shortest path from all nodes that pass through the focal node; and closeness centrality (CC), the inverse of the sum of the shortest distance between the focal node and every other node in the network (Rayfield et al. 2011). We normalized all metrics by the total number of sites (i.e. nodes) in the catchment. Ecologically, sites with high values for these network metrics are generally located in central positions in the landscape, and thus are critical for dispersal pathways across the river network (Erős et al. 2011, Rayfield et al. 2011). Formulas and examples of these metrics are presented in the Supplementary material Appendix 2 Fig. 1A, 2A, 3A.

Statistical analysis

To estimate the amount of variation explained by environmental and spatial variables for each catchment we applied variation partitioning analysis (VPA; Borcard et al. 1992).

This method uses a community (sites-by-species) matrix as the response variable and sites-by-environmental variables [E] and sites-by-spatial variables [S] as explanatory predictors. To create the community matrix [Y], we applied Hellinger-transformation on fish densities as recommended by Legendre and Gallagher (2001) for zero-inflated data in ordination analysis. Then, all environmental variables [E] were tested independently on [Y] in global models using redundancy analysis (RDA). When these global models were significant, a forward selection procedure was applied on the environmental matrix (Blanchet et al. 2008). Note that forward selection was not used for the spatial model [S] with network metrics due to the small number of predictors (Peres-Neto and Cumming 2010). We then applied a RDA-based VPA, which decomposes additively the total variation in the response data into four variance fractions: environment variation independent of spatial variation [E], spatially-structured environmental variation [E + S], spatial variation independent of environmental variation [S], and residual variation [R]. To estimate these fractions we used the adjusted redundancy statistic (R^2_{adj}), which controls for the number of predictors and sample size (Peres-Neto et al. 2006), thus allowing comparisons across metacommunities. Finally, we ran permutation tests (n = 1000) on partial RDAs

to estimate the significance ($p < 0.05$) of [E] and [S]. We also ran partial RDAs with both environmental and spatial variables as predictors and site sampling year as a covariable to determine whether the effect of these variables was robust to compositional change over time.

In order to contrast our graph-based approach to the traditional methods used in stream metacommunity ecology, we also performed two additional analyses. First, we tested the NPH with a VPA using spatial eigenfunction analysis to model spatial processes (Landeiro et al. 2011, Schmera et al. 2018). We applied principal coordinate of neighboring matrices (PCNM; Griffith and Peres-Neto 2006) on watercourse distances, an undirected method that is recommended for constrained ordination (Dray et al. 2006). Second, we tested the NPH by evaluating the relationship between community dissimilarity with environmental and water course distance among sites (Brown and Swan 2010, Jamoneau et al. 2018) using Mantel and partial Mantel tests with 999 permutations (Smouse et al. 1986). A more detailed description of both analyses is provided in the Supplementary material Appendix 3 and 4. From here we refer to the VPA using river network configuration metrics as $VP_{NETWORK}$ and VPA using spatial eigenfunction analysis as VP_{PCNM}. No acronym was used for Mantel tests.

Similar to previous studies evaluating the NPH (Brown and Swan 2010, Schmera et al. 2018), we interpreted the output of the two VPAs and two Mantel tests in relation to the NPH predictions as: the NPH was supported (*Su*) when [E] was statistically significant in both headwater and downstream reaches, and when [S] was only statistically significant for downstream reaches; the NPH was not supported (*NSu*) if [E] was not statistically significant in either headwater or downstream sites, or if [S] was not statistically significant in downstream reaches (Table 1). Finally, the assumption that only environmental factors are important in structuring communities within headwater reaches has yet to be rigorously tested (Schmera et al. 2018). For instance, headwater communities may be so isolated that dispersal limitation may increase the importance of stochasticity and non-equilibrium dynamics (Erős 2017) and, consequently, the signal of spatial processes (Heino et al. 2015, Göthe et al. 2017). In this case, species sorting would cease to be the exclusive structuring process in these communities. With this in mind, we explored an additional test whereby the NPH was partially supported

(*PSu*) when both [E] and [S] were statistically significant for both headwater and downstream sites (Table 1).

Regional variables and analysis

To determine whether regional properties influence the support for the NPH , we assembled a catchment-scale dataset including number of sites, geographical coordinates, indices of topographic and environmental heterogeneity, and regional-scale network metrics (Table 2).

To account for differences in sample sizes (Sály and Erős 2016), we included the total number of sites (total_SITES) as well as the total number of headwaters (total_HEADW) and downstream sites (total_RIVER) from each catchment. We also included catchment surface (SA) area as the relative importance of assembly processes may vary with spatial extent (Heino et al. 2015) and we computed their latitude (LAT) and longitude (LONG) as the average latitude and longitude of their constituent sites, respectively.

Topographic heterogeneity was measured in two ways: coefficient of variation in altitude (cv_ALT) and slopes (cv_SLO) across reaches in each catchment. Higher values for these indices suggest a more heterogeneous topography. Environmental heterogeneity (EH) was calculated following Henriques-Silva et al. (2013) as the average environmental Euclidean distance among sites within each catchment. We included bioclimatic variables describing annual mean, seasonality and extreme values of temperature and precipitation (i.e. BIO1, BIO2, BIO3, BIO4, BIO5, BIO6, BIO7, BIO12, BIO13, BIO14, BIO15; Supplementary material Appendix 1 Table A3), elevation and all stream variables. All variables were standardized and we computed two indices for each catchment: total environmental heterogeneity (total_EH), which was computed using all sites, and relative environmental heterogeneity (relative_EH), which was computed as the ratio between the EH of headwaters to the EH of downstream reaches (i.e. in this case EH was computed for headwater and downstream sites separately). The first index can be interpreted as the 'length' of the environmental gradient, where larger values indicate larger EH. The second can be interpreted as such: values near 1 suggest that headwater and downstream sites have similar EH, values > 1 indicate that headwaters are more heterogeneous than mainstems and values below 1 suggest the opposite.

Table 1. Evaluation of the network position hypothesis (NPH) based on the statistical significance ($\alpha = 0.05$) of pure environmental [E] and pure spatial [S] fractions in the variation partitioning analyses performed on headwater and downstream sites. p=p-value. The NPH is not supported in either of the cases described in the three last rows of the table. The symbol '–' indicates that the p-value for this fraction does not matter due to a crucial condition not being met for another fraction.

[E]		[S]		NPH
Headwater	Downstream	Headwater	Downstream	
p ≤ 0.05	**p ≤ 0.05**	p > 0.05	**p ≤ 0.05**	Supported (*Su*)
p ≤ 0.05	**p ≤ 0.05**	**p ≤ 0.05**	**p ≤ 0.05**	Partially supported (*PSu*)
p > 0.05	–	–	–	Not supported (*NSu*)
–	p > 0.05	–	–	Not supported (*NSu*)
–	–	–	p > 0.05	Not supported (*NSu*)

Table 2. Regional catchment properties and indices computed for regional analysis.

Regional variable	Description
total_SITES	total number of sites sampled
total_HEADW	total number of headwater sites sampled
total_RIVER	total number of downstream sites sampled
LAT	average latitude of constituent sites
LONG	average longitude of constituent sites
SA	log-transformed catchment surface area
cv_ELE	coefficient of variation in sites' elevation
cv_SLO	coefficient of variation in sites' slope
total_EH	catchment environmental heterogeneity
relative_EH	environmental heterogeneity of headwater relative to downstream sites
DE	centralization-network degree
CC	centralization-network closeness centrality
BC	centralization-network betweenness centrality

Finally, we described river network topology by upscaling the three node metrics described previously (i.e. DE, CC and BC) to the entire network (i.e. catchment) using the centralization method established by Freeman (1979). The centralization of node properties represents the variation in these node properties across the network. In other words, it measures the extent to which the connectivity of a graph is organized around particular focal nodes (Freeman 1979). These values vary from 0 to 1, in which a 1 represents a star-shaped graph (i.e. all nodes are connected to one central node; Supplementary material Appendix 2 Fig. A4a) and 0 a complete graph (i.e. every pair of node is connected through a unique edge; Supplementary material Appendix 2 Fig. A4b). Thus, high values on one of these metrics suggest that the stream metacommunity is composed of many peripheral communities connected to only a few central communities.

We used a classification tree analysis (CTA) to determine whether, and how, regional variables influence the support of the NPH predictions. CTA is a flexible and robust predictive modeling approach, allowing the use of different types of explanatory variables. The CTA also deals with non-linear relationships and high-order interactions (De'ath and Fabricius 2000). CTA is nonparametric, and thus collinearity does not prevent reliable parameter estimation (Cutler et al. 2007). Given the small number of observations (i.e. 28 catchments), splits were allowed as long as resulting leaves had a minimum of 5 observations. The tree was evaluated with a 10-fold cross-validation procedure and we pruned the tree by selecting the split with the least cross-validated error to avoid overfitting the data. The response variable was the support for the NPH (i.e. *Su*, *PSu*, *NSu*) whereas all regional variables described previously (Table 2) were used as predictors. The percent variation (R^2) explained by the CTA was calculated as follows: $R^2 = 1$ – relative error (Sharma et al. 2012). Given that we were able to validate the NPH in a much larger number of catchments using our graph-based method (see results below), we limited the regional analysis only to the NPH classification obtained from the $VP_{NETWORK}$.

We used the 'igraph' package (Csardi and Nepusz 2006) to compute the network metrics, the 'usdm' package (Naimi et al. 2014) to compute the variance inflation factor, the 'adespatial' package (Dray et al. 2017) to perform the forward selection procedure and the 'vegan' package to create the *PCNM*s, and to perform the RDA, Mantel tests and VPAs (Oksanen et al. 2015). Finally, we used 'rpart' (Therneau et al. 2015) and 'rattle' (Williams 2011) packages to conduct and plot the classification tree analysis, respectively.

Data deposition

Results

We found that both [E] and [S] fractions obtained from the variation partition using river network configuration to model spatial connectivity ($VP_{NETWORK}$) exhibited a large variability across catchments (Fig. 2). However, when we pooled the results from all catchments and compared headwater and downstream metacommunities, we did not find any statistical difference for either [E] (Wilcoxon test; $W = 406$, $p = 0.82$) or [S] (Wilcoxon test; $W = 503$; $p = 0.07$) while the difference in [E + S] was marginally significant (Wilcoxon test; $W = 273$, $p = 0.051$). $VP_{NETWORK}$ on headwater sites resulted in an average [E] of 0.131 ± 0.103, [S] of 0.031 ± 0.057, and [E + S] of 0.089 ± 0.088 (Supplementary material Appendix 5 Table A4). The same analysis performed on downstream sites resulted in an average [E] of 0.133 ± 0.076, [S] of 0.044 ± 0.038, and [E + S] of 0.169 ± 0.072 (Supplementary material Appendix 5 Table A5). Variation partition using *PCNM* (VP_{PCNM}) to model spatial connectivity on headwater sites resulted in average [E] of 0.082 ± 0.216, [S] of 0.014 ± 0.142 and [E + S] of 0.140 ± 0.167 (Supplementary material Appendix 5 Table A6). The same analysis performed on downstream sites resulted in an average [E] of 0.129 ± 0.052, [S] of 0.051 ± 0.045 and [E + S] of 0.172 ± 0.107 (Supplementary material Appendix 5 Table A7). Similarly to $VP_{NETWORK}$, we found no statistical difference between headwater and downstream metacommunities for either [E] (Wilcoxon test; $W = 488$, $p = 0.118$), [S] (Wilcoxon test; $W = 429$, $p = 0.552$) or [E + S] (Wilcoxon test; $W = 497$, $p = 0.086$) fractions. We also found that Mantel and partial-Mantel tests mirrored the results from the variation partitioning analysis, exhibiting large variability among catchments (Supplementary material Appendix 6 Table A10 and A11).

We found that $VP_{NETWORK}$ supported the NPH predictions in 10 catchments when considering [E] and [S] (Fig. 1, Supplementary material Appendix 5 Table A8). Moreover, we observed that the NPH was partially supported in 7 catchments, but not supported in the remaining 11 (Fig. 1). Partial RDAs with sampling year revealed that compositional change over time was only significantly important in 1 headwater
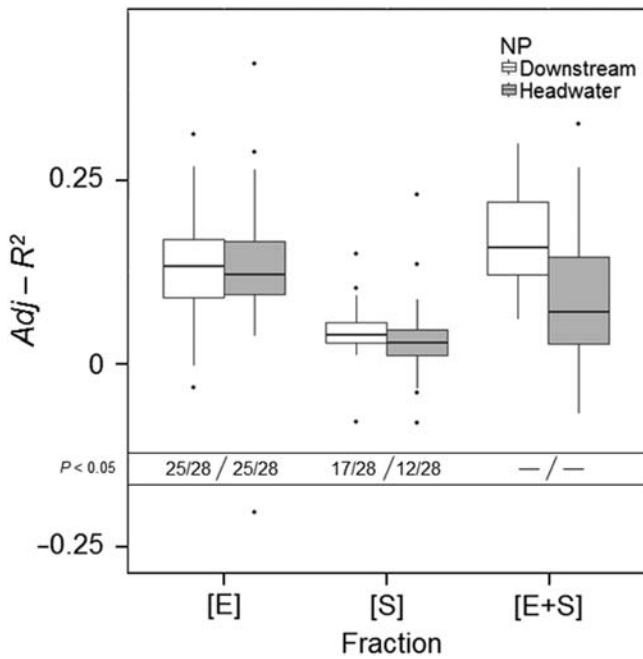
Figure 2. Boxplots depicting the adjusted R^2 of the pure environmental fraction [E], spatially-structured environmental fraction [E + S] and pure spatial fraction [S] obtained from the variation partitioning analysis performed on headwater and downstream sites across all 28 catchments using river network configuration metrics ($VP_{NETWORK}$). The numbers below the boxplots represent the proportion of catchments in which the fractions were statistically significant ($\alpha = 0.05$). The bottom and top boundary of the boxplots represent the 25 and 75% percentiles, respectively, while the line within them represents the median. Top and bottom whiskers outside the boxplots represent ± 1.5 times the distance between the first and third quartiles. Data beyond whiskers are outliers and plotted as dots. The adjusted R^2 values used in this figure are presented in Supplementary material Appendix 5 Table A8.

and 3 downstream metacommunities; adjusted-R^2 varied between 0.01 and 0.02. There was no clear bias towards any NPH classification; sampling year affected one *Su,* one *NSu,* and two *PSu* catchments (Supplementary material Appendix 7 Table A12, A13). Both VP_{PCNM} and Mantel tests resulted in similarly mixed support of the NPH; VP_{PCNM} with 3, 7 and 18 catchments with full, partial or no support, respectively (Supplementary material Appendix 5 Table A9), Mantel tests with 6, 9 and 13 catchments with full, partial or no support, respectively (Supplementary material Appendix 6 Table A10) and partial-Mantel tests result with 5, 1 and 22 catchments with full, partial or no support, respectively (Supplementary material Appendix 6 Table A11).

Our CTA explained $R^2 = 64.7\%$ of the variation in NPH support across catchments (Supplementary material Appendix 8 Table A14). The analysis resulted in a tree with 2 splits and 3 terminal leaves (Fig. 3). The first split was based on regional closeness centrality (CC), in which catchments with values ≥ 0.14 were more likely to not support the NPH (82% of the cases; Fig. 3). Catchments with lower values
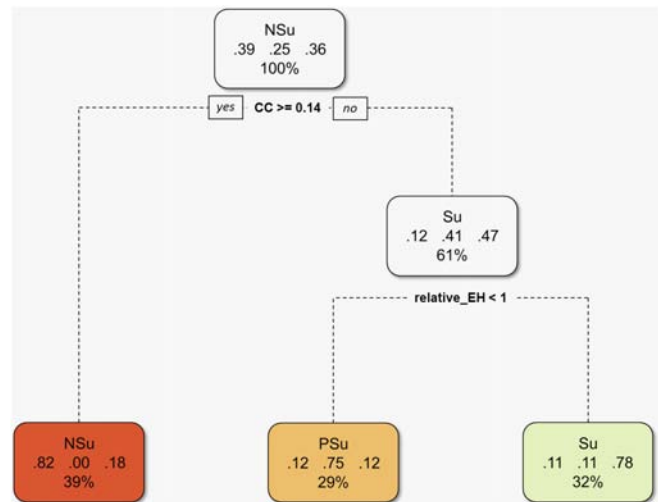

Figure 3. Results for classification trees analysis (CTA) on the predictions of the network position hypothesis (NPH). Splitting of each node was determined by a linear combination of variables where cases that meet the criteria at each junction are split off to the left. For instance, the first split in the tree indicate that all the observations going to the left have a CC ≥ 0.14 whereas the ones branching to the right have a CC < 0.14. The name of each node represents the most probable NPH classification. The three numbers below the name represent the probability of classification *NSu,* *PSu* and *Su* (from left to right) based on the splitting criteria. The last number (%) represents the proportion of observations in that node. For instance, the first terminal node (from left to right) shows that most of the observations are classified as *NSu,* the probability to be classified as *NSu,* *PSu* or *Su* is 82%, 0 and 18%, respectively, and it contains 39% of the total number of observations (i.e. catchments). CC = closeness centrality – centralization; relative_EH = relative environmental heterogeneity; *NSu* = not supported (in red); *PSu* = partially supported (in orange); *Su* = supported (in light green). Refer to the Methods section to understand how these three classifications were assigned.

of CC were further subdivided by relative_EH; catchments with values >1 were more likely to fully support (*Su*) the NPH predictions (78% of the cases), and catchments with relative_EH < 1 were more likely to partially support (*PSu*) the NPH predictions (75% of the catchments; Fig. 3).

Discussion

We evaluated the NPH and were the first to test the validity of this hypothesis for multiple fish metacommunities in river networks. We had three primary findings. First, we found a lack of general support for the NPH predictions regardless of the method considered. Second, when we accounted for the potential influence of dispersal processes through river configuration metrics, we found that the NPH was fully or partially supported in 36 and 25% catchments, respectively, while the support was weaker using the traditional methods (VP_{PCNM}, *Su* = 10%, *PSu* = 25%; Mantel test, *Su* = 21%, *PSu* = 25%; partial Mantel test, *Su* = 21%, *PSu* = 3.5%).

Third, we demonstrated that the relative influence of network position on community assembly could depend on other characteristics such as regional environmental heterogeneity and river network connectivity. We expand on these three findings below.

While the NPH did not receive widespread support in our study, the classification tree analysis provided interesting insight into the drivers of these different responses. For example, the majority of catchments with relatively higher centralization (i.e. regional) closeness centrality (CC) values did not support the NPH predictions. Higher regional CC suggests higher among-site variation in connectivity, meaning that dispersal pathways are constrained to go through a few sites in the river network (Rayfield et al. 2011), which in turn decreases the importance of spatial predictors across communities. Indeed, we found that most catchments that did not support the NPH predictions had non-significant spatial predictors for downstream reaches, while only three of them had non-significant environmental predictors in either headwater or downstream reaches. This suggests that the lack of support for NPH predictions was mostly due to non-significant effects of dispersal on downstream metacommunities rather than the lack of environmental sorting in headwaters. Ultimately, regional network connectivity is crucial for the effect of dispersal to be detected in downstream sites, and could explain why certain studies failed to find support for the NPH even for species that use stream corridors to disperse (Schmera et al. 2018).

Furthermore, spatial connectivity explained a significant proportion of the community variation in 17 downstream metacommunities and, surprisingly, in 12 headwater metacommunities. This may potentially explain why a quarter of the studied catchments were found to partially support the NPH predictions. Moreover, we found that spatial variables explained a higher proportion of the variation than environmental variables in a few headwater metacommunities (e.g. Charente, Durance, Meuse; Fig. 1, Supplementary material Appendix 5 Table A8). Given the dendritic nature of river networks, headwater streams may be geographically close to one another, and consequently experience similar environmental conditions, but also be far from one another in watercourse distance (Fig. 4a). This could influence organisms like fishes that must disperse through the river network (Schmera et al. 2018). In this case, spatial rather than environmental variables would explain the variation in community assemblages across headwater reaches (Fig. 4a). Our findings suggest that even though headwater reaches are more isolated than those further downstream, not all are equally isolated. Hence, the assumption that spatial isolation consequently leads to the dominance of species sorting processes (e.g. Fig. 4b) should be relaxed and potentially further explored. If the variation in isolation influences their community structure (e.g. Fig. 4a) then the signal of spatial processes may reflect dispersal limitation in headwater communities (Thompson and Townsend 2006, Göthe et al. 2017).

We also found that most catchments that were relatively more homogeneously connected (i.e. second split; Fig. 2) but exhibited EH_relative < 1 provided a partial support for the NPH. In these catchments, downstream reaches had more
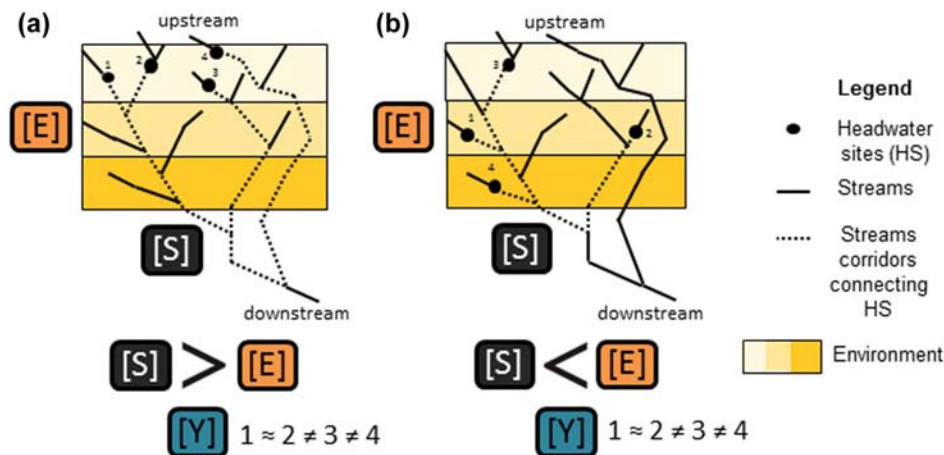

Figure 4. Theoretical example describing how dispersal and environmental factors may influence headwater communities (black circles). The colors in the black boxes represent different environmental conditions. Streams are represented by full and dashed lines. Dashed lines specifically describe the stream corridors along the river network that connect the headwater sites. (a) In the first example, communities 1 and 2 have similar species composition and differ from both communities 3 and 4. The species composition of communities 3 and 4 are also different amongst themselves. The variation in species composition [Y] across the four headwater sites cannot be explained by their environmental conditions [E], which are similar for all sites given that they are geographically near to each other. Rather, how connected these sites are along the river network [S] should better explain the variation in community composition across these communities. (b) In the second example, communities 1 and 2 have similar species composition and differ from both communities 3 and 4. The species composition of communities 3 and 4 are also different amongst themselves. In this case the variation in environmental conditions [E] rather than connectivity [S] should better explain the differences found in their community compositions [Y]. The example (b) but not (a) follows the first NPH prediction, yet both are plausible. These interpretations hold only for taxa that disperse exclusively through stream corridors.

heterogeneous environments than upstream reaches. Given that headwaters were less environmentally heterogeneous, spatial processes explained a significant proportion of the variation in their fish communities, potentially via dispersal limitation (e.g. Fig. 4a). While environmental heterogeneity is often greater in headwaters than further downstream (Clarke et al. 2008), it is not necessarily always the case and should not be assumed (Schmera et al. 2018). In contrast, we observed that the majority of catchments with relative_EH > 1 fully supported the NPH, suggesting that the environmental heterogeneity in headwater sites should be sufficiently high for environmental filtering to be the exclusive assembly process in these communities (Heino et al. 2012). Hence, the results from the regional analysis highlight the importance of considering broad catchment properties into stream metacommunity studies (Sály et al. 2011), and could explain apparent context dependency in communities' response to environmental and spatial processes (Tonkin et al. 2016, 2018).

The two traditional methods, more commonly used in stream metacommunity evaluations, failed as well to provide general support for the NPH predictions. While it is not clear whether this finding is a result of issues with the NPH or with the methods themselves, recent studies have shown that spatial distance alone may not be sufficient to assess the effect of dispersal processes in metacommunities. For example, explicit dispersal measures such as wind directionality (for freshwater zooplankton diapausing eggs; Horváth et al. 2016), oceanic currents (for marine polychaete larvae; Moritz et al. 2013) and landscape resistance within wetlands (for floodplain fishes; Fernandes et al. 2014) have revealed the effect of dispersal processes that would otherwise remain undetected by geographical distance or spatial eigenfunction predictors. Likewise, metrics that better represent the complex structure of dendritic networks should be used to detect dispersal processes in riverine metacommunities, especially for organisms that disperse exclusively through river corridors (Erős et al. 2011, Tonkin et al. 2018).

Our findings suggest that evaluating the NPH solely through the statistical significance of environmental and spatial variables across headwater and downstream sites is inadequate for characterizing the multiple processes that act jointly in stream metacommunities. Perhaps a more nuanced and appropriate approach would be to evaluate relative effect sizes (Stephens et al. 2007); i.e. the change in relative importance of environmental and spatial processes with respect to site network position. For instance, the ratio of spatial compared to environmental variables should increase from headwaters to downstream (e.g. headwaters [S]/[E] < downstream [S]/[E]). However, using our data, applying that approach would still result in only 17 of the 28 catchments supporting the NPH predictions (Supplementary material Appendix 9 Table A15). Nevertheless, interpreting these results remains challenging because methodologists have yet to develop tools that will allow to distinguish unequivocally the effect of dispersal limitation from mass-effects in

[S] on non-experimental research frameworks. Ultimately though, improving our ability to generalize in stream metacommunity theories will require additional investigations, and consideration and evaluation of different facets of biodiversity (Heino et al. 2013, Blanchet et al. 2014, Cilleros et al. 2016, Villéger et al. 2017) in relation to spatial configuration and environmental heterogeneity of river networks.

In conclusion, our evaluation of the NPH with multiple metacommunities did not provide widespread support for this hypothesis, regardless of the methodology used. Even within a taxonomic group that disperses exclusively through stream corridors, the patterns found across the studied metacommunities were diverse. We also demonstrate that the influence of the site network position on community assembly processes may depend on broader characteristics of the river network such as regional environmental heterogeneity and connectivity. Thus, these regional catchment properties should be taken into account when establishing sampling protocols to test general metacommunity theory in river networks (Sály and Erős 2016, Tonkin et al. 2018), or apply its principles to stream conservation or restoration (Swan and Brown 2017).

References

Altermatt, F. et al. 2013. River network properties shape α-diversity and community similarity patterns of aquatic insect communities across major drainage basins. – J. Biogeogr. 40: 2249–2260.

Blanchet, F. G. et al. 2008. Forward selection of explanatory variables. – Ecology 89: 2623–2632.

Blanchet, S. et al. 2014. Regional vs local drivers of phylogenetic and species diversity in stream fish communities. – Freshwater Biol. 59: 450–462.

Borcard, D. et al. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Borcard, D. et al. 2011. Numerical ecology with R. – Springer.

Brown, B. L. and Swan, C. M. 2010. Dendritic network structure constrains metacommunity properties in riverine ecosystems. – J. Anim. Ecol. 79: 571–580.

Buisson, L. et al. 2008. Climate change hastens the turnover of stream fish assemblages. – Global Change Biol. 14: 2232–2248.

Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and contemporary approaches. – Univ. of Chicago Press.

Cilleros, K. et al. 2016. Taxonomic and functional diversity patterns reveal different processes shaping European and Amazonian stream fish assemblages. – J. Biogeogr. 43: 1832–1843.

Clarke, A. et al. 2008. Macroinvertebrate diversity in headwater streams: a review. – Freshwater Biol. 53: 1707–1721.

Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. – Ecol. Lett. 8: 1175–1182.

Cottenie, K. et al. 2003. Zooplankton metacommunity structure: regional vs. local porcesses in highly interconnected ponds. – Ecology 84: 991–1000.

Csardi, H. and Nepusz, T. 2006. The igraph sofwater package for complex network research. – Interjournal, Complex Systems 1695.

Cutler, D. R. et al. 2007. Random forests for classification in ecology. – Ecology 88: 2783–2792.

Datry, T. et al. 2016a. Metacommunity patterns across three Neotropical catchments with varying environmental harshness. – Freshwater Biol. 61: 277–292.

Datry, T. et al. 2016b. Determinants of local and regional communities in intermittent and perennial headwaters of the Bolivian Amazon. – Freshwater Biol. 61: 1335–1349.

De'ath, G. and Fabricius, K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. – Ecology 81: 3178–3192.

Dray, S. et al. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). – Ecol. Model. 196: 483–493.

Dray, S. et al. 2017. ADESPATIAL: multivariate multiscale spatial analysis. – R package ver. 0.0-8.

Erős, T. 2017. Scaling fish metacommunities in stream networks: synthesis and future research avenues. – Commun. Ecol. 18: 72–86.

Erős, T. et al. 2011. Network thinking in riverscape conservation – a graph-based approach. – Biol. Conserv. 144: 184–192.

Erős, T. et al. 2012. Characterizing connectivity relationships in freshwaters using patch-based graphs. – Landscape Ecol. 27: 303–317.

Erős, T. et al. 2017. Effect of landscape context on fish metacommunity structuring in stream networks. – Freshwater Biol. 62: 215–228.

ESRI 2014. ArcMap 10.3. – Environmental Systems Research Inst., USA.

Fagan, W. F. 2002. Connectivity, fragmentation, and extinction risk in dendritic metapopulations. – Ecology 83: 3243–3249.

Fernandes, I. M. et al. 2014. Spatiotemporal dynamics in a seasonal metacommunity structure is predictable: the case of floodplain-fish communities. – Ecography 37: 464–475.

Freeman, L. C. 1979. Centrality in social networks: conceptual clarification. – Social Networks 1: 215–239.

Göthe, E. et al. 2017. Environmental and spatial controls of taxonomic versus trait composition of stream biota. – Freshwater Biol. 62: 397–413.

Griffith, D. A. and Peres-Neto, P. R. 2006. Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. – Ecology 87: 2603–2613.

Heino, J. et al. 2012. Context dependency and metacommunity structuring in boreal headwater streams. – Oikos 121: 537–544.

Heino, J. et al. 2013. A macroecological perspective of trait patterns in stream communities. – Freshwater Biol. 58: 1539–1555.

Heino, J. et al. 2015. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. – Freshwater Biol. 60: 845–869.

Henriques-Silva, R. et al. 2013. A community of metacommunities: exploring patterns in species distributions across large geographical areas. – Ecology 94: 627–639.

Henriques-Silva, R. et al. 2018. Data from: A comprehensive examination of the network position hypothesis across multiple river metacommunities. – Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.bp31rm7>.

Hijmans, R. J. et al. 2005. Very high resolution interpolated climate surfaces for global land areas. – Int. J. Climatol. 25: 1965–1978.

Horváth, Z. et al. 2016. Wind dispersal results in a gradient of dispersal limitation and environmental match among discrete aquatic habitats. – Ecography 39: 726–732.

Jamoneau, A. et al. 2018. Beta diversity of diatom species and ecological guilds: response to environmental and spatial mechanisms along the stream watercourse. – Freshwater Biol. 63: 62–73.

Keith, P. et al. 2011. Les poissons d'eau doce de France. – Collection Inventaires & Biodiversité.

Landeiro, V. L. et al. 2011. Spatial eigenfunction analyses in stream networks: do watercourse and overland distances produce different results? – Freshwater Biol. 56: 1184–1192.

Legendre, P. and Gallagher, E. 2001. Ecologically meaningful transformations for ordination of species data. – Oecologia 129: 271–280.

Leibold, M. A. et al. 2004. The metacommunity concept: a framework for multi-scale community ecology. – Ecol. Lett. 7: 601–613.

Logue, J. B. et al. 2011. Empirical approaches to metacommunities: a review and comparison with theory. – Trends Ecol. Evol. 26: 482–491.

MacArthur, R. H. and Wilson, E. O. 1963. An equilibrium theory of insular zoogeography. – Evolution 17: 373–387.

Moritz, C. et al. 2013. Disentangling the role of connectivity, environmental filtering, and spatial structure on metacommunity dynamics. – Oikos 122: 1401–1410.

Naimi, B. et al. 2014. Where is positional uncertainty a problem for species distribution modelling? – Ecography 37: 191–203.

Oberdorff, T. et al. 2001. A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. – Freshwater Biol. 46: 399–415.

Oksanen, O. et al. 2015. vegan: community ecology package. – R package ver. 2.2-1, <http://CRAN.R-project.org/package=vegan>.

Olden, J. D. et al. 2001. Spatial isolation and fish communities in drainage lakes. – Oecologia 127: 572–585.

Pella, H. et al. 2012. Le réseau hydrographique théorique (RHT) français et ses attributs environnementaux. – Géomorphologie: relief, processus, environnement 18: 317–336.

Peres-Neto, P. R. and Cumming, G. S. 2010. A multi-scale framework for the analysis of fish metacommunities. – In: Gido, K. and Jackson, D. (eds), Community ecology of stream fishes: concepts, approaches, and techniques. American Fisheries Society, pp. 235–262.

Peres-Neto, P. R. et al. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. – Ecology 87: 2614–2625.

Poulet, N. et al. 2011. Time trends in fish populations in metropolitan France: insights from national monitoring data. – J. Fish Biol. 79: 1436–1452.

Pulliam, H. R. 1988. Sources, sinks, and population regulation. – Am. Nat. 132: 652–661.

Rayfield, B. et al. 2011. Connectivity for conservation: a framework to classify network measures. – Ecology 92: 847–858.

Sály, P. and Erős, T. 2016. Effect of field sampling design on variation partitioning in a dendritic stream network. – Ecol. Complex. 28: 187–199.

Sály, P. et al. 2011. The relative influence of spatial context and catchment- and site-scale environmental factors on stream fish assemblages in a human-modified landscape. – Ecol. Freshwater Fish 20: 251–262.

Schmera, D. et al. 2018. Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. – Freshwater Biol. 63: 74–85.

Sharma, S. et al. 2012. Effects of spatial scale and choice of statistical model (linear versus tree-based) on determining species–habitat relationships. – Can. J. Fish. Aquat. Sci. 69: 2095–2111.

Smith, T. W. and Lundholm, J. T. 2010. Variation partitioning as a tool to distinguish between niche and neutral processes. – Ecography 33: 648–655.

Smouse, P. E. et al. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. – Syst. Zool. 35: 627–632.

Snelder, T. H. et al. 2011. Empirical modelling of large scale patterns in river bed surface grain size. – Geomorphology 127: 189–197.

Stephens, P. A. et al. 2007. Inference in ecology and evolution. – Trends Ecol. Evol. 22: 192–197.

Swan, C. M. and Brown, B. L. 2017. Metacommunity theory meets restoration: isolation may mediate how ecological communities respond to stream restoration. – Ecol. Appl. 27: 2209–2219.

Therneau, T. et al. 2015. Recursive partitioning and regression trees. – R package ver. 4.1-11.

Thompson, R. and Townsend, C. 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.

Tonkin, J. D. et al. 2016. Context dependency in biodiversity patterns of central German stream metacommunities. – Freshwater Biol. 61: 607–620.

Tonkin, J. D. et al. 2017. Metacommunity structuring in Himalayan streams over large elevational gradients: the role of dispersal routes and niche characteristics. – J. Biogeogr. 44: 62–74.

Tonkin, J. D. et al. 2018. The role of dispersal in river network metacommunities: patterns, processes, and pathways. – Freshwater Biol. 63: 141–163.

Urban, D. and Keitt, T. H. 2001. Landscape connectivity: a graph-theoretic perspective. – Ecology 82: 1205–1218.

Vidal, J.-P. et al. 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. – Int. J. Climatol. 30: 1627–1644.

Villéger, S. et al. 2017. Functional ecology of fish: current approaches and future challenges. – Aquat. Sci. 79: 783–801.

Vogt, J. et al. 2007. A pan-European river and catchment database. – EC–JRC, Report EUR 22920 EN, Luxembourg.

Williams, G. J. 2011. Data mining with Rattle and R: the art of excavating data fro knowledge discovery. Use R! – Springer.

Zhao, K. et al. 2017. Metacommunity structure of zooplankton in river networks: roles of environmental and spatial factors. – Ecol. Indic. 73: 96–104.

Supplementary material (Appendix ECOG-03908 at <www.ecography.org/appendix/ecog-03908>). Appendix 1–9.