High compositional dissimilarity among small communities is decoupled from environmental variation

Cristina Mariana Jacobi[*], Tadeu Siqueira

Institute of Biosciences, São Paulo State University (UNESP), Avenida 24 A 1515, Rio Claro, SP 13506-900, Brazil

*Corresponding author e-mail: jacobicris@gmail.com

ORCID iDs: 0000-0001-5660-7934; 0000-0001-5069-2904

Abstract. Communities composed of small populations are predicted to be strongly influenced by stochastic demographic events and, thus, less affected by environmental selection than those composed of large populations. However, this prediction has only been tested with computer simulations, simplified controlled experiments, and limited observational data. Here, using multiple datasets on fish abundance in 541 streams we tested (1) if communities composed of small populations are more spatially variable and (2) if they are less related to the environment variation than communities composed of large populations. We used process-based simulations to identify β-diversity metrics and community-environment measures that were appropriate to investigate the role of assembly processes along a gradient of community size. We show that variation in species composition among small communities is higher than among large communities and that the strength of community-environment relationships is weaker in small communities. Our results indicate that community size affects the strength of ecological drift and environmental selection in metacommunities. We thus suggest that further declines in the size of populations and ecosystems can make spatial variation in biodiversity more unpredictable.

Keywords: β-diversity, community size, drift, metacommunity, niche selection

Introduction

Small populations are relatively more prone to be affected by stochastic events than large populations (Otto and Whitlock 1997). While in larger populations stochastic demographic events tend to be balanced by the greater number of individuals, small populations have a higher probability to deviate from mean expectations (Melbourne 2012). For example, a small population can go locally extinct if all its individuals die before reproduction, an event that is less likely to occur in large populations. Thus, species relative abundance in communities composed of small populations tend to vary more in space and time independent of environmental variation – ecological drift (Vellend 2016). Although much progress has been made recently towards the robustness of these predictions, this has been made mainly through simulations (Orrock and Watling 2010), simplified experiments (Gilbert and Levine 2017), and the analysis of limited observational data (Siqueira et al. 2020). Here we expand previous findings by not only focusing on spatial variation in species composition, but also by investigating the role of the environmental factors as drivers of that variation (i.e., community-environment relationship) along a gradient of community size in multiple independent data sets from different regions of the world.

In general, we can assume that while environmental selection deterministically assemble species through density-dependent and independent mechanisms, ecological drift causes variation in local species relative abundance through demographic stochasticity (Vellend 2016). Ecologists are increasingly recognizing that stochastic and deterministic processes interact with each other in many ways to form multispecies communities. For example, experimental research found that small plant communities showed high variation in species relative abundances and high levels of local extinction, resulting in lower local diversity and high variation in species composition (Gilbert and Levine 2017). Observational research found higher spatial variation in species composition among small communities of stream insects that resembled a stochastic model (Siqueira et al. 2020). These results are in line with model simulations that suggest that the effects of environmental selection can be overridden by the effects of ecological drift in small communities (Orrock and Watling 2010).

One prediction regarding the effects of ecological drift on communities that has received less attention concerns the decoupling of spatial variation in species composition and in environmental factors. If the role of ecological drift is strong in small communities, then community-environment relationships should be weak in metacommunities composed of small communities. That is, there should be a positive relationship between the strength of environmental selection and community size. This is not easy to test for various reasons, but mainly because of statistical issues, including lack of sufficient spatial replication (e.g., Siqueira et al. 2020) and zero-inflated data. The data required to test this hypothesis must include various replicated metacommunities from which individual measures of the strength of the community-environment relationship are estimated (e.g., Bini et al. 2014). Here, we used a global database of riverine fish abundances (Comte et al. 2020) to test the idea that community size mediates the roles of ecological drift and environmental selection in driving community spatial variation (β-diversity). Specifically, we analyzed 32 metacommunities in four biogeographic realms (Australasia, Nearctic, Neotropics and Palearctic) that include 541 local communities.

First, we tested if spatial β-diversity was negatively related to community size. Our aim here was to test the generality of the findings by Siqueira et al. (2020) with more comprehensive data and considering recent advances in β-diversity metrics. We expected that β-diversity would be higher among smaller communities and lower among larger communities, even after accounting for the potential effects of environmental heterogeneity and spatial extent. However, quantifying changes in β-diversity along a gradient of community size is challenging because almost all β-diversity metrics depend on both sampling effort and sample size. Thus, we must consider that community size itself can bias the estimates of β-diversity due to differences in the regional species pool size, species abundance distributions (SAD), and the number of individuals (Chase and Knight 2013, Engel et al. 2021). This is undesirable because a theoretically expected negative relationship between β-diversity and community size can be found simply because smaller communities are less likely to share species than larger communities due to sampling effects.

To select a β-diversity metric that was not affected by sampling effects related to community size, we used a process-based simulation model to simulate metacommunities with the same SAD, richness, and number of communities. The simulated metacommunities differed only in their overall size and were assembled without any environmental filter. A null relationship between β-diversity of simulated metacommunities and community size would indicate that the β-diversity metric does not change with community size without invoking environmental selection.

Second, we tested the importance of community size as a mediator of the relative role of environmental selection versus demographic stochasticity by regressing a measure of the strength of the community-environment relationship within each metacommunity against the median community size. Here, we used community incidence data to measure the strength of the community-environment relationship instead of abundance data to avoid that the relationship could increase with community size simply because statistical power increases with abundance (Gwinn et al. 2016). According to previous research (Magurran and Henderson 2003, Coyle et al. 2013), we expected that infrequent and occasional species should have a weak association with the environmental conditions in the community. Thus, we selected the core species of each community to measure the strength of the community-environment relationship, since they are abundant and widespread, and are more likely to be affected by niche-selection (Magurran 2007). We expected that the spatial variation in species composition of larger metacommunities would have a stronger relationship with the environment indicating the preponderance of deterministic environmental selection over demographic stochasticity in communities composed of larger populations. For this second prediction, we also used a process-based simulation model to investigate if the strength of the community-environment relationship would not increase with community size just because larger communities provide more statistical power to detect an association.

Material and Methods

Species data

We used publicly available fish abundance data sampled in riverine ecosystems in different continents, but mainly in Europe and North America (RivFishTIME – Comte et al. 2020). In the RivFishTIME database, sampling sites are organized within river basins. Each site we selected was assigned to a Strahler stream order within the HydroRIVERS network (Lehner and Grill 2013). We also performed a visual inspection of orders attributed to sites located more than 1000m away from the network, changing to the appropriate order, if needed. Then, we defined a community as the set of species occurring in a site and a metacommunity as the set of sites within basin delineations according to HydroBASINS (level 7 - Lehner and Grill 2013). We just considered sites that included direct fish counts and selected the most recent sampling event for each metacommunity that included at least 10 communities in first to third-order streams, excluding basins with less than five species. These procedures resulted in 32 metacommunities. Data selection was made in R (version 4. 2. 1) using the packages mapview (Appelhans et al. 2021) raster (Hijmans 2021), rgdal (Bivand et al. 2021), sf (Pebesma 2018), sp (Pebesma and Bivand 2005, Bivand et al. 2013) and tidyverse (Wickham et al. 2019).

Environmental data

For consistency and because local water variables (e.g., pH, temperature, dissolved oxygen) were not available, we obtained environmental variables relevant to freshwater ecosystems by snapping site geographic coordinates in two global remote-sensing datasets, HydroATLAS (Linke et al. 2019) and free-flowing rivers (FFR) (Grill et al 2019). Both datasets consider a reach as the smallest unit in the hydrological network, being characterized as the line segment between two confluences. Spatial data gathered by HydroATLAS include mean annual natural discharge in m^3/s (Döll et al 2003), average elevation (m) (Robinson et al. 2014), average annual air temperature (ºC), average annual precipitation (mm) (Hijmans et al. 2005), and average human footprint index (Venter et al. 2016) within the local catchment that drains directly into the reach. From the FFR dataset we obtained the connectivity status index (CSI) that was calculated based on five pressure factors (river fragmentation, flow regulation, sediment trapping, water consumption and infrastructure development). These factors are associated with the four dimensions of freshwater connectivity (longitudinal, lateral, vertical and temporal). The CSI provides a characterization of connectivity to every individual river reach, ranging from 0% to 100% connected. We also used a categorical variable (CSI_FF2) based on CSI that indicates if a river reach belongs to a river with free-flowing, good connectivity or degraded status. Together, data from HydroATLAS and FFR were used as our environmental variables since they reflect hydrologic, physiographic, climatic and anthropogenic features of the freshwater environment. These variables were also selected in R using packages mapview, raster, rgdal, sf, sp and tidyverse.

Community descriptors

We measured the size of each community as the number of individuals by site and we called community size the median community size within each metacommunity. Thus, larger metacommunities were those formed by communities with more individuals and smaller metacommunities were those formed by communities with fewer individuals.

We used a process-based simulation model to select a β-diversity metric that was not dependent on sampling effects or directly affected by community size. To do this, we simulated spatially explicit metacommunities without the influence of environmental selection with the package mobsim (May et al. 2021. Simulated metacommunities had the same species abundance distributions (SAD, log-normal), richness, and the number of local communities, but differ in terms of total size (taken from a uniform distribution that varied according to the range of number of individuals sampled in the empirical metacommunities, Supporting Information Table S1). We also represented the intraspecific aggregation caused by dispersal limitation considering that all communities were equally likely to be colonized by the initial member of each species, but once they were, subsequent members of that species were more likely to colonize surrounding locations. Metacommunities were simulated with different numbers of species and sites, according to values in the empirical data (Supporting Information, Table S1). For each of these combinations, we estimated different β-diversity metrics and regressed these β-diversity metrics against community size. Among the many β-diversity metrics we tested, rank-abundance difference was the only one that was not affected by community size in a neutral-like simulation scenario, and thus it was the metric selected to be used in this study. In the supporting information, we describe further metrics that failed the tests (Supporting Information, Table S2).

To describe variation in species rank-abundance patterns for the empirical metacommunities, we measured the median rank difference of each metacommunity as the difference in the rank of species between all pairs of communities within a metacommunity and divided it by the number of unique species in each pair. Rank difference values were bound between 0 and 0.5 (maximum rank change). The R package codyn (Hallett et al. 2020) was used to measure rank-abundance differences. We fitted a multiple linear Gaussian model to describe the relationship between rank differences and community size, including metacommunity environmental heterogeneity and spatial extent as covariates (see Supporting Information Fig. S2 for model fit assumptions). These covariates were added to take into account that larger regions could be more environmentally heterogeneous and have less connected communities that could confound the effects of community size on β-diversity. We quantified environmental heterogeneity within metacommunities using the standardized environmental variables of each community in a distance-based analysis of homogeneity of multivariate dispersions (Euclidian distance). This was done using the betadisper function from the vegan package (Oksanen et al. 2020). To measure the spatial extent of each metacommunity we calculated their centroids and measured the average distance of each of their communities to it. This analysis was also based on Euclidean distance calculated in meters with the geosphere package (Hijmans, 2021).

Community-environment relationships

To investigate the strength of the community-environment relationships within metacommunities and relate to the median community size we also first employed a process-based model to define which statistic would not be affected by community size itself. To do this, we simulated metacommunities that were equally affected by demographic stochasticity and environmental selection regardless of their size. The simulated metacommunities were colonized by species with similar niche breadth (0.2) and that followed a Gaussian response to one environmental variable. Species environmental optimum ranges were taken from a uniform distribution ranging from 0 to 5. We simulated ecological drift by drawing the abundance of every species in each local community from a Poisson distribution whose mean was given by the deterministic influence of environmental selection. We assigned different sizes to the metacommunities estimating species expected abundance in optimum environments from a random interval. Then we modeled the strength of the community-environment relationship within metacommunities. An unbiased measure of the strength of community-environment relationships should not be related to community size.

We tested a Generalized Linear Model (GLM) to relate the strength of the community-environment relationship with community size, but we found that the relationship increased with community size (Fig. S1a, Supporting Information), indicating that the GLM explanatory power increased with community size even for metacommunities assembled by equal contributions of drift and niche selection. Our second approach was to run a Hierarchical Modelling of Species Communities (HMSC) (Ovaskainen 2017), using the package HMSC (Tikhonov et al. 2021). HMSC is a powerful method as it can increase the predictive power identifying biotic interactions and the influence of missing environmental covariates (Tikhonov et al. 2020). The posterior distribution of the HMSC model was sampled with four Markov Chain Monte Carlo (MCMC) chains. Each chain was run for 150.000 iterations, of which the first 50.000 were removed as burn-in. We obtained 1000 samples per chain that were recorded every 100:th step (thin=100) of the iterations. Then, we measured the explanatory power of the model obtaining a median coefficient of discrimination (Tjur R^2) (Tjur 2009) for each simulated metacommunity that was regressed against their community size. We did not find a relationship between the HMSC Tjur R^2 and community size (Fig. S1b, Supporting information). Thus, we used the HMSC approach to answer our question using the same steps described above. We used a probit model to relate the response matrix, composed of the incidence of the core species of each community (10% most abundant), to the matrix of environmental variables. We also examined the MCMC convergence by examining the potential scale reduction factors (Gelman and Rubin 1992) of the model parameters.

Results

As expected, we found a negative relationship between rank-abundance difference and community size (Fig. 1a, Table 1). Neither environmental heterogeneity nor spatial extent were related to this β-diversity metric, indicating that 38.8% of the variation in rank-abundance difference was due uniquely to community size. These results indicate that fish composition in small communities was more different among each other within a metacommunity compared to metacommunities composed of larger communities.

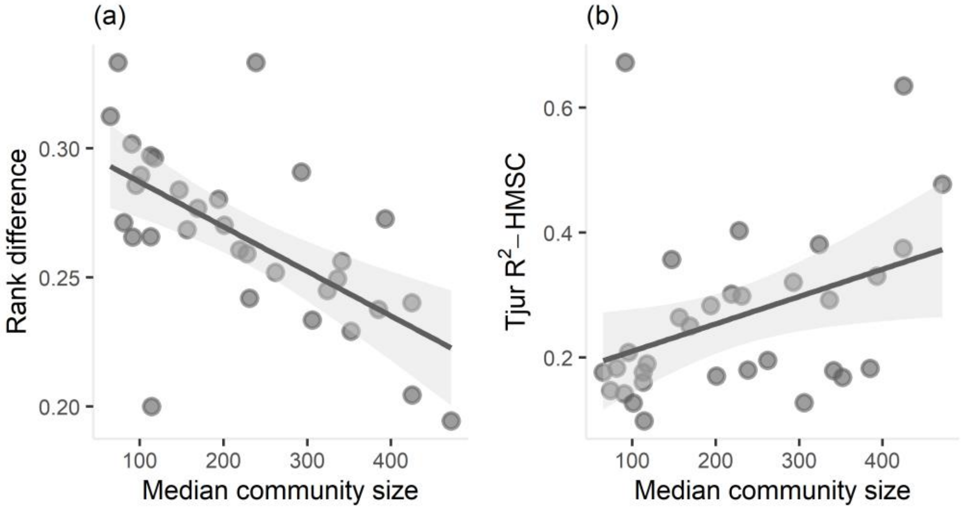

Figure 1. Relationships between a) rank-abundance difference and median community size; and b) between the strength of the community-environment relationship within metacommunities against median community size. Gray bands represent the 95% confidence interval around the regression line.

The MCMC convergence of the HMSC models was satisfactory, with the potential scale reduction factors for the β-parameters varying between 0.999 and 1.014. The coefficient of determination (Tjur R^2) varied between 0.10 and 0.67 (median = 0.20; standard deviation [sd] = 0.14). Considering only the total amount of variance in species abundance explained by environmental variables, average annual air temperature (median = 0.19; sd = 0.05), average elevation (median = 0.18; sd = 0.05), and connectivity status index (CSI; median = 0.16; sd = 0.07) explained the largest proportion of variance across all species (Supporting Information; Table S3). The highest values of explained variance, however, were found for mean annual natural discharge (0.47) and average human footprint index (0.42; Table S3). As predicted, we found a positive relationship between *Tjur* R^2 and community size (Fig. 1b; standardized coefficient = 0.386, t-value = 2.29, R^2 = 0.149, p = 0.029), indicating that the strength of the community-environment relationship was lower in metacommunities composed of smaller communities.

Discussion

The persistence and stability of populations and metapopulations are positively related to their size (Otto and Whitlock 1997, Reed and Hobbs 2004, Terui et al. 2018). Expanding on this perspective and on previous findings (Siqueira et al. 2020) our analyses of intercontinental data support the prediction that spatial variation in species composition among small communities is higher than among large communities. We found support for the prediction that the strength of the relationship between species composition and environmental variation is weak in metacommunities composed of small communities. Taken together, these results indicate that community size mediates the role of ecological drift and environmental selection as drivers of metacommunity spatial dynamics. More specifically, we suggest that ecological drift plays a substantial role in the spatial dynamics of small fish communities, decoupling species-environment relationships. Such decoupling indicates that further declines in the size of fish communities due to environmental change might add more stochasticity to their dynamics, making them less predictable in the face of future environmental scenarios.

Spatial variation in species composition occurs due to changes in various aspects of community structure such as abundance, evenness or richness (Avolio et al. 2019). We found a negative relationship between changes in species distribution abundance (rank-abundance differences) and median community size. This result means that species switched positions in the rank-abundance curve (from rarer to more common and vice-versa) more often in metacommunities composed of small communities than in those composed of large communities. Spatial variation in species abundance within metacommunities occur due to a combination of processes, including dispersal, selection, and demographic stochasticity (Vellend 2016). Our results indicate that whereas changes in species abundance in large communities were mostly because of environmental variation (mainly temperature, elevation, and discharge), changes in species abundance in small communities were mostly caused by stochastic processes. In this sense, our work can be considered an extension of previous work that identified a positive relationship between population size and population stability (Reed and Hobbs 2004) and ecosystem size (and complexity) and metapopulation stability (Terui et al. 2018). Although there have been suggestions of a relationship between spatial and temporal variation in biodiversity properties (e.g., spatial-temporal beta diversity, Steiner and Leibold 2004), the degree to which such correlations are causal is still to be resolved (Stegen et al. 2013).

Random and unpredicted changes in species composition can happen entirely or partially due to demographic events that are independent of environmental selection (Shoemaker 2020, Vellend 2010). For example, smaller communities are more likely to be affected by births and deaths, changing species relative abundance independently of species fitness (Vellend 2010). Previous studies based on mechanistic simulation models suggested that when random demographic events play a major role in community dynamics, species competitive differences become less important to shape community composition and inferior competitors become increasingly represented in local communities

(Orrock and Waltling 2010, Orrock and Fletcher 2005). Siqueira et al. (2020) also found that small insect communities exhibit higher variation in species composition and that it was similar to a null expectation. As community size increases, the role of demographic events that are random regarding species identity should decrease relative to deterministic assembly. In this situation, metacommunity dynamics may be mainly governed by species sorting and patch dynamics (Orrock and James 2010) leading to deterministic exclusion of the inferior competitors and lowering β-diversity in large metacommunities.

The positive relationship between the strength of community-environment coupling and community size indicates that environmental selection plays a minor role in the dynamics of small communities. This relationship further suggests that environmental factors might have limited power to predict species identity or species abundances in small communities. Evidence of weak relationships between species composition and environmental factors is abundant in community ecology. For example, a comprehensive analysis of stream invertebrate communities across the continental United States showed that β-diversity was weakly related to environmental distances in many of the ecoregions studied (Bini et al. 2014). Heino et al. (2015) also assessed the role of environmental factors in driving β-diversity in stream insect communities in different regions of the world and found that environmental variables were poor predictors of species composition, explaining on average 13% of the variation in community spatial variation. Although we did not re-analyze data from previous studies reporting weak community-environment relationships, our results suggest that, besides the already known causes (e.g., “snapshot” sampling, lack of statistical power, unaccounted legacy effects), ecological drift acting more strongly upon small communities might be an additional cause for such weak relationships.

As it is usually the case in observational, correlative studies, ours also have limitations. Stream communities can be structured by many environmental factors that vary at the local scale, such as nutrient concentration, turbidity, pH (Allan et al. 2021). The set of variables we used is not necessarily the best one to represent community-environment relationships, as these variables were not measured at the same spatial scale used to sample communities. But, even with such data, we detected a positive relationship between the strength of the community-environment relationship and community size. We thus speculate that the inclusion of additional local environmental variables could affect the observed relationship in two ways: (1) the model explanatory power will increase for all metacommunities as more variables are added to it. In this case, the relationship would remain positive, with a similar slope but higher intercept. (2) The role of stochasticity might be so strong in small communities that the strength of the community-environment relationship would not increase for them, but only for the large communities. In this case, the slope of the regression line would change to a more positive relationship, while the intercept would remain the same. Despite these limitations, some of our models explained more than 30% (and in some cases up to 60%) of the variance in abundance. This result indicates that large scale variables such as average annual air temperature, mean annual natural discharge, and average human footprint index can be used to explain variation in fish abundance (e.g., Chen and Olden 2020, Comte et al. 2021), especially for species in large communities. Another issue to bear in mind is that by relying solely on the median, we might have missed how other aspects of community size influence β-diversity and community-environment relationships. For example, other aspects of the distribution of population sizes within communities (e.g., range; standard deviation) could represent different processes that interact with demographic stochasticity – an issue that deserves further investigation.

Relying on a robust data set that spans 541 communities, across 32 river basins in four biogeographic realms, and robust methods to quantify spatial variation in species composition, we provide evidence for a general pattern in nature – species composition varies more among small communities than among large communities. This pattern is likely driven by the stronger influence of demographic stochastic events on small than on larger communities, as the dynamics of small communities are weakly linked to spatial variation in environmental factors. On the other hand, metacommunities composed of large communities seem to be assembled more deterministically as their species composition were more related to environmental variation. As the effects of demographic stochasticity should manifest also, maybe even more strongly, on community temporal dynamics, future studies should analyze the relationship between temporal β-diversity and community size. We also suggest the prevalence of the relationships we describe here should be tested with data from terrestrial metacommunities, where community size might be more directly linked to ecosystem size. In a changing world where ecosystems have been fragmented and reduced, high variation in species composition among small communities that decoupled from environmental variation points to further challenges to ecosystem management.

References

Allan, J. D., Castillo, M. M. and Capps K. A. 2021. Stream Ecology: Structure and Function of Running Waters. *Springer Nature Switzerland AG*.

Appelhans, T., Detsch, F., Reudenbach, C. and Woullauer, S. 2021. mapview: Interactive Viewing of Spatial Data in R. R package version 2.10.0. - <https://CRAN.R-project.org/package=mapview>.

Avolio, M. L., Carroll, I. T., Collins, S. L., Houseman, G. R., Hallett, L. M., Isbell, F., Koerner, S. E., Komatsu, K. J., Smith, M. D. and Wilcox, K. R. 2019. A comprehensive approach to analyzing community dynamics using rank abundance curves. *Ecosphere*, 10: e02881.

Bini, L. M., Landeiro, V. L., Padial, A. A., Siqueira, T. and Heino, J. 2014. Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States. *Ecology*, 95: 1569-1578.

Bivand, R., Keitt, T. and Rowlingson, B. 2021. rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-23. - <https://CRAN.R-project.org/package=rgdal>.

Bivand, R., Pebesma, E. and Gomez-Rubio. 2013. Applied spatial data analysis with R. Second edition, Springer, NY. <https://asdar-book.org/>.

Chase, J. M. and Knight, T. M. 2013. Scale-dependent effect sizes of ecological drivers on biodiversity: Why standardised sampling is not enough. *Ecol. Lett.*, *16*: 17-26.

Chen, k., Olden, J. D. 2020. Threshold responses of riverine fish communities to land use conversion across regions of the world. *Glob. Change Biol*., 26: 4952-4965.

Comte, L., Carvajal-Quintero, J., Tedesco, P. A., Giam, X., Brose, U., Erõs, T., Filipe, A. F., Fortin, M. J., Irving, K., Jacquet, C., Larsen, S., Sharma, S., Ruhi, A., Becker, F. G., Casatti, L., Castaldelli, G., Dala-Corte, R. B., Davenport, S. R., Franssen, N. R., García-Berthou, E., Gavioli, A., Gido, K. B., Jimenez-Segura, L., Leitão, R. P., McLarney, B., Meador, J., Milardi, M., Moffatt, D. B., Occhi, T. V. T., Pompeu, P. S., Propst, D. L., Pyron, M., Salvador, G. N., Stefferud, J. A., Sutela, T., Taylor, C., Terui, A., Urabe, H., Vehanen, T., Vitule, J. R. S., Zeni, J. O. and Olden, J. D. 2020. RivFishTIME: A global database of fish time-series to study global change ecology in riverine systems. *Glob. Ecol. Biogeogr*., 30: 38-50.

Comte, L., Olden, J. D., Tedesco, P. A., Ruhi, A. and Giam, X., 2021. Climate and land-use changes interact to drive long-term reorganization of riverine fish communities globally. *Proc. Natl. Acad. Sci.*, 118: e2011639118.

Coyle, J. R., Hurlbert, A. H. and White, E. P. 2013. Opposing Mechanisms Drive Richness Patterns of Core and Transient Bird Species. *Am Nat*., 181:4.

Döll, P., Kaspar, F. and Lehner, B. 2003. A global hydrological model for deriving water availability indicators: model tuning and validation. *J. Hydrol.*, 270: 105-134.

Engel, T., Blowes, S. A., McGlinn, D. J., May, F., Gotelli, N. J., McGill, B. J. and Chase, J. M. 2021. Using coverage-based rarefaction to infer non-random species distributions. *Ecosphere,* 12:e03745.

Gelman, A. and Rubin, D. B. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci*., 7: 457-472.

Gilbert, N. and Levine, J. M. 2017. Ecological drift and the distribution of species diversity. *Proc. R. S. B.,* 284: 20170507.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., Babu, S., Borrelli, P., Cheng, L., Crochetiere, H., Macedo, H. E., Filgueiras, R., Goichot, M., Higgins, J., Hogan, Z., Lip, B.,, McClain, M. E., Meng, J., Mulligan, M., Nilsson, C., Olden, J. D., Opperman, J. J., Petry, P., Liermann, C. R., Sáenz, L., Salinas-Rodrpiguez, S., Schelle, P., Schmitt, R. J. P., Snider, J., Tan, F., Tockner, K., Valdujo, P. H., van Soesbergen, A. and Zarfl, C. 2019. Mapping the world's free-flowing rivers. *Nature,* 569: 215-221.

Gwinn, D. C., Allen, M. S., Bonvechio, K. I., Hoyer, M. V. and Beesley, L. S. 2016. Evaluating estimators of species richness: the importance of considering statistical error rates. *Methods Ecol Evol*, 7: 294-302.

Hallett, L., Avolio, M. L., Carroll, I. T., Jones, S. K., MacDonald, A. A. M., Flynn, D. F. B., Slaughter, P., Ripplinger, J., Collins, S. L., Gries, C. and Jones, M. B. 2020. *codyn: Community Dynamics Metrics*. R package version 2.0.5.

Heino, J., Melo, A. S., Bini, L. M., Altermatt, F., Al-Shami, S. A., Angeler, D. G., Bonada, N., Brand, C., Callisto, M., Cottenie, K., Dangles, O., Dudgeon, D., Encalada, A., Göthe, E., Grönroos, M., Hamada, N., Jacobsen, D., Landeiro, V. L., Ligeiro, R., Martins, R. T., Miserendino, M. L., Rawi, C. S. Md., Rodrigues, M. E., Roque, F. O., Sandin, L., Schmera, D., Sgarbi, L. F., Simaika, J. P., Siqueira, T., Thompson, R. M. and Townsend, C. R. 2015. A comparative analysis reveals weak relationships between ecological factors and beta diversity of stream insect metacommunities at two spatial levels*. Ecol. Evol*., 5: 1235-1248.

Hijmans, R. 2021. _geosphere: Spherical Trigonometry_. R package version 1.5-14, <https://CRAN.R-project.org/package=geosphere>.

Hijmans, R.J. 2021. raster: Geographic Data Analysis and Modeling. R package version 3.4-13.

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

Lehner, B. and Grill, G. 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol*. *Process*., 27: 2171-2186.

Linke, S., Lehner, B., Dallaire, C. O., Ariwi, J., Gril, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F. and Thieme, M. 2019. Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Sci. Data*, 6: 283.

Magurran, A. and Henderson, P. 2003. Explaining the excess of rare species in natural species abundance distributions. *Nature,* 422: 714-716.

Magurran, A. E. 2007. Species abundance distributions over time. *Ecol.Lett.*, 10: 347-354.

May, F. 2021. mobsim: Spatial Simulation and Scale-Dependent Analysis of Biodiversity Changes. R package version 0.2.0, <https://github.com/MoBiodiv/mobsim>.

Melbourne, B. A. 2012. Stochasticity, Demographic. Encyclopedia of Theoretical Ecology, edited by Alan Hastings and Louis Gross, Berkeley: University of California Press, 706-712.

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan Stier, A., Ter Braak, C. and Weedon, J. 2020. vegan: Community Ecology Package. R package version 2: 5-7. <https://CRAN.R-project.org/package=vegan>.

Orrock, J. L. and Fletcher, R. J. Jr. 2005. Changes in community size affect the outcome of competition. *Am Nat*., 166: 107-11.

Orrock, J. L. and Watling, J. I. 2010. Local community size mediates ecological drift and competition in metacommunities. *Proc. R. S. B*., 277: 2185-2191.

Otto, S. P. and Whitlock, M. C. 1997. Fixation of beneficial mutations in a population of changing size. *Genetics*, 146: 723-733.

Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin, T. and Abrego, N. 2017 How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecol. Lett.*, 20: 561-576.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10: 439-446.

Pebesma, E. J. and Bivand, R. S. 2005. Classes and methods for spatial data in R. *R News*, 5: 9-13. <https://cran.r-project.org/doc/Rnews/>.

Reed, D. H. and Hobbs, G. R. 2004. The relationship between population size and temporal variability in population size. *Anim. Conserv. forum.*, 7: 1-8.

Robinson, N., Regetz, J. and Guralnick, R. P. 2014. EarthEnv-DEM90: a nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. ISPRS J. *Photogramm. Remote. Sens.*, 87: 57-67.

Shoemaker, L. G., Sullivan, L. L., Donohue, I., Cabral, J. S., Williams, R. J., Mayfield, M. M., Chase, J. M., Chu, C., Stanley Harpole, W., Huth, A., HilleRisLambers, J., James, A. R. M., Kraft, N. J. B., May, Felix., Muthukrishnan, Satterlee, S., Taubert, F., Wang, X., Wiegand., Yang, Q. and. Abbott, K. C. 2020. Integrating the underlying structure of stochasticity into Community ecology. *Ecology*, 101: e02922.

Siqueira, T., Saito, V. S., Bini, L. M., Melo. A. S., Petsch, D. K., Landeiro, V. L., Tolonen, K. T., Jyrkänkallio-Mikkola, J., Soininen and Heino, J. 2020. Community size can affect the signals of ecological drift and niche selection on biodiversity. *Ecology*, 101: e03014.

Steiner, C. F. and Leibold, M. A. 2004. Cyclic assembly trajectories and scale-dependent productivity-diversity relationships. *Ecology*, 85: 107-113.

Stegen, J. C., Freestone, A. L., Crist, T. O., Anderson, M. J., Chase, J. M., Comita, L. S., Cornell, H. V., Davies, K. F., Harrison, S. P., Hurlbert, A. H., Inouye, B. D., Kraft, N. J. B., Myers, J. A., Sanders, N. J., Swenson, N. G. and Vellend, M. 2013. Drivers of spatial and temporal turnover. *Glob. Ecol. Biogeogr.*, 22: 202-212.

Terui, A., Ishiyama, N., Urabe, H., Ono, S., Finlay, J. O. and Nakamura, F. 2018. Metapopulation stability in branching river networks. *P. Natl. Acad. Sci. USA*, 115: E5963-E5969.

Tikhonov, G., Opedal, Ø.H., Abrego, L., Lehikoinen, A., de Jonge, M. M. J., Oksanen, J. and Ovaskainen, O. 2020. Joint species distribution modelling with the R-package HMSC. *Methods Ecol. Evol.*, 11: 442-447.

Tikhonov, G., Ovaskainen, O., Oksanen, J., de Jonge, M., Opedal, O. and Dallas, T. 2021. hmsc: Hierarchical Model of Species Communities. R package version 3.0-11. <https://CRAN.R-project.org/package=Hmsc>.

Tjur, T. 2009. Coefficients of determination in logistic regression models—A new proposal: the coefficient of discrimination. *Am. Stat.*, 63: 366-372.

Vellend, M. and Agrawal, A. 2010. Conceptual synthesis in community ecology. *Q. Rev. Biol.*, 85: 183-206.

Vellend, M. 2016. The theory of ecological communities. Mpb series: Edition 57. *Princeton University Press*, Princeton, New Jersey, USA.

Venter, O., Sanderson, E. W., Magrach, A., Allan, J. R., Beher, J., Jones, K. R., Possingham, H. P., Laurance, W. F., Wood, P., Fekete, B. M., Levy, M. A. and Watsom, J. E. M. 2016. Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data*, 3: 160067.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Lin Pedersen, T., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K. and Yutani, H. 2019. Welcome to the tidyverse. *J. of Open Source Softw.*, 4: 1686.

Table 1. Model statistics of the relationship of rank-abundance difference with community size (CS), environmental variation (EV) and spatial extent (SE) in riverine fish metacommunities. SD coeff. = standardized regression coefficient; R^2 and Adj. R^2 = coefficient of determination and adjusted coefficient of determination, respectively.

Predictor	**SD coeff.**	**t-value**	**p-value**	**R^2**	**Adj R^2**
CS	-0.657	-4.291	<0.001	0.388	0.368
EV	0.044	0.303	0.764		
SE	0.112	0.732	0.470		

Data Accessibility

R codes and data used here are publicly available on Zenodo: https://doi.org/10.5281/zenodo.6504817

Acknowledgements

We thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) - Finance Code 001 for funding. C.M.J. is also grateful to Luciano N. Lima for the support in spatial analysis. T.S. was supported by grant #21/00619-7, São Paulo Research Foundation (FAPESP) and by grant #309496/2021-7, Brazilian National Council for Scientific and Technological Development (CNPq).

Supporting Information

1. Material and methods

Table S1. General features of each dataset used in the empirical analyses at both regional and local scales.

Regional				Local			
Metacommunity ID	**# sites**	**# individuals**	**γ-diversity**	**α-diversity (mean)**	**α-diversity (median)**	**# individuals (mean)**	**# individuals (median)**
2070021230	10	4538	12	7	6	454	471.5
2070021360	15	5559	15	7	7	371	393
2070469050	12	5398	19	5	2	450	238.5
2070470930	10	15956	21	8	8	1596	293
2070513970	29	4112	32	7	7	142	101
5070070110	29	4340	31	9	8	150	113
5070070390	12	3636	23	10	10	303	336
5070070440	21	4307	30	9	8	205	194
5070070450	13	3990	26	10	11	307	324
5070468810	10	2087	14	5	4	209	91.5
6070015200	22	2411	17	7	8	110	114
6070709820	16	4374	27	9	8	273	113
6070713620	22	3161	40	10	9	144	95.5
7070042610	12	2633	44	18	18	219	201
7070317870	11	3323	41	15	14	302	219
7070327640	12	756	6	2	2	63	65
7070360890	31	2460	7	2	2	79	74
7070415980	10	2971	33	10	8	297	117.5
7070450470	13	1220	21	7	7	94	81
7070451570	18	2882	27	9	9	160	147
7070453370	32	3967	31	7	6	124	90.5
7070620210	21	8395	47	20	20	400	306
7070630060	11	5356	34	16	16	487	425
7070642870	21	8429	36	14	15	401	352
7070643030	11	2704	34	15	15	246	228
7070647250	18	5814	46	16	17	323	262
7070647350	14	3651	42	16	15	261	231
7070648760	10	3928	44	18	18	393	424.5
7070651920	12	1978	41	18	19	165	156.5
7070656320	22	8021	43	16	16	365	341
7070663220	23	10421	29	15	15	453	385
7070684690	18	4501	48	17	16	250	169

1.1 Process-based simulation models

1.1.1 Spatial variation in species composition

As our first goal was to investigate how spatial variation in species composition within metacommunities varied along a gradient of median community size, we needed to make sure that our metrics representing spatial variation in species composition would not be affected by sampling effort and sample size. For example, larger communities have a higher probability to share more species and thus exhibit a lower variation in species composition than smaller communities because they have more individuals, but not necessarily because they are more similar.

So, using the package mobsim (May et al 2018), we simulated metacommunities with the same SAD (log-normal distribution), species richness and number of communities. They were assembled without any environmental filter and differed only in their overall size (taken from a uniform distribution that varied between 756 and 15956 individuals, according to the range of individuals sampled in the empirical metacommunities (Table S1). We represented the intraspecific aggregation caused by dispersal limitation considering that all communities were equally likely to be colonized by the initial member of each species, but once they were, subsequent members of that species were more likely to colonize surrounding locations. To investigate the potential effect of metacommunity characteristics in the tested metrics, we varied the number of species and sites in the simulations. The values of these parameters covered the range of variation of the empirical metacommunities (6-48 species and 10-32 sites, Table S1).

We then measured six β-diversity metrics within each simulated metacommunity and regressed β-diversity values against the median community size of the simulated metacommunities. Because we hypothesized that environmental selection should be stronger in larger communities and ecological drift should be stronger in smaller communities, we expected that an unbiased β-diversity metric would have no relationship with community size in the simulation, as environmental selection played no role in community assembly.

Our first candidate β-diversity metric was Baselga’s (2010) multiple-site Total Beta Diversity based on presence-absence data. This metric accounts both for the spatial turnover and the nestedness components of β-diversity. It was carried out in R with the function beta.multi (Baselga et al. 2022), using Jaccard dissimilarity. The second candidate β-diversity metric was C-score (Stone & Roberts 1990), which provides an average measure of co-occurrence for all species pairs. Higher C-score values indicate fewer co-occurrence between species in the matrix. The C-score was measured with the package bipartite (Dormann et al. 2008) in R. Next, we used two β-diversity metrics that could be corrected for under-sampling: the Jaccard-Chao and the Shannon. The correction applied to these metrics uses accumulation curves and should be more effective than the original metrics to remove γ-dependence (Cao et al 2021). The fifth candidate metric was a recently developed metric called βc (Engel et al 2021), which was developed to represent differences in species composition even when there are differences in the size of the species pool and associated sampling effects.

Finally, as changes in species relative abundances are a good predictor of community change (Dornelas et al. 2014, Jones et al. 2017, Avolio et al. 2019), we measured how the rank-abundance of communities varied among each other within each metacommunity, and regressed their median against median community size.

1.1.2 Community-environment relationship

To investigate our second prediction, we had to be sure that a measure of the strength of the community-environment relationship would not itself be a statistical consequence of community size. For example, it is possible that a strong community-environment relationship emerges in large communities simply because it might be easier to model more individuals than a matrix full of zeros, irrespective of the underlying assembly process. Thus, we used a second process-based simulation in which metacommunities of different sizes were assembled by the same amount of demographic stochasticity and environmental selection. Now, we simulated metacommunities composed of species that followed a Gaussian response to a single environmental variable. They had the same niche breadth (0.2) but differed in their environmental optima (i.e., niche position; taken from a uniform distribution ranging from 0 to 5). We simulated ecological drift by drawing the abundance of every species in each local community from a Poisson distribution which mean was given by the deterministic influence of environmental selection. We assigned different sizes to the metacommunities estimating species expected abundance in optimum environments from a random interval.

We estimated the strength of the community-environment relationship, but only considering the incidence of the 10% most abundant species of each community. We tested two modeling approaches to estimate the strength of the community-environment relationship. First, we used a generalized linear model (GLM) and pseudo R^2 values as a measure of the strength of the community-environment relationship. We then used Hierarchical Modeling of Species Communities (HMSC) (Ovaskainen 2017) and Tjur R^2 values as a measure of the strength of the community-environment relationship. An appropriate measure of community-environment relationship would show no relationship with community size in simulated metacommunities under equal contribution of drift and selection.

2. Results

2.1 Spatial variation in species composition

We found that the first five (except rank-abundance difference) β-diversity metrics were related (varying from negative to positive) to median community size (Table S2). These results indicate that these five β-diversity metrics are affected by median community size independently of the underlying assembly process.

We observed no relationship between rank-abundance difference and community size for the range of parameter values in our simulations (Table S2). So, as the lack of relationship was consistent, we concluded that rank difference was the most appropriate β-diversity metric that we could use in our study to investigate the relationship between variation in species composition and community size.

Table S2. Model statistics of the relationships between β-diversity metrics (Jaccard, C-score, Jaccard-Chao, Shannon, βc and rank difference) and community size for 30 simulated metacommunities covering the range of variation of species and sites from the empirical metacommunities. SD coeff. = standardized regression coefficient; R^2 and Adj. R^2 = coefficient of determination and adjusted coefficient of determination, respectively. Some metrics were not calculated when the number of species and sites were small (empty lines).

β-diversity metric	**SD coeff.**	**F-statistic**	**p-value**	**R^2**	**Adj R^2**
Metacommunities simulated with 6 species and sites					
Jaccard - Baselga					
C-score					
Jaccard-Chao					
Shannon	-0.316	3.114	0.088	0.100	0.068
Bc					
Rank difference	-0.152	0.664	0.422	0.023	-0.012
Metacommunities simulated with 10 species and sites					
Jaccard - Baselga	-0.749	35.7	<0.001	0.560	0.545
C-score					
Jaccard-Chao	-0.617	17.17	<0.001	0.380	0.358
Shannon	-0.176	0.894	0.352	0.031	-0.004
Bc	-0.583	11.85	0.002	0.34	0.311
Rank difference	0.073	0.150	0.702	0.005	-0.030
Metacommunities simulated with 20 species and sites					
Jaccard - Baselga	-0.857	77.17	<0.001	0.734	0.724
C-score	0.350	3.903	0.058	0.122	0.091
Jaccard-Chao	-0.706	27.88	<0.001	0.499	0.481
Shannon	-0.346	3.82	0.061	0.120	0.089
Bc	-0.642	19.68	<0.001	0.413	0.3918
Rank difference	0.221	1.434	0.241	0.049	0.015
Metacommunities simulated with 30 species and sites					
Jaccard - Baselga	-0.927	171.3	<0.001	0.859	0.854
C-score	-0.819	56.95	<0.001	0.670	0.659
Jaccard-Chao	-0.919	152.1	<0.001	0.844	0.839
Shannon	-0.737	33.21	<0.001	0.543	0.526
Bc	-0.226	1.513	0.229	0.052	0.017
Rank difference	0.252	1.894	0.180	0.063	0.030
Metacommunities simulated with 40 species and sites					
Jaccard - Baselga	-0.954	282.3	<0.001	0.910	0.906
C-score	-0.895	112.2	<0.001	0.800	0.793
Jaccard-Chao	-0.938	204.4	<0.001	0.879	0.875
Shannon	-0.674	23.26	<0.001	0.454	0.434
Bc	0.513	9.996	0.004	0.263	0.237
Rank difference	0.144	0.595	0.447	0.021	-0.014
Metacommunities simulated with 48 species and sites					
Jaccard - Baselga	-0.929	175.7	<0.001	0.863	0.858

β-diversity metric	SD coeff.	F-statistic	p-value	R^2	Adj R^2
C-score	-0.870	87.57	<0.001	0.758	0.749
Jaccard-Chao	-0.930	179.2	<0.001	0.865	0.860
Shannon	-0.783	44.38	<0.001	0.613	0.599
Bc	0.569	11.47	0.002	0.323	0.295
Rank difference	-0.069	0.136	0.715	0.005	-0.031

2.2 Community-environment relationships

We found that the GLM pseudo-R2 increased with community size (SD. coeff. = 0.678, R^2 = 0.459, p = <0.001, Fig. S1a), whereas there was no evidence of a relationship between HMSC Tjur R^2 and community size (SD. coeff. = 0.247, R^2 = 0.061, p = 0.188, Fig. S1b). So, we decided to use the HMSC to test our second prediction.

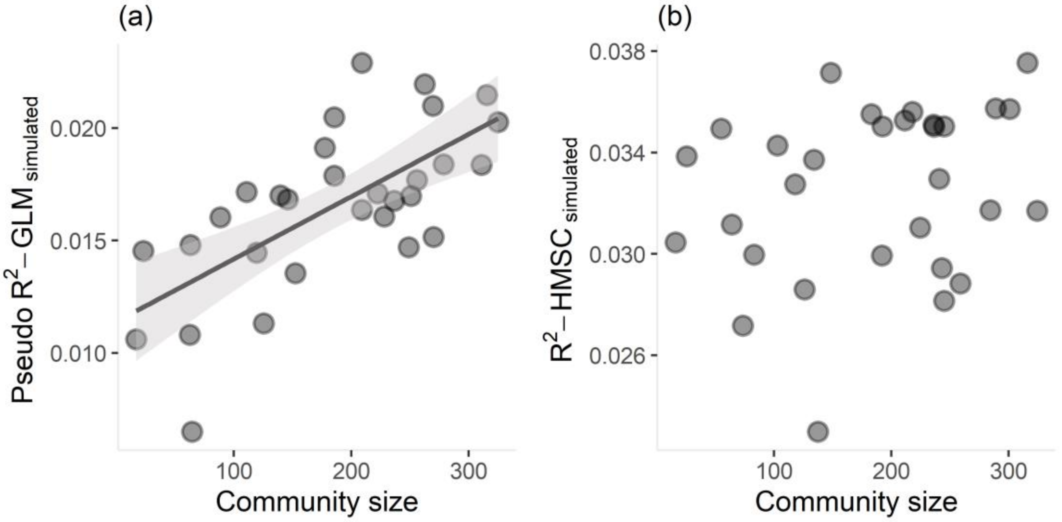

Figure S1. Relationship of the strength of the community-environment relationships within simulated metacommunities measured with a Generalized Linear Model (GLM) (a) and with Hierarchical Modelling of Species Communities (HMSC) (b) against median community size. Gray bands represent the 95% confidence interval around the regression line.

2.3 Model fit assumptions

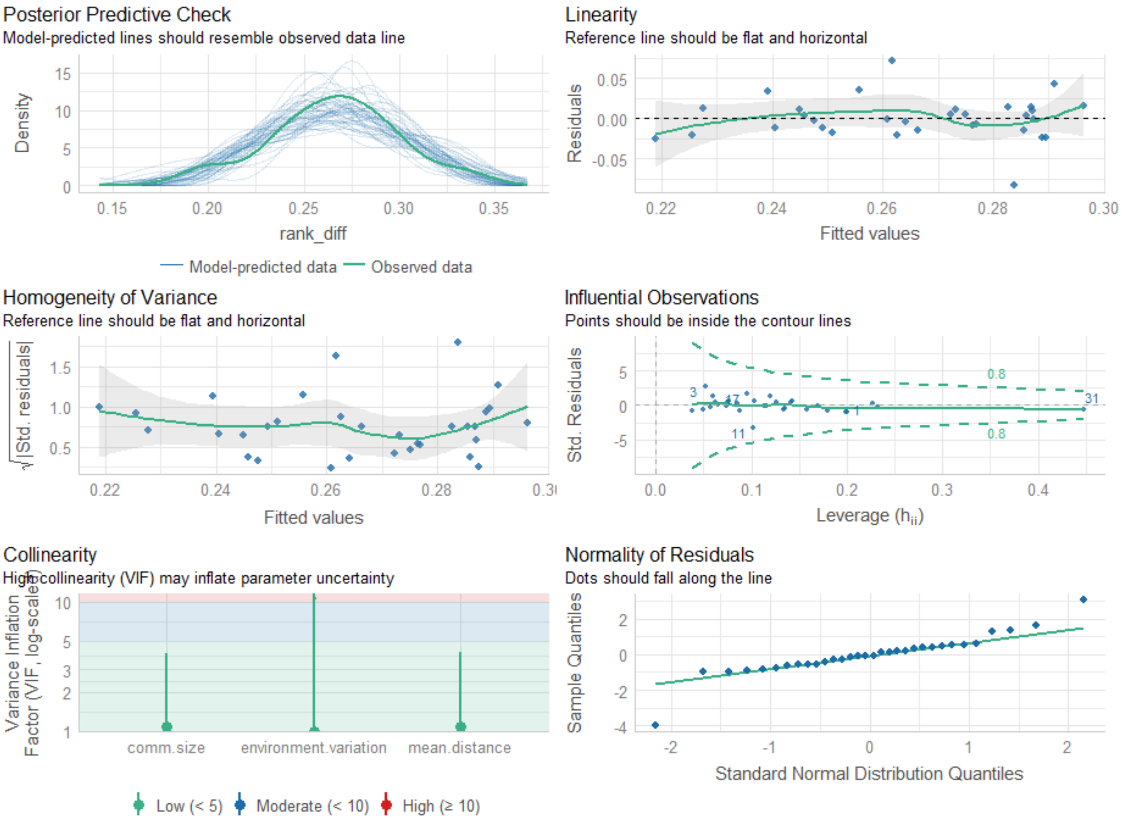

Figure S2. Model fit assumptions of the multiple linear model used to describe rank difference values as a function of community size, environmental heterogeneity and spatial extent of the metacommunities. All plots were done using the performance package (Lüdecke et al. 2021) in R.

2.4 Relative importance of each environmental variable to explain community composition

Table S3. Partitioning species' variance explained by the environmental variables across all metacommunities, considering the 10% most abundant species of each community. Environmental variables: discharge (Dis), elevation (Ele), temperature (Tmp), precipitation (Pre), human footprint (hft), connectivity status index (csi) and a measure that indicates if a river reach belongs to a river with free-flowing, good connectivity or degraded status (csiff2).

Species	Dis	Ele	Tmp	Pre	hft	csi	csiff2
Metacommunity ID 2070021230							
Cottus gobio	0.151	0.206	0.220	0.113	0.223	0.086	0.000
Salmo trutta	0.119	0.164	0.194	0.130	0.274	0.119	0.000
Phoxinus phoxinus	0.168	0.215	0.276	0.104	0.156	0.081	0.000
Metacommunity ID 2070021360							
Salmo trutta	0.335	0.106	0.162	0.069	0.056	0.235	0.037
Cottus gobio	0.209	0.134	0.159	0.061	0.134	0.239	0.065
Phoxinus phoxinus	0.183	0.195	0.185	0.090	0.111	0.178	0.058
Barbatula barbatula	0.146	0.140	0.307	0.144	0.072	0.141	0.049
Perca fluviatilis	0.155	0.142	0.191	0.146	0.103	0.203	0.061
Metacommunity ID 2070469050							
Salmo trutta	0.171	0.182	0.193	0.168	0.141	0.112	0.033
Salmo salar	0.119	0.194	0.193	0.142	0.130	0.188	0.035
Cottus gobio	0.171	0.204	0.165	0.173	0.138	0.109	0.040
Barbatula barbatula	0.110	0.197	0.206	0.152	0.112	0.183	0.038
Squalius cephalus	0.168	0.186	0.173	0.164	0.141	0.103	0.064
Metacommunity ID 2070470930							
Cottus gobio	0.291	0.145	0.183	0.123	0.132	0.081	0.044
Phoxinus phoxinus	0.327	0.118	0.168	0.151	0.122	0.077	0.036
Barbatula barbatula	0.243	0.148	0.154	0.176	0.148	0.090	0.041
Perca fluviatilis	0.200	0.164	0.164	0.156	0.136	0.129	0.051
Squalius cephalus	0.187	0.161	0.176	0.161	0.123	0.146	0.047
Metacommunity ID 2070513970							
Barbatula barbatula	0.154	0.277	0.286	0.134	0.075	0.074	0.000
Gobio gobio	0.105	0.239	0.245	0.127	0.118	0.166	0.000
Phoxinus phoxinus	0.137	0.281	0.281	0.129	0.080	0.092	0.000
Rutilus rutilus	0.054	0.247	0.215	0.157	0.239	0.087	0.000
Squalius cephalus	0.118	0.236	0.247	0.117	0.109	0.173	0.000
Rhodeus sericeus	0.085	0.210	0.317	0.111	0.195	0.081	0.000
Alburnus alburnus	0.166	0.200	0.200	0.155	0.109	0.170	0.000

Abramis brama	0.105	0.221	0.207	0.177	0.111	0.178	0.000
Leucaspius delineatus	0.136	0.203	0.198	0.142	0.184	0.138	0.000
Umbra krameri	0.111	0.246	0.251	0.128	0.130	0.134	0.000
Metacommunity ID 5070070110							
Gobiomorphus australis	0.134	0.152	0.167	0.063	0.117	0.307	0.059
Hypseleotris compressa	0.150	0.392	0.090	0.174	0.058	0.093	0.043
Gambusia holbrooki	0.214	0.227	0.115	0.054	0.095	0.211	0.084
Retropinna semoni	0.134	0.122	0.224	0.055	0.068	0.340	0.056
Mugil cephalus	0.160	0.125	0.175	0.071	0.097	0.301	0.071
Hypseleotris galii	0.196	0.137	0.157	0.080	0.073	0.286	0.072
Pseudomugil signifer	0.157	0.173	0.139	0.090	0.064	0.307	0.070
Hypseleotris klunzingeri	0.146	0.186	0.145	0.078	0.072	0.300	0.072
Melanotaenia duboulayi	0.135	0.097	0.177	0.104	0.168	0.254	0.064
Anguilla reinhardtii	0.166	0.123	0.136	0.094	0.092	0.316	0.072
Mogurnda adspersa	0.135	0.125	0.133	0.105	0.061	0.339	0.102
Metacommunity ID 5070070390							
Xiphophorus maculatus	0.118	0.161	0.139	0.098	0.140	0.213	0.132
Gambusia holbrooki	0.130	0.152	0.139	0.062	0.094	0.388	0.035
Oreochromis mossambicus	0.264	0.168	0.139	0.083	0.120	0.138	0.089
Hypseleotris klunzingeri	0.144	0.171	0.142	0.078	0.141	0.234	0.091
Mogurnda adspersa	0.151	0.180	0.150	0.131	0.139	0.120	0.128
Hypseleotris compressa	0.106	0.170	0.120	0.076	0.144	0.284	0.099
Melanotaenia duboulayi	0.153	0.168	0.133	0.127	0.121	0.215	0.083
Metacommunity ID 5070070440							
Philypnodon grandiceps	0.139	0.164	0.193	0.091	0.125	0.233	0.055
Retropinna semoni	0.137	0.161	0.217	0.055	0.284	0.093	0.054
Gambusia holbrooki	0.182	0.185	0.159	0.096	0.096	0.261	0.021
Pseudomugil signifer	0.136	0.179	0.167	0.068	0.100	0.303	0.046
Melanotaenia duboulayi	0.152	0.148	0.165	0.111	0.181	0.204	0.039
Hypseleotris klunzingeri	0.127	0.176	0.175	0.074	0.125	0.248	0.074
Trachystoma petardi	0.136	0.182	0.175	0.081	0.112	0.261	0.052
Anguilla reinhardtii	0.154	0.158	0.214	0.061	0.116	0.240	0.058
Hypseleotris compressa	0.168	0.206	0.180	0.066	0.091	0.255	0.034
Xiphophorus hellerii	0.160	0.170	0.168	0.067	0.107	0.278	0.050
Metacommunity ID 5070070450							
Retropinna semoni	0.141	0.192	0.200	0.052	0.169	0.216	0.030
Gambusia holbrooki	0.152	0.202	0.180	0.110	0.184	0.137	0.034
Hypseleotris compressa	0.158	0.175	0.164	0.171	0.142	0.157	0.034
Hypseleotris klunzingeri	0.191	0.182	0.179	0.092	0.137	0.164	0.056
Hypseleotris galii	0.173	0.192	0.176	0.096	0.147	0.179	0.037
Melanotaenia duboulayi	0.176	0.214	0.196	0.081	0.136	0.155	0.042
Metacommunity ID 5070468810							
Mogurnda adspersa	0.092	0.141	0.139	0.156	0.419	0.054	0.000
Hypseleotris galii	0.162	0.123	0.120	0.206	0.293	0.095	0.000
Gambusia holbrooki	0.148	0.227	0.226	0.109	0.130	0.160	0.000
Metacommunity ID 6070015200							

Rineloricaria aequalicuspis	0.087	0.265	0.271	0.228	0.084	0.065	0.000
Cyanocharax itaimbe	0.096	0.244	0.226	0.245	0.089	0.100	0.000
Pareiorhaphis hypselurus	0.084	0.274	0.274	0.229	0.073	0.066	0.000
Pareiorhaphis nudulus	0.176	0.249	0.266	0.206	0.056	0.047	0.000
Epactionotus bilineatus	0.099	0.219	0.232	0.212	0.139	0.099	0.000
Metacommunity ID 6070709820							
Poecilia reticulata	0.467	0.163	0.072	0.103	0.118	0.078	0.000
Knodus moenkhausii	0.238	0.198	0.050	0.097	0.129	0.288	0.000
Hypostomus ancistroides	0.333	0.154	0.095	0.184	0.138	0.096	0.000
Astyanax lacustris	0.373	0.149	0.124	0.105	0.175	0.075	0.000
Aspidoras fuscoguttatus	0.275	0.156	0.094	0.154	0.137	0.185	0.000
Gymnotus sylvius	0.282	0.170	0.107	0.113	0.118	0.211	0.000
Moenkhausia sanctaefilomenae	0.295	0.183	0.101	0.104	0.122	0.196	0.000
Serrapinnus notomelas	0.260	0.256	0.096	0.116	0.156	0.116	0.000
Metacommunity ID 6070713620							
Piabina argentea	0.134	0.151	0.156	0.162	0.221	0.176	0.000
Hypostomus ancistroides	0.147	0.160	0.187	0.188	0.195	0.123	0.000
Knodus moenkhausii	0.265	0.164	0.233	0.192	0.072	0.074	0.000
Corydoras aeneus	0.162	0.153	0.222	0.154	0.131	0.177	0.000
Astyanax lacustris	0.099	0.163	0.249	0.202	0.192	0.095	0.000
Poecilia reticulata	0.092	0.142	0.129	0.150	0.288	0.199	0.000
Oligosarcus pintoi	0.196	0.163	0.155	0.150	0.144	0.192	0.000
Gymnotus sylvius	0.130	0.156	0.238	0.158	0.130	0.187	0.000
Serrapinnus notomelas	0.125	0.255	0.141	0.188	0.128	0.163	0.000
Hemigrammus marginatus	0.158	0.159	0.162	0.185	0.126	0.210	0.000
Imparfinis schubarti	0.125	0.135	0.110	0.184	0.214	0.233	0.000
Metacommunity ID 7070042610							
Luxilus albeolus	0.093	0.134	0.208	0.342	0.089	0.134	0.000
Etheostoma nigrum	0.121	0.172	0.201	0.250	0.136	0.120	0.000
Nocomis leptocephalus	0.096	0.272	0.190	0.253	0.089	0.099	0.000
Notropis altipinnis	0.111	0.143	0.201	0.271	0.110	0.163	0.000
Lepomis auritus	0.087	0.173	0.192	0.223	0.135	0.190	0.000
Etheostoma olmstedi	0.214	0.142	0.180	0.173	0.120	0.173	0.000
Anguilla rostrata	0.216	0.160	0.194	0.161	0.104	0.166	0.000
Lythrurus matutinus	0.165	0.202	0.178	0.220	0.089	0.145	0.000
Noturus insignis	0.124	0.155	0.206	0.236	0.123	0.156	0.000
Notropis procne	0.094	0.149	0.192	0.295	0.088	0.182	0.000
Cyprinella analostana	0.103	0.166	0.193	0.254	0.104	0.180	0.000
Enneacanthus gloriosus	0.114	0.153	0.248	0.155	0.125	0.204	0.000
Metacommunity ID 7070317870							
Percina maculata	0.204	0.065	0.080	0.147	0.338	0.113	0.052
Notropis dorsalis	0.191	0.093	0.130	0.176	0.178	0.118	0.115
Etheostoma nigrum	0.219	0.110	0.097	0.133	0.125	0.197	0.120
Margariscus margarita	0.205	0.068	0.143	0.150	0.144	0.209	0.082
Luxilus cornutus	0.219	0.077	0.071	0.197	0.156	0.215	0.065
Notropis volucellus	0.204	0.092	0.115	0.171	0.130	0.168	0.120

Notemigonus crysoleucas	0.199	0.075	0.111	0.125	0.152	0.220	0.116
Chrosomus eos	0.205	0.104	0.075	0.145	0.173	0.215	0.083
Metacommunity ID 7070327640							
Oncorhynchus clarkii	0.120	0.274	0.232	0.092	0.204	0.079	0.000
Salvelinus confluentus	0.133	0.217	0.171	0.102	0.287	0.091	0.000
Salmo trutta	0.171	0.244	0.186	0.107	0.182	0.111	0.000
Salvelinus fontinalis	0.238	0.229	0.174	0.113	0.125	0.121	0.000
Metacommunity ID 7070360890							
Prosopium williamsoni	0.106	0.234	0.282	0.177	0.086	0.115	0.000
Oncorhynchus clarkii	0.109	0.175	0.194	0.076	0.229	0.217	0.000
Salmo trutta	0.058	0.188	0.201	0.156	0.240	0.157	0.000
Salvelinus confluentus	0.125	0.203	0.186	0.154	0.168	0.164	0.000
Salvelinus fontinalis	0.144	0.171	0.247	0.077	0.144	0.217	0.000
Oncorhynchus mykiss	0.098	0.255	0.314	0.135	0.101	0.098	0.000
Metacommunity ID 7070415980							
Rhinichthys atratulus	0.176	0.108	0.104	0.112	0.234	0.266	0.000
Campostoma anomalum	0.102	0.126	0.153	0.140	0.202	0.276	0.000
Notropis stramineus	0.105	0.127	0.148	0.141	0.207	0.271	0.000
Pimephales promelas	0.103	0.143	0.128	0.216	0.164	0.246	0.000
Notropis dorsalis	0.157	0.157	0.099	0.252	0.136	0.199	0.000
Semotilus atromaculatus	0.162	0.140	0.096	0.142	0.227	0.234	0.000
Hybognathus hankinsoni	0.213	0.140	0.102	0.171	0.152	0.222	0.000
Catostomus commersonii	0.209	0.126	0.114	0.179	0.154	0.218	0.000
Culaea inconstans	0.142	0.184	0.111	0.153	0.152	0.258	0.000
Umbra limi	0.172	0.140	0.109	0.173	0.171	0.235	0.000
Lepomis cyanellus	0.170	0.128	0.123	0.180	0.163	0.236	0.000
Metacommunity ID 7070450470							
Etheostoma olmstedi	0.084	0.198	0.226	0.208	0.145	0.138	0.000
Etheostoma nigrum	0.089	0.197	0.212	0.219	0.146	0.137	0.000
Semotilus atromaculatus	0.048	0.206	0.233	0.237	0.139	0.138	0.000
Cottus bairdii	0.105	0.183	0.204	0.210	0.159	0.138	0.000
Rhinichthys cataractae	0.151	0.186	0.167	0.163	0.198	0.135	0.000
Rhinichthys obtusus	0.067	0.175	0.178	0.179	0.228	0.173	0.000
Metacommunity ID 7070451570							
Campostoma anomalum	0.118	0.249	0.200	0.263	0.090	0.080	0.000
Etheostoma caeruleum	0.085	0.244	0.200	0.269	0.100	0.102	0.000
Semotilus atromaculatus	0.130	0.249	0.254	0.226	0.073	0.068	0.000
Catostomus commersonii	0.127	0.220	0.221	0.213	0.130	0.089	0.000
Rhinichthys obtusus	0.168	0.247	0.243	0.194	0.094	0.055	0.000
Rhinichthys cataractae	0.182	0.208	0.208	0.163	0.092	0.148	0.000
Metacommunity ID 7070453370							
Catostomus commersonii	0.092	0.171	0.190	0.227	0.179	0.104	0.038
Rhinichthys cataractae	0.137	0.237	0.243	0.162	0.067	0.132	0.022
Semotilus atromaculatus	0.192	0.179	0.221	0.267	0.064	0.059	0.018
Pimephales promelas	0.214	0.195	0.229	0.138	0.115	0.083	0.026
Lepomis cyanellus	0.211	0.196	0.217	0.145	0.120	0.084	0.027

Notropis atherinoides	0.140	0.211	0.226	0.200	0.079	0.115	0.029
Rhinichthys obtusus	0.213	0.196	0.205	0.090	0.069	0.203	0.025
Pimephales notatus	0.085	0.161	0.169	0.155	0.155	0.253	0.023
Etheostoma nigrum	0.203	0.153	0.180	0.100	0.123	0.218	0.023
Cottus bairdii	0.206	0.181	0.225	0.109	0.079	0.174	0.026
Salvelinus fontinalis	0.132	0.206	0.242	0.116	0.112	0.164	0.028
Metacommunity ID 7070620210							
Noturus insignis	0.082	0.279	0.200	0.251	0.092	0.065	0.031
Nocomis leptocephalus	0.086	0.282	0.226	0.257	0.064	0.062	0.023
Etheostoma flabellare	0.177	0.216	0.244	0.228	0.062	0.052	0.021
Notropis chiliticus	0.116	0.213	0.198	0.229	0.083	0.126	0.035
Luxilus cerasinus	0.087	0.237	0.239	0.258	0.086	0.068	0.025
Notropis procne	0.077	0.257	0.255	0.251	0.059	0.072	0.029
Lepomis auritus	0.088	0.233	0.276	0.248	0.064	0.067	0.024
Lepomis cyanellus	0.127	0.217	0.224	0.243	0.085	0.072	0.034
Etheostoma nigrum	0.097	0.238	0.191	0.258	0.081	0.102	0.034
Cyprinella analostana	0.099	0.257	0.252	0.256	0.046	0.067	0.024
Lepomis macrochirus	0.109	0.227	0.240	0.230	0.092	0.070	0.031
Metacommunity ID 7070630060							
Notropis chiliticus	0.310	0.146	0.091	0.170	0.091	0.177	0.014
Nocomis leptocephalus	0.144	0.224	0.223	0.113	0.102	0.177	0.016
Luxilus coccogenis	0.227	0.157	0.141	0.146	0.135	0.171	0.024
Clinostomus funduloides	0.195	0.169	0.123	0.218	0.147	0.120	0.029
Noturus insignis	0.126	0.167	0.126	0.101	0.221	0.236	0.023
Metacommunity ID 7070642870							
Cottus bairdii	0.099	0.170	0.106	0.157	0.104	0.284	0.081
Nocomis micropogon	0.078	0.145	0.211	0.214	0.104	0.192	0.057
Rhinichthys atratulus	0.127	0.216	0.133	0.132	0.107	0.185	0.100
Luxilus coccogenis	0.114	0.248	0.183	0.100	0.127	0.131	0.097
Notropis lutipinnis	0.171	0.164	0.143	0.134	0.134	0.196	0.058
Rhinichthys cataractae	0.143	0.226	0.143	0.190	0.107	0.095	0.095
Percina evides	0.105	0.267	0.128	0.135	0.187	0.107	0.071
Campostoma anomalum	0.115	0.194	0.141	0.167	0.130	0.106	0.147
Notropis leuciodus	0.149	0.216	0.190	0.130	0.133	0.120	0.062
Metacommunity ID 7070643030							
Nocomis leptocephalus	0.146	0.194	0.285	0.240	0.054	0.080	0.000
Clinostomus funduloides	0.132	0.172	0.270	0.185	0.114	0.126	0.000
Lepomis auritus	0.064	0.203	0.170	0.191	0.155	0.218	0.000
Etheostoma olmstedi	0.249	0.189	0.182	0.168	0.085	0.127	0.000
Lepomis macrochirus	0.190	0.214	0.187	0.179	0.139	0.090	0.000
Cyprinella pyrrhomelas	0.260	0.152	0.201	0.182	0.109	0.096	0.000
Notropis chiliticus	0.194	0.189	0.233	0.170	0.091	0.122	0.000
Metacommunity ID 77070647250							
Nocomis leptocephalus	0.390	0.078	0.060	0.083	0.052	0.303	0.033
Luxilus albeolus	0.304	0.137	0.107	0.114	0.111	0.167	0.059
Lepomis auritus	0.176	0.099	0.225	0.164	0.137	0.148	0.050

Luxilus cerasinus	0.341	0.101	0.088	0.109	0.127	0.182	0.054
Fundulus rathbuni	0.205	0.169	0.119	0.119	0.077	0.253	0.058
Semotilus atromaculatus	0.300	0.171	0.109	0.123	0.099	0.135	0.062
Gambusia holbrooki	0.290	0.122	0.186	0.115	0.088	0.151	0.046
Etheostoma olmstedi	0.363	0.120	0.079	0.175	0.061	0.148	0.054
Notropis procne	0.218	0.118	0.145	0.132	0.179	0.159	0.050
Lepomis macrochirus	0.229	0.153	0.094	0.116	0.177	0.185	0.046
Notropis hudsonius	0.235	0.145	0.134	0.146	0.090	0.192	0.058
Lythrurus ardens	0.229	0.147	0.133	0.143	0.088	0.205	0.055
Notropis altipinnis	0.250	0.129	0.115	0.130	0.143	0.172	0.060
Metacommunity ID 7070647350							
Etheostoma olmstedi	0.071	0.177	0.115	0.170	0.128	0.339	0.000
Erimyzon oblongus	0.094	0.138	0.148	0.111	0.226	0.284	0.000
Nocomis leptocephalus	0.108	0.156	0.176	0.089	0.193	0.277	0.000
Notropis altipinnis	0.063	0.129	0.082	0.249	0.202	0.275	0.000
Lepomis auritus	0.100	0.184	0.136	0.108	0.172	0.300	0.000
Notropis alborus	0.078	0.143	0.220	0.143	0.166	0.251	0.000
Lepomis macrochirus	0.300	0.141	0.185	0.100	0.146	0.129	0.000
Lepomis cyanellus	0.095	0.171	0.132	0.178	0.134	0.290	0.000
Gambusia holbrooki	0.134	0.181	0.171	0.137	0.163	0.213	0.000
Metacommunity ID 7070648760							
Lepomis macrochirus	0.313	0.136	0.149	0.090	0.127	0.131	0.053
Lepomis auritus	0.271	0.107	0.146	0.133	0.166	0.131	0.046
Nocomis leptocephalus	0.278	0.171	0.256	0.067	0.074	0.118	0.036
Noturus insignis	0.273	0.130	0.225	0.077	0.087	0.143	0.066
Luxilus albeolus	0.319	0.153	0.139	0.072	0.117	0.140	0.060
Etheostoma flabellare	0.241	0.130	0.174	0.114	0.116	0.161	0.065
Nocomis raneyi	0.222	0.177	0.143	0.096	0.096	0.190	0.078
Cyprinella analostana	0.217	0.242	0.133	0.128	0.084	0.133	0.063
Notropis procne	0.209	0.250	0.138	0.127	0.081	0.134	0.059
Metacommunity ID 7070651920							
Notropis cummingsae	0.053	0.114	0.294	0.135	0.117	0.287	0.000
Etheostoma olmstedi	0.130	0.168	0.174	0.157	0.126	0.245	0.000
Notemigonus crysoleucas	0.099	0.148	0.202	0.157	0.128	0.266	0.000
Lepomis auritus	0.081	0.138	0.303	0.102	0.108	0.268	0.000
Gambusia holbrooki	0.123	0.138	0.222	0.160	0.248	0.109	0.000
Lepomis macrochirus	0.133	0.121	0.221	0.148	0.140	0.236	0.000
Noturus insignis	0.108	0.183	0.182	0.145	0.128	0.253	0.000
Anguilla rostrata	0.151	0.149	0.166	0.139	0.179	0.216	0.000
Notropis altipinnis	0.096	0.202	0.191	0.145	0.145	0.222	0.000
Metacommunity ID 7070656320							
Luxilus coccogenis	0.106	0.203	0.279	0.156	0.155	0.101	0.000
Lepomis auritus	0.084	0.224	0.220	0.246	0.149	0.078	0.000
Nocomis leptocephalus	0.049	0.224	0.238	0.330	0.083	0.076	0.000
Semotilus atromaculatus	0.130	0.223	0.183	0.199	0.173	0.093	0.000
Campostoma anomalum	0.323	0.156	0.221	0.159	0.077	0.063	0.000

Cyprinella chloristia	0.117	0.311	0.240	0.165	0.084	0.083	0.000
Fundulus rathbuni	0.104	0.232	0.255	0.235	0.071	0.104	0.000
Etheostoma olmstedi	0.151	0.244	0.197	0.203	0.103	0.102	0.000
Lepomis macrochirus	0.189	0.228	0.211	0.174	0.112	0.086	0.000
Notropis chlorocephalus	0.062	0.183	0.196	0.383	0.119	0.056	0.000
Clinostomus funduloides	0.186	0.206	0.193	0.157	0.113	0.145	0.000
Metacommunity ID 7070663220							
Nocomis leptocephalus	0.236	0.226	0.248	0.085	0.079	0.127	0.000
Etheostoma brevispinum	0.085	0.320	0.241	0.162	0.112	0.079	0.000
Cyprinella chloristia	0.077	0.309	0.198	0.105	0.220	0.091	0.000
Cyprinella pyrrhomelas	0.094	0.197	0.236	0.145	0.265	0.063	0.000
Hybopsis hypsinotus	0.095	0.220	0.202	0.106	0.185	0.191	0.000
Semotilus atromaculatus	0.110	0.313	0.230	0.151	0.114	0.082	0.000
Lepomis auritus	0.112	0.255	0.208	0.239	0.120	0.066	0.000
Notropis scepticus	0.082	0.280	0.222	0.163	0.169	0.083	0.000
Metacommunity ID 7070684690							
Semotilus atromaculatus	0.128	0.196	0.237	0.086	0.102	0.214	0.038
Nocomis leptocephalus	0.067	0.276	0.127	0.042	0.081	0.393	0.014
Notropis chiliticus	0.125	0.323	0.127	0.041	0.034	0.336	0.014
Notropis altipinnis	0.163	0.211	0.157	0.090	0.088	0.253	0.039
Etheostoma olmstedi	0.073	0.235	0.261	0.114	0.059	0.233	0.024
Lepomis auritus	0.186	0.186	0.206	0.048	0.076	0.273	0.025
Perca flavescens	0.119	0.236	0.227	0.085	0.156	0.121	0.056
Noturus insignis	0.123	0.234	0.220	0.085	0.156	0.125	0.057
Lepomis marginatus	0.126	0.199	0.224	0.104	0.055	0.253	0.040
Semotilus lumbee	0.172	0.171	0.195	0.083	0.140	0.201	0.039
Notropis cummingsae	0.158	0.191	0.237	0.071	0.065	0.241	0.038
Lepomis macrochirus	0.105	0.158	0.208	0.110	0.160	0.221	0.036
Aphredoderus sayanus	0.112	0.149	0.201	0.121	0.173	0.207	0.037

References

Avolio, M. L, Carroll, I. T., Collins, S. L., Houseman, G. R., Hallett, L. M., Isbell, F., Koerner, S. E., Komatsu, K. J., Smith, M. D. and Wilcox, K. R. 2019. A comprehensive approach to analyzing community dynamics using rank abundance curves. *Ecosphere*, 10: e02881.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.*, 19: 134-143.

Baselga, A., Orme, D., Villeger, S., De Bortoli, J., Leprieur, F. and Logez, M. 2022._betapart: Partitioning Beta Diversity into Turnover and Nestedness Components_. R package version 1.5.6. <https://CRAN.R-project.org/package=betapart>.

Cao, K., Svenning, J. -C., Yan, C., Zhang, J., Mi, X. and Ma, K. 2021. Undersampling correction methods to control γ-dependence for comparing β-diversity between regions. *Ecology*, 102: 9.

Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C. and Magurran, A. E. 2014. Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344: 296-9.

Dormann, C. F., Gruber, B. and Fruend, J. 2008. Introducing the bipartite Package: Analysing Ecological Networks. R news Vol 8/2, 8-11.

Engel, T., Blowes, S. A., McGlinn, D. J., May, F., Gotelli, N. J., McGill, B. J. and Chase, J. 2021. Using coverage-based rarefaction to infer non-random species distributions. *Ecosphere*, 12:e03745.

Jones, S. K., Ripplinger, J. and Collins, S. L. 2017. Species reordering, not changes in richness, drives long-term dynamics in grassland communities. *Ecol. Lett.*, 20: 1556-1565.

Lüdecke, D., Mattan, S., Ben-Shachar, Patil, I., Waggoner, P. and Makowski, D. 2021. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *J. Open Source Software*, 6(60): 3139.

May, F., Gerstner, K., McGlinn, D. J., Xiao, X. and Chase, J. M. 2018. mobsim: An R package for the simulation and measurement of biodiversity across spatial scales. *Methods Ecol Evol*., 9: 1401-1408.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. and Abrego, N. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecol. Lett.*, 20: 561-576.

Stone, L. and Roberts, A. 1990. The checkerboard score and species distributions. Oecologia 85: 74–79.