High compositional dissimilarity among small communities is decoupled from environmental variation: a multicontinental analysis of riverine fish

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. Small communities are predicted to be strongly influenced by stochastic demographic events and, thus, less affected by environmental selection than large communities. However, this prediction has only been tested with computer simulations, simplified controlled experiments, and limited observational data. Using multicontinental data on riverine fish and considering recent advances in β-diversity metrics, we tested if communities composed of small populations are more spatially variable and more affected by ecological drift than communities composed of large populations. We show that variation in species composition among small communities was higher than among large communities and similar to stochastic assembly. We also show that the strength of species-environment relationships is weaker in small communities. Our results indicate that community size can affect the strength of ecological drift and environmental selection in metacommunities, and that further declines in the size of populations and ecosystems can make spatial variation in biodiversity more unpredictable.

Key words: β-diversity, community size, drift, metacommunity, selection

Introduction

Small populations are relatively more prone to be affected by stochastic events than large populations [1]. For example, a small population can go locally extinct if all its individuals die before reproducing, 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 [2]. Although much progress has been made recently towards the robustness of these predictions, this has been made mainly through simulations [3], simplified experiments [4], and the analysis of limited observational data [5]. Here we expand previous findings by not only focusing on spatial variation in species composition, but also by investigating the strength of community-environment coupling 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 together through density dependent and independent mechanisms, ecological drift causes variation in local species relative abundance through demographic stochasticity [2]. 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 [4]. Observational research found higher spatial β-diversity among small communities of stream insects that was like a stochastic model [5]. These results are in line with model simulations that suggest that the effects of environmental selection can be override by the effects of ecological drift in small communities [3].

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 lack of statistical power [e.g., 5]. The data required to test this hypothesis must be composed of various replicated metacommunities from which individual measures of the strength of community-environment relationship is estimated [e.g., 6].

Here, we test the hypothesis that community size mediates the roles of ecological drift and environmental selection in driving community spatial variation in metacommunities. To reach our goal we used a global database of riverine fish abundances [7]. Specifically, we analyzed 2332 local communities, across 113 river basins in four biogeographic realms (Australasia, Nearctic, Neotropics and Palearctic).

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*. [5] with more compressive data and considering recent advances in β-diversity metrics. To do this, we regressed β-diversity values against metacommunity size. We expected that β-diversity would be higher among smaller communities and lower among larger communities. As we used data gathered from different studies, which had different goals and sampling methods, we calculated two metrics of β-diversity (β-Shannon and β-Jaccard-Chao) that are less affected by undersampling and γ-dependence than other indices, making them a better option to compare β-diversity across metacommunities in different regions [8]. In addition, we estimated how evenness, richness and species rank-abundance differed among communities and regressed them with community size. These metrics quantify how communities vary in space in their species relative abundance and the number of species and thus can be useful to elucidate β-diversity. We also expected a greater variability of evenness, richness, and rank-abundance among smaller communities.

Second, we tested if observed relationships between β-diversity and community size were like those expected under ecological drift. To do this, we simulated β-diversity values in a null model entirely based on a scenario of ecological drift. We expected that the deviation of observed to simulated β-diversity would be closer to zero among smaller communities than among larger communities, indicating greater similarity to random assembly. Third, we tested the role 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 community size. We expected that 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.

Material and Methods

Species data

We used publicly available fish abundance data sampled in riverine ecosystems in different continents, but mainly Europe and North America (RivFishTIME [7]). In the RivFishTIME data base, sampling sites are organized within river basins. Each site we selected was assigned to a Strahler stream order within the HydroRIVERS network [9]. 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 [9]). We 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 113 metacommunities, being 81 with density data (individuals/100 m^2) and 32 with count data. All analyzes described below were ran separately for these different abundance types. Data selection was made in R [version 1.4.1717; 10] using the packages mapview [11] raster [12], rgdal [13], sf [14], sp [15, 16] and tidyverse [17]

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 [18] and free-flowing rivers (FFR) [19]. 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 [20], average of elevation (m) [21], average of annual air temperature (ºC), average of annual precipitation (mm) [22], and average of the human footprint index [23] 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 in 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 defined community size as the total number of individuals sampled in a site and metacommunity size as the median of all community sizes within a metacommunity. So, metacommunities with larger communities were those with more individuals per site.

To measure β-diversity we considered that the sites were sampled with different methods or with different sampling effort. Not considering this can be problematic because any comparisons among metacommunities may be affected by undersampling and regional diversity dependence (γ-dependence [24]). For this reason, null-models have been frequently used to account for γ-dependence of β-diversity metrics [25, 26]. However, using simulated and empirical data, Cao *et al*. [8] showed that undersampling corrections using diversity accumulation curves are more effective to remove γ-dependence. They found that the corrected β-Shannon diversity index was least dependent on γ-diversity and that the corrected Jaccard-Chao index applied to null model removed γ-dependence more effectively than either the correction alone or the null model alone. So, to deal with undersampling and regional diversity dependence we used two approaches: first, we calculated the corrected β-Shannon diversity using a diversity accumulation curve by asymptotically estimating both true α– and γ-Shannon diversity [8, 27]. The second approach was based on the corrected Jaccard–Chao index. The β-Shannon and β-Jaccard-Chao indexes were calculated for count (β-Shannon$_{count}$, β-Jaccard-Chao$_{count}$) and density (β-Shannon$_{dens}$, β-Jaccard-Chao$_{dens}$) data. Thus, we obtained one β-diversity value of each index for each metacommunity and regressed these values against community size.

To get a complementary understanding of how species composition differed within metacommunities, we also estimated how Simpson's evenness, species richness and rank-abundance differed among communities [28] within metacommunities and regressed these estimates against community size. We measured evenness for each community and then we calculated the median evenness within each metacommunity, that could range from 0 (uneven community) to 1 (perfectly even community). Richness difference was calculated comparing the difference of the number of species between all pairs of communities inside a metacommunity divided by the number of unique species in each pair. From these pairwise values we calculated the median richness difference of each metacommunity, where larger values indicate greater changes in species richness among communities. Rank-abundance difference measures how species relative abundances change among communities relative to each other. We calculated this metric measuring the difference of the rank of species between all pairs of communities within a metacommunity divided by the number of unique species in each pair. Rank-abundance difference values were bound between

0 and 0.5 (maximum rank change). The package entropart [29] was used to estimate β-Shannon index while the β-Jaccard-Chao index was estimated with vegan [30]. The package codyn [31] was used to estimate differences in evenness, richness, and rank-abundance.

Null models

We complemented the analysis of β-diversity metrics based on observed data with null models to investigate whether observed β-diversity resembled to values generated by a stochastic assembly process. The null model generated communities with random species composition but with the same number of individuals and species richness observed in each real community and with the same species pool in each metacommunity. Then, we calculated the corrected β-Shannon and β-Jaccard-Chao index for these new species compositions. This was repeated 1000 times to obtain a distribution of values for each metacommunity. The difference between the observed β-diversity and the mean β-diversity obtained in the null model was divided by the standard deviation of the simulated β-diversity values. So, there was a standardized β-deviation value for each metacommunity, where values close to zero indicate greater similarity with the random species composition generated by the null model, while higher values indicate greater dissimilarity with the null model. We regressed deviation values of each metric (Shannon β-deviation$_{\text{count}}$; Shannon β-deviation$_{\text{dens}}$; Jaccard-Chao β-deviation$_{\text{count}}$; Jaccard-Chao β-deviation$_{\text{dens}}$) with community size. We used the permatfull function in vegan package to simulate null communities.

Community-environment relationships

To investigate the strength of community-environment relationships within metacommunities we used Hierarchical Modelling of Species Communities (HMSC [32]) to model how the occurrences and abundances of each species in each community were related to environmental variables. We fitted a HMSC model with both count ($HMSC_{count}$) and density ($HMSC_{dens}$) data, using the package HMSC in R [33]. We sampled the posterior distribution with two Markov Chain Monte Carlo (MCMC) chains. Each chain was run for 15000 iterations, of which the first 5000 were removed as burn-in. We obtained 1000 samples per chain that were recorded every 10:th step (thin=10) of the iterations. We examined MCMC convergence by examining the potential scale reduction factors [34] of the model parameters. Then, we measured the predictive power of the models through the coefficient of discrimination (Tjur R^2[35]) for each species in the metacommunity and then we calculated a median Tjur R^2 value of all species for each metacommunity that was regressed against community size.

Results

Community descriptors

As expected, we found a negative relationship between β-$Shannon_{count}$ and community size (figure 1a; std. coeff. = -0.63, R^2 = 0.39, electronic supplementary material, table S1), β-$Shannon_{dens}$ and community size (figure 1b; std. coeff. = -0.21, R^2 = 0.04, electronic supplementary material, table S1), and between β-Jaccard-$Chao_{count}$ and community size (figure 1c; std. coeff. = -0.72, R^2 = 0.52, electronic supplementary material, table S1). There was also a weak evidence of a negative relationship between β-Jaccard-

$\text{Chao}_{\text{dens}}$ and community size (figure 1d; std. coeff. = -0.18, R^2 = 0.03, electronic supplementary material, table S1). We also found $\text{Evenness}_{\text{count}}$ (electronic supplementary material, figure S1a; std. coeff. = -0.48, R^2 = 0.23, electronic supplementary material, table S2), $\text{Evenness}_{\text{dens}}$ (electronic supplementary material, figure S1b; std. coeff. = -0.55, R^2 = 0.30, electronic supplementary material, table S2) and Rank $\text{diff}_{\text{count}}$ (electronic supplementary material, figure S1e; std. coeff. = -0.62, R^2 = 0.38, electronic supplementary material, table S2) were negatively related to community size. There was no evidence of relationship between community size and Rank $\text{diff}_{\text{dens}}$ (electronic supplementary material, figure S1f; electronic supplementary material, table S2) and richness difference (electronic supplementary material, figure S1c, d; electronic supplementary material, table S2). Taken together, these relationships suggest that fish composition in small communities were more different among each other within a metacommunity, compared to metacommunities formed by larger communities.

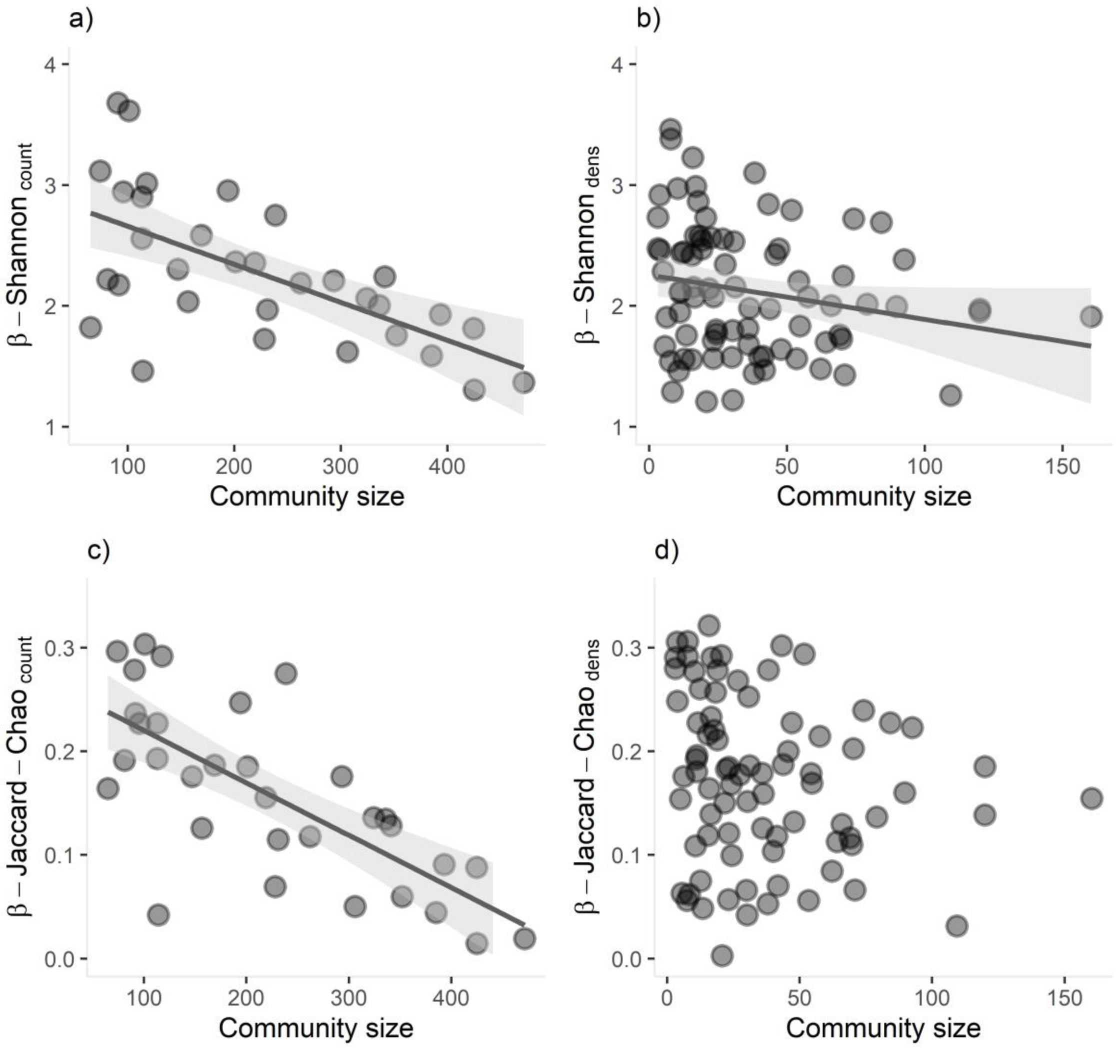

Figure 1. Relationships between β-diversity metrics and community size in riverine fish metacommunities. a) count data (β-Shannon$_{count}$), b) density data (β-Shannon$_{dens}$), c) count data (β-Jaccard-Chao$_{count}$) and d) density data (β-Jaccard-Chao$_{dens}$). Gray bands represent the 95% confidence interval around the regression line.

β-diversity null models

As expected, Shannon β-deviation$_{dens}$ (figure 2b; std. coeff. = 0.73, $R^2 = 0.53$, electronic supplementary material, table S1), Jaccard-Chao β-deviation$_{count}$ (figure 2c; std. coeff. = 0.71, $R^2 = 0.50$; electronic supplementary material, table S1) and Jaccard-Chao β-deviation$_{dens}$ (figure

2d; std. coeff. = 0.46, R^2 = 0.21, electronic supplementary material, table S1) were positively related to community size, indicating that in small communities, β-deviation values were closer to zero than in larger communities. There was only weak evidence of relationship between Shannon β-deviation$_{\text{count}}$ and community size (figure 2a; std. coeff. = 0.33, R^2 = 0.11, electronic supplementary material, table S1).

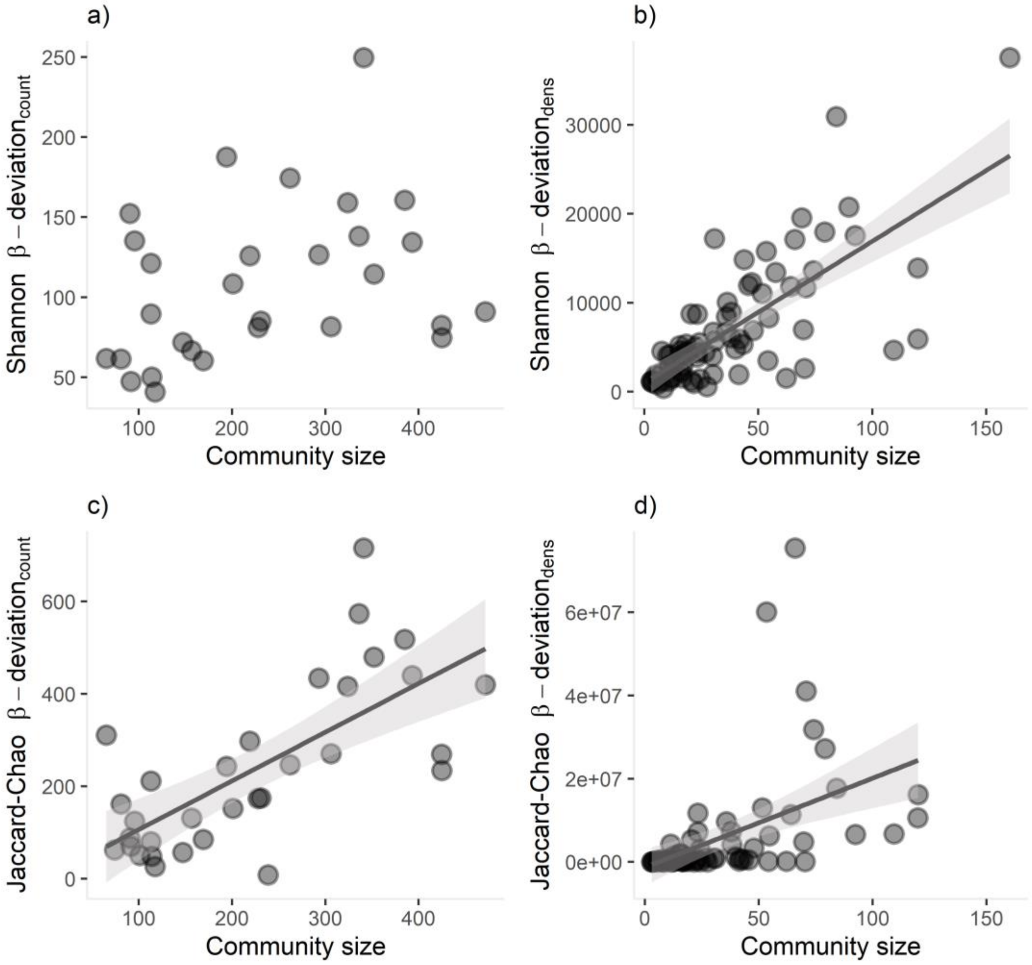

Figure 2. β-deviation relationships with community size in riverine fish metacommunities for Shannon β-deviation: a) count data and b) density data, and Jaccard-Chao β-deviation: c) count data and d) density data. Gray bands represent the 95% confidence interval around the regression line.

Community-environment relationships

The MCMC convergence of the HMSC models was satisfactory, with the potential scale reduction factors for the β-parameters varying between 1.0 and 1.05. As predicted, we found a positive relationship between the coefficient of discrimination (Tjur R^2) obtained between $HMSC_{count}$ and community size (figure 3a; SD. coeff. = 0.41, $F_{1,30} = 6.08$, $R^2 = 0.16$, $p = 0.01$), indicating that the strength of community-environment relationship was lower in metacommunities composed of smaller communities. On the other hand, there was no evidence of relationship between strength of community-environment relationship and community size for $HMSC_{dens}$ (figure 3b; $F_{1,79} = 2.89$, $R^2 = 0.03$, $p = 0.09$).

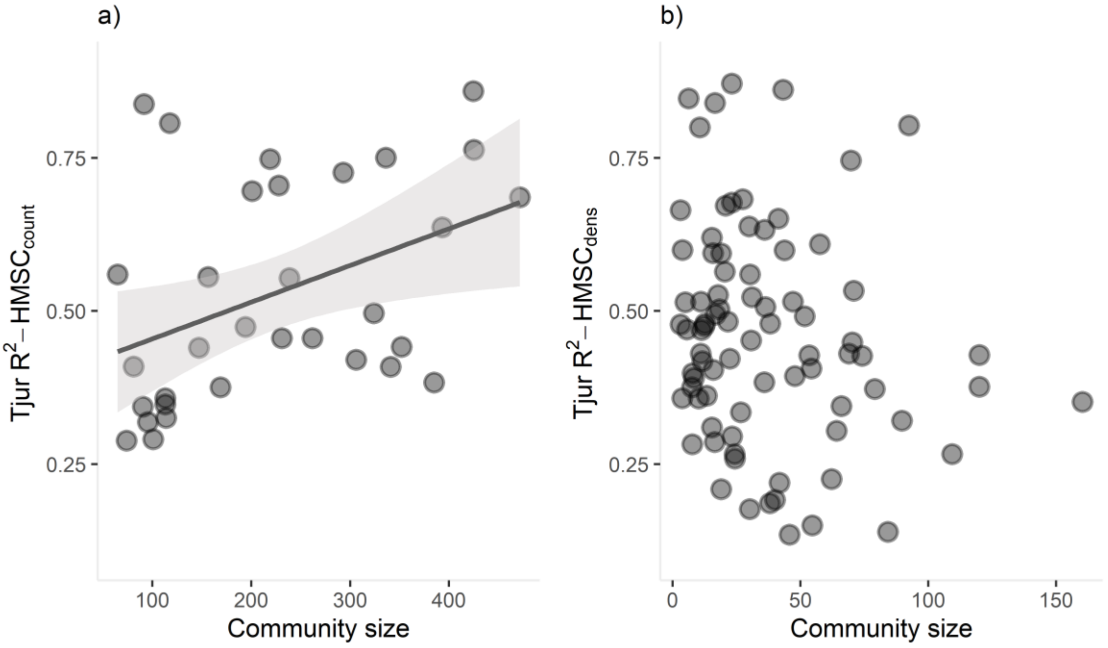

Figure 3. Strength of community-environment relationships within riverine fish metacommunities regressed against community size using: a) count (TjurR^2 - $HMSC_{count}$) and b) density (TjurR^2 - $HMSC_{dens}$) data. The gray band represents the 95% confidence interval around the regression line.

Discussion

Our analyses of intercontinental data support the prediction that spatial variation in species composition among small communities is higher than among large communities and similar to the dynamics expected under stochastic assembly. We also found moderate 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 mediate 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 forcing. Such decoupling implies that further declines in the size of fish communities due to environmental change can add more stochasticity to their dynamics, making them less predictable in the face of future environmental scenarios.

The negative relationship between observed β-diversity and community size, both for β-Shannon and β-Jaccard-Chao indices, indicates that species composition varies more among small communities than among large communities. Spatial variation in species composition can occur due to changes in various aspects of community structure [28]. For example, previous research has shown that high temporal beta diversity can occur despite no decrease in species richness [36, 37, 38], suggesting that changes in species relative abundances might be a better indicator of community change. This match with our findings, as we found that the negative relationship between β-diversity and community size seemed to be a result of mechanisms changing species relative abundance, but not species richness.

A negative relationship between β-diversity and community size alone does not provide strong evidence for the role of stochasticity in small communities. But we also found that β-diversity among small communities was similar to β-diversity of communities simulated under a purely stochastic process. This is in line with the results reported by Siqueira *et al*. [5] for stream insect communities in the neotropics. Random and unpredicted changes in species composition can happen in small communities entirely or partially due to demographic events that are independent of environmental selection [39, 40]. For example, smaller communities are more likely to be affected by births and deaths, changing species relative abundance independently of species fitness [40]. Thus, if such demographic events play a major role in community dynamics, species competitive differences become less important to shape community composition [3, 41]. As community size increases, the role of demographic events that are random regarding species identity should decrease relative to deterministic assembly, as indicated by the higher β-deviation values we found among larger communities, which indicate that observed β-diversity was different from stochastic assembly expectations.

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 limit power to predict species identity or species abundances in small communities. Weak relationships between species composition and environment factors are abundant in community ecology [6, 42, 43]. 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. Heino *et al*. [43] 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 studies, ours also have limitations. Stream communities can be structured by many environmental variables, such as nutrient concentration, turbidity, pH [44] and the ones that we used to relate with community composition (e.g., air temperature, precipitation) are not necessarily the best variables to represent community-environment relationships, as they were not measured at the same spatial scale used to sample communities. However, even with such data, we detected a positive relationship between the strength of community-environment relationship and community size. Thus, the inclusion of local environmental variables would likely make this relationship stronger, providing better support for our hypothesis.

Relying on a robust data set that spans 2332 communities, across 113 river basins in four biogeographic realms, and robust methods to quantify beta diversity, we provide evidence for a general pattern in nature – species composition vary more among small communities than among large communities. This pattern seems to be driven by the stronger influence of demographic stochastic events on small communities, as their dynamics are weakly linked to spatial variation in environmental factors. As the effects of demographic stochasticity should manifest also, maybe more strongly, on community temporal dynamics, our next steps will concentrate on analyzing the relationship between temporal beta diversity and community size. We also suggest that further studies should test the prevalence of the relationships we describe here in terrestrial metacommunities, where community size might be more directly linked to ecosystem size. In a changing world where ecosystems have been fragmented and reduced, and species overexploited, managers might need to put a major focus on actions that increase population size as, for example, species reintroductions [45].

Data Accessibility

R codes and data used here are publicly available on Zenodo: https://doi.org/10.5281/zenodo.6504818

Author contributions

CMJ and TS conceived the ideas and designed the study; CMJ selected data and performed the analysis; CMJ and TS interpreted the results; CMJ wrote the manuscript with substantial contribution from TS.

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).

References

1. Otto SP., Whitlock MC. 1997 Fixation of beneficial mutations in a population of changing size. *Genetics*, 146,723-733.
2. Vellend M. 2016 The theory of ecological communities. Mpb series: Edition 57. *Princeton University Press*, Princeton, New Jersey, USA.
3. Orrock JL., Watling JI. 2010 Local community size mediates ecological drift and competition in metacommunities. *Proc. R. S. B.*, 277, 2185-2191. (doi: https://doi.org/10.1098/rspb.2009.2344)

4. Gilbert N., Levine JM. 2017 Ecological drift and the distribution of species diversity. *Proc. R. S. B.,* 284, 20170507. (doi: https://doi.org/10.1098/rspb.2017.0507)
5. Siqueira T., Saito VS., Bini LM., Melo AS., Petsch DK., Landeiro VL., Tolonen KT., Jyrkänkallio-Mikkola J., Soininen J., Heino J. 2020 Community size can affect the signals of ecological drift and niche selection on biodiversity. *Ecology*, 101, e03014. (doi: https://doi.org/10.1002/ecy.3014)
6. Bini LM., Landeiro VL., Padial AA., Siqueira T., Heino J. 2014 Nutrient enrichment is related to two facets of beta diversity for stream invertebrates across the United States. *Ecology*, 95, 1569-1578. (doi: https://doi.org/10.1890/13-0656.1)
7. Comte L., Carvajal-Quintero J., Tedesco PA., Giam X., Brose U., Erős T., Filipe AF, Fortin M-J, Irving k, Jacquet C., e*t* al. 2020 RivFishTIME: A global database of fish time-series to study global change ecology in riverine systems. *Glob. Ecol. Biogeogr.*, 30, 38-50. (doi: https://doi.org/10.1111/geb.13210)
8. Cao K., Svenning JC., Yan C., Zhang J., Mi X., Ma K. 2021 Undersampling correction methods to control γ-dependence for comparing β-diversity between regions. *Ecology*, 102, 9. (doi: https://doi.org/10.1002/ecy.3448)
9. Lehner B., 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. (doi: https://doi.org/10.1002/hyp.9740)

10. R Core Team. 2021 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
11. Appelhans T., Detsch F., Reudenbach C, Woellauer S. 2021 mapview: Interactive Viewing of Spatial Data in R. R package version 2.10.0.
12. Hijmans RJ. 2021 raster: Geographic Data Analysis and Modeling. R package version 3.4-13.
13. Bivand R., Keitt T., Rowlingson B. 2021 rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-23.
14. Pebesma, E. 2018 Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10, 439-446. (doi: 10.32614/RJ-2018-009)
15. Pebesma EJ., Bivand RS. 2005 Classes and methods for spatial data in R. *R News*, 5, 9-13.
16. Bivand RS., Pebesma E., Gomez-Rubio V. 2013 Applied spatial data analysis with R, Second edition. Springer, NY.
17. Wickham H., Averick M., Bryan J., Chang W., McGowan L. D'A., François R. 2019 Welcome to the tidyverse*. J. of Open Source Softw*., 4, 1686. (doi: 10.21105/joss.01686)
18. Linke S., Lehner B., Dallaire CO., Ariwi J., Grill G., Anand M., Beames P., Burchard-Levine V., Maxwell S., *et al*. 2019 Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Sci. Data*, 6, 283. (doi: https://doi.org/10.1038/s41597-019-0300-6)
19. Grill G., Lehner B., Thieme M., Greenen B., Tickner D., Antonelli F., Babu S., Borrelli P., Cheng L., Crochetiere H., *et al.* 2019 Mapping the

world's free-flowing rivers. *Nature,* 569, 215-221. (doi: https://doi.org/10.1038/s41586-019-1111-9)

20. Döll P., Kaspar F., Lehner B. 2003 A global hydrological model for deriving water availability indicators: model tuning and validation. *J. Hydrol.*, 270, 105-134. (doi: https://doi.org/10.1016/S0022-1694(02)00283-4)
21. Robinson N., Regetz J., Guralnick RP. 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. (doi: https://doi.org/10.1016/j.isprsjprs.2013.11.002)
22. Hijmans RJ., Cameron SE., Parra JL., Jones PG., Jarvis A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol*., 25, 1965–1978. (doi: https://doi.org/10.1002/joc.1276)
23. Venter O., Sanderson EW., Magrach A., Allan JR., Beher J., Jones KR., Possingham HP., Laurance WF., Wood P., Fekete BM *et al*. 2016 Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data*, 3, 160067. (doi: 10.1038/sdata.2016.67)
24. Beck J., Holloway JD., Schwanghart W. 2013 Undersampling and the measurement of beta diversity. *Methods Ecol Evol*, 4, 370-382. (doi: https://doi.org/10.1111/2041-210x.12023)
25. Chase JM., Myers JA. 2011 Disentangling the importance of ecological niches from stochastic processes across scales. *Philos. Trans. R. Soc. B*, 366, 2351–2363. (doi: https://doi.org/10.1098/rstb.2011.0063)
26. Kraft NJB., Comita LS., Chase JM., Sanders NJ., Swenson NG., Crist TO., Stegen JC, Vellend M., Boyle B., Anderson MJ., *et al*. 2011

Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, 333, 1755-1758. (doi: 10.1126/science.1208584)

27. Chao A., Wang YT., Jost L. 2013 Entropy and the species accumulation curve: a novel entropy estimator via discovery rates of new species. *Methods Ecol. Evol*., 4, 1091-1100. (doi:10.1111/2041-210X.12108)
28. Avolio ML., Carroll IT., Collins SL., Houseman GR., Hallett LM., Isbell F., Koerner SE., Komatsu KJ., Smith MD, Wilcox, KR. 2019 A comprehensive approach to analyzing community dynamics using rank abundance curves. *Ecosphere*, 10, e02881. (doi: https://doi.org/10.1002/ecs2.2881)
29. Marcon E., Herault B. 2015 entropart: An R Package to Measure and Partition Diversity. *J. Stat. Softw*., 67, 1-26. (doi: 10.18637/jss.v067.i08)
30. Oksanen J., Blanchet FG., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin PR., O'Hara RB., Simpson GL., Solymos P., *et al*. 2020 vegan: Community Ecology Package. R package version 2, 5-7.
31. Hallett L., Avolio M., Carroll I., Jones S., MacDonald A., Flynn D., Ripplinger J., Slaughter P., Corinna G., Collins SL. 2020 codyn: Community Dynamics Metrics. R package version 2.0.5. (doi: https://doi.org/10.1111/2041-210X.12569)
32. Ovaskainen O., Tikhonov G., Norberg A., Blanchet FG., Duan L., Dunson D., Roslin T., Abrega 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. (doi: https://doi.org/10.1111/ele.12757)

33. Tikhonov G., Ovaskainen O., Oksanen J., de Jonge M., Opedal O., Dallas T. 2021 hmsc: Hierarchical Model of Species Communities. R package version 3.0-11.

34. Gelman A., Rubin DB. 1992 Inference from iterative simulation using multiple sequences. *Stat. Sci*., 7, 457-472. (doi: 10.1214/ss/1177011136)

35. Tjur T. 2009 Coefficients of determination in logistic regression models—A new proposal: the coefficient of discrimination. *Am. Stat.*, 63, 366–372. (doi: https://doi.org/10.1198/tast.2009.08210)

36. Dornelas M., Gotelli NJ., McGill B., Shimadzu H., Moyes F., Sievers C., Magurran A. 2014 Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344, 296-9. (doi: 10.1126/science.1248484)

37. Jones SK., Ripplinger J., Collins SL. 2017 Species reordering, not changes in richness, drives long-term dynamics in grassland communities. *Ecol. Lett*., 20, 1556-1565. (doi: https://doi.org/10.1111/ele.12864)

38. Hillebrand H., Blasius B., Borer ET., Chase JM., Downing JA., Eriksson BK., Filstrup CT., Harpole WS, Hodapp D., Larsen S. et al. 2018 Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *J. Appl. Ecol.*, 55: 169– 184. (doi: https://doi.org/10.1111/1365-2664.12959)

39. Shoemaker LG., Sullivan LL., Donohue I., Cabral JS., Williams RJ., Mayfield MM., Chase JM., Chu C., Harpole WS., Huth A., *et al*. 2020 Integrating the underlying structure of stochasticity into Community ecology. *Ecology*, 101, e02922. (doi: 10.1002/ecy.2922)

40. Vellend M. 2010 Conceptual synthesis in community ecology. *Q. Rev. Biol*., 85, 183-206. (doi: 10.1086/652373)

41. Orrock JL., Fletcher RJ Jr. 2005 Changes in community size affect the outcome of competition. *Am Nat*., 166, 107-11. (doi:10.1086/430641)

42. Nabout JC., Siqueira T., Bini LM., Nogueira IS. 2009 No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta Oecol.*, 35, 720–726. (doi: https://doi.org/10.1016/j.actao.2009.07.002)

43. Heino J., Melo AS., Bini LM., Altermatt F, Al-Shami SA., Angeler DG., Bonada N., Brand C., Callisto M., Cottenie K., *et al*. 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. (doi: https://doi.org/10.1002/ece3.1439)

44. Allan JD., Castillo MM., Capps KA. 2021 Stream ecology: structure and function of running waters. *Springer Nature*.

45. Jourdan J., Plath M., Tonkin JD., Ceylan M., Dumeier AC., Gellert G., Graf W., Hawkins CP., Kiel E., Lorenz AW., *et al*. 2019 Reintroduction of freshwater macroinvertebrates: challenges and opportunities. *Biol. Rev*., 94, 368-387. (doi: 10.1111/brv.12458)

Supplemental Material

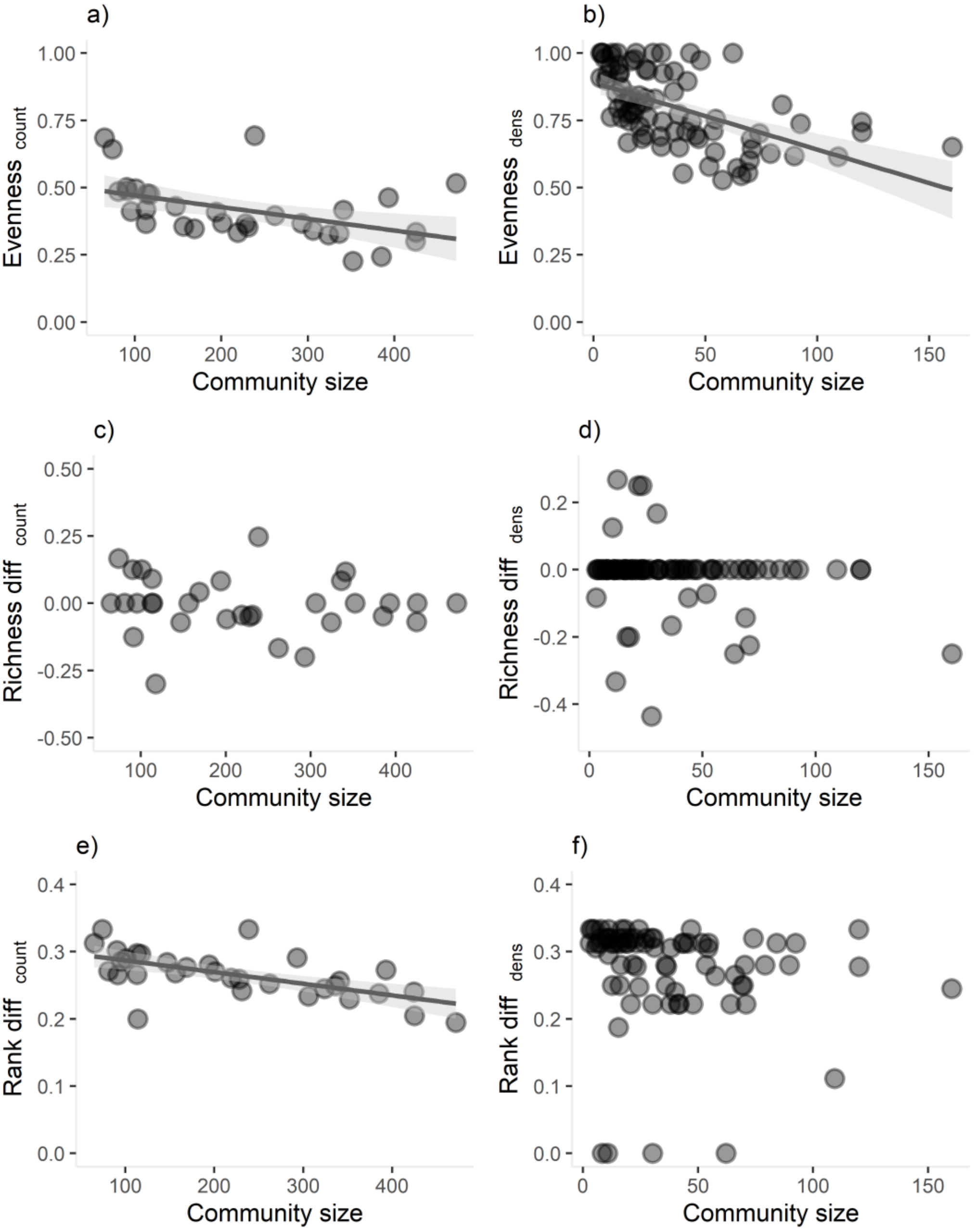

Figure S1. Relationships between community compositional differences and community size in riverine fish metacommunities for count data: a) evenness, c) richness difference, e) rank difference; and for density data b) evenness, d) richness difference, f) rank difference. The gray bands represent the 95% confidence interval around the regression line.

Table S1. Summary statistics describing the relationship between *β-diversity and β-deviation indices (β-Shannon and β-Jaccard-Chao) with community* size for count and density data. SD coeff. = standardized regression coefficient; R2 and Adj. R2 = coefficient of determination and adjusted coefficient of determination of the regressions, respectively.

Response variable	SD coeff.	R^2	Adj R^2	F-statistic	p-value
β-Shannon$_{count}$	-0.63	0.39	0.37	19.78	0.0001
β-Shannon$_{dens}$	-0.21	0.04	0.03	3.79	0.055
β-Jaccard-Chao$_{count}$	-0.72	0.52	0.50	32.65	<0.0001
β-Jaccard-Chao$_{dens}$	-0.18	0.03	0.02	2.9	0.0921
Shannon β-deviation$_{count}$	0.33	0.11	0.08	3.47	0.0734
Shannon β-deviation$_{dens}$	0.73	0.53	0.53	91.82	<0.0001
Jaccard-Chao β-deviation$_{count}$	0.71	0.50	0.48	30.7	<0.0001
Jaccard-Chao β-deviation$_{dens}$	0.46	0.21	0.20	18.26	<0.0001

Table S2. Summary statistics describing the relationship between community compositional difference *(evenness, richness difference and rank difference) and community size* for count and density data. SD coeff. = standardized regression coefficient; R2 and Adj. R2 = coefficient of determination and adjusted coefficient of determination of the regressions, respectively.

Response variable	SD coeff.	R^2	Adj R^2	F-statistic	p-value
Evenness$_{count}$	-0.48	0.23	0.20	8.99	0.005
Evenness$_{dens}$	-0.55	0.30	0.29	35.14	<0.0001
Richness diff$_{count}$	-0.09	0.009	-0.02	0.27	0.6011
Richness diff$_{dens}$	-0.19	0.03	0.02	3.16	0.0793
Rank diff$_{count}$	-0.62	0.38	0.36	19.03	0.0001
Rank diff$_{dens}$	-0.16	0.02	0.01	2.24	0.1381