scientific reports

Check for updates

OPEN

The importance of common and the irrelevance of rare species for partition the variation of community matrix: implications for sampling and conservation

Leandro Schlemmer Brasil[1,2✉], Thiago Bernardi Vieira[3], André Felipe Alves Andrade[4], Rafael Costa Bastos[2,3], Luciano Fogaça de Assis Montag[1,2,3] & Leandro Juen[1,2,3]

In community ecology, it is important to understand the distribution of communities along environmental and spatial gradients. However, it is common for the residuals of models investigating those relationships to be very high (> 50%). It is believed that species' intrinsic characteristics such as rarity can contribute to large residuals. The objective of this study is to test the relationship among communities and environmental and spatial predictors by evaluating the relative contribution of common and rare species to the explanatory power of models. Our hypothesis is that the residual of partition the variation of community matrix (varpart) models will decrease as rare species get removed. We used several environmental variables and spatial filters as varpart model predictors of fish and Zygoptera (Insecta: Odonata) communities in 109 and 141 Amazonian streams, respectively. We built a repetition structure, in which we gradually removed common and rare species independently. After the repetitions and removal of species, our hypothesis was not corroborated. In all scenarios, removing up to 50% of rare species did not reduce model residuals. Common species are important and rare species are irrelevant for understanding the relationships among communities and environmental and spatial gradients using varpart. Therefore, our findings suggest that studies using varpart with single sampling events that do not detect rare species can efficiently assess general distributional patterns of communities along environmental and spatial gradients. However, when the objectives concern conservation of biodiversity and functional diversity, rare species must be carefully assessed by other complementary methods, since they are not well represented in varpart models.

Understanding the distribution of biological communities along environmental and spatial gradients has had several theoretical advances such as the Niche concept[1], Neutral Theory[2], and the Meta communities synthesis[3]. Parallel to these advances, the development of methods that quantify the importance of spatial and environmental filters as community predictors[4–6] have been essential for interpreting the distribution patterns of species in communities[7].

Niche theory[1], that predicts species are distributed according to the environmental conditions[8] and biotic interactions[9], which determine their distribution along the environmental gradient. Thus, species distribution is solely explained by the conditions and resources present in the studied sites. In contrast, Neutral theory posits that species distribution does not depend of environmental conditions or resources, and that spatially closer sites would have similar communities[10] due to historical processes such as vicariance and dispersal[2]. Therefore, according to niche theory predictions, the main filters for species distributions are the relationship between species and their environment, along with local extinctions where conditions and resources are inadequate[1]. Alternatively, for Neutral theory, the main filters are species dispersal and the probabilities (random events) of speciation and

[1]Programa de Pós-Graduação em Zoologia, Universidade Federal do Pará, Belém, Pará, Brasil. [2]Laboratório de Ecologia e Conservação, Universidade Federal do Pará, Belém, Pará, Brasil. [3]Programa de Pós-Graduação em Ecologia e Conservação, Universidade Federal do Pará, Belém, Pará, Brasil. [4]Theory, Metacommunity and Landscape Ecology Lab, Universidade Federal de Goias, Goiânia, Goias, Brasil. ✉email: brasil_biologia@hotmail.com
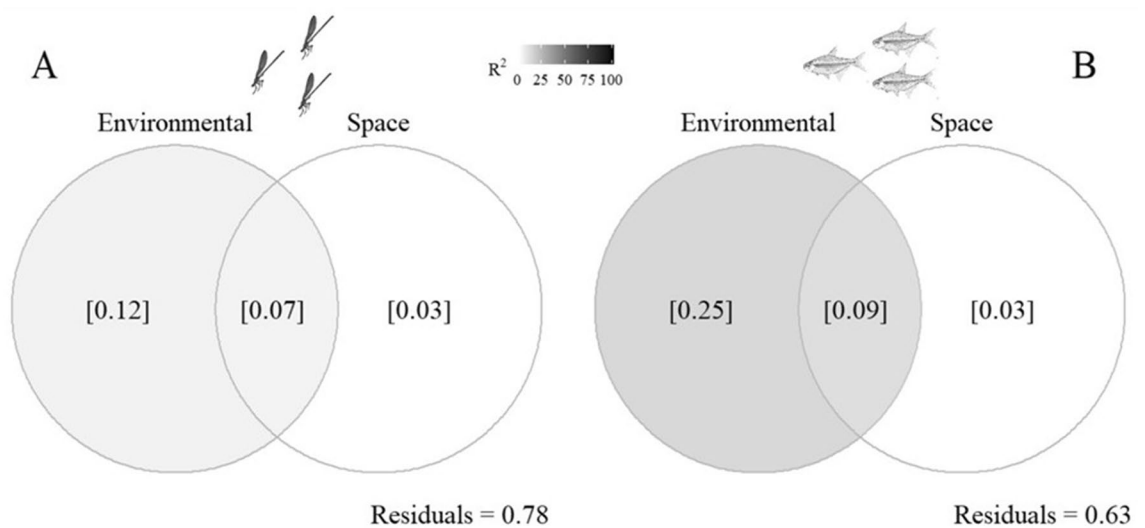
Figure 1. Variation partitioning (pRDA) analyzing the effects of environmental predictors, spatial predictors, the interaction term between environmental and spatial predictors, and the residual portion on all species of Zygoptera and fish communities.

extinction[2]. Considering these different structuring forces, to understand the distribution patterns of species in a landscape, it is necessary to understand the environmental and spatial mechanisms associated with species[7].

A way of quantifying the importance of those mechanisms is using variation partitioning models[5] between environmental and spatial predictors[11]. These models have four elements: the portion explained exclusively by the spatial component, the portion explained exclusively by the environmental component, the portion explained by the interaction between environmental and spatial predictors, and the residual portion, not explained by the model[12]. Though this is a widely accepted method for investigating the effects of niche and neutral mechanisms in community ecology, a recent meta-analysis[7] revealed it is common for the residual portion of the model to be higher than the explained portion ($> 50\%$)[5]. This was shown in very metacommunities with fish[13], Odonata[14,15] and termites[16] in Amazonian streams, macroinvertebrates in the south[17] and southeast[18] of Brazil, and in macrophytes[19] and beetles[20] in temperate regions. The high residual variation is associated with environmental or spatial predictors that were not included in the models[5,7] and/or with a subset of species that have antagonistic responses to the environmental and/or spatial gradients due to autoecological characteristics[21].

One potential reason for such an antagonistic response is the categorization of species as rare or common, which tend to respond to environmental gradients in an idiosyncratic way[18,22]. Because of that, by analyzing all species in the community together, it is possible to increase the model residual portion. In this context, species that are common and have a large spatial distribution do not have high dispersal limitations and are generalists, possessing environmental plasticity that allows them to survive in different conditions[23], some of which are considered adverse. While, rare species, on the other hand, have more restricted distributions[24] and knowledge about their relationship with environmental variables is generally limited. For these reasons, rare species are often not considered in community analyses or have few statistical relevance to community patterns (e.g. in direct ordination analyses)[4].

To understand the mechanisms and patterns important for community assembly it is essential to create a multitaxon approach, allowing for greater extrapolation power of the results[25]. Considering studies in stream ecology to jointly evaluate taxa that use different parts of the habitat (e.g.: fish, exclusively aquatic organisms, and Zygoptera adults, which live in the riparian vegetation[13,26]) may provide robust results which could be extrapolated for other organisms of the aquatic biota living in similar habitats.

Thus, our objective is to test the relationship among communities and the environmental and spatial predictors by evaluating the relative contribution of common and rare species to the power of explanation of the models. Our hypothesis is that the residual portion of variation partitioning models will decrease as rare species are removed from the analyses. Rare species would inflate the residual portion as they make the models less parsimonious, since they would add too many species that are not relevant in the relationship among community patterns and spatial and environmental patterns.

Results

Environmental and spatial predictors explained together 37% and 22% of the variation for Zygoptera and fish, respectively. For fish (25%) and Zygoptera (12%), most of the explanation was for environmental predictors. The spatial component explained only 3% for fish and Zygoptera (Fig. 1).

The gradual removal of rare species (up to 50% of total community richness) had little effect on the percentage of variation explained by environmental and spatial predictors for Zygoptera ($\Delta R^2 = -0.004$, $p = 0.053$) and fish ($\Delta R^2 = -0.003$, $p = 0.060$) communities. However, the gradual removal of common species (up to 50% of total community richness) strongly altered the percentage of variation explained by environmental and spatial
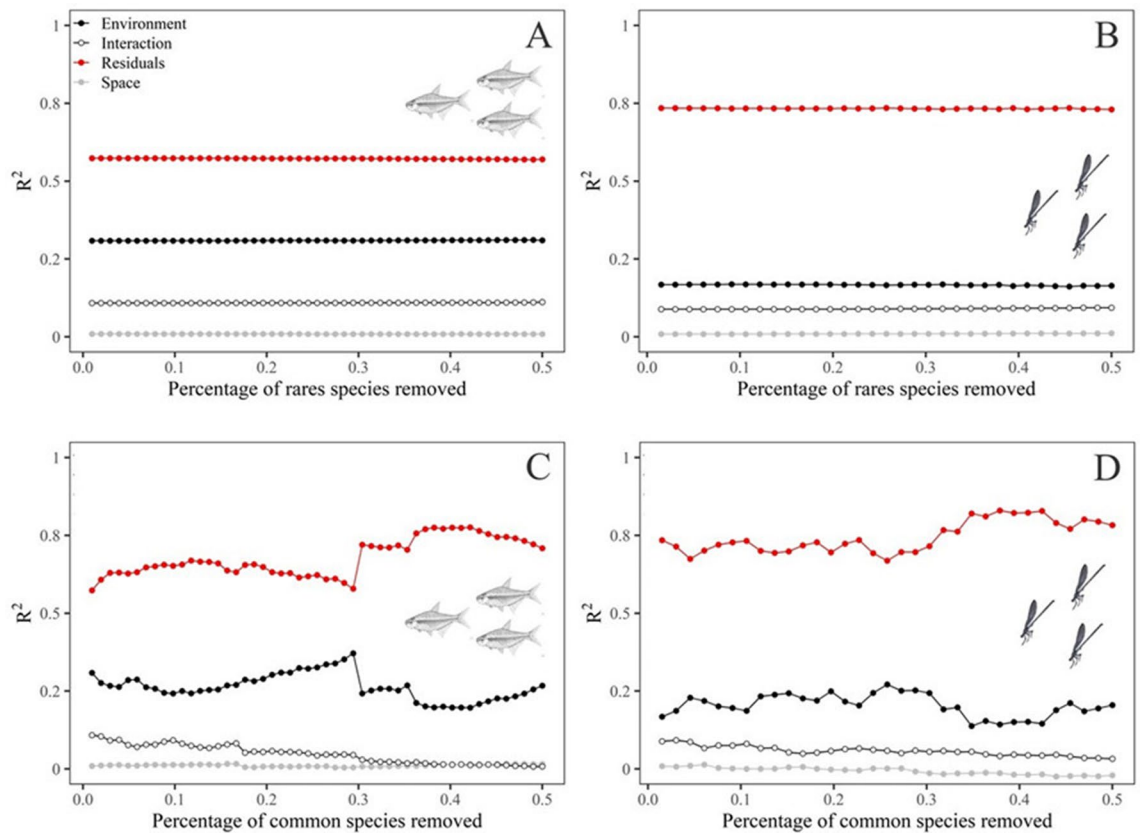
Figure 2. Correlogram showing the R^2 values of the variation partitioning (pRDA) analyzing the effects of environmental predictors (black circles), spatial predictors (empty circles), the interaction termbetween environmental and spatial (grey circles), and the residual portion (red circles) on fish and Zygoptera communities. A = fish communities with abundance data gradually removing rare species. B = Zygoptera communities with abundance data gradually removing rare species. C = fish communities with presence-absence data gradually removing rare species. D = Zygoptera communities with abundance data gradually removing common species.

predictors of fish ($\Delta R^2 = 0.135$, p = 0.002) communities, but did not significantly differ from the alteration in the explanation of Zygoptera ($\Delta R^2 = 0.048$, p = 0.536) communities when species are removed at random when species are removed at random (Fig. 2A:D).

Discussion

Our hypothesis that the residual portion of the variation partitioning models would decrease as rare species were removed was not corroborated. In all scenarios, for both fish and Zygoptera, removing up to 50% of rare species did not reduce residuals in a significant way. However, for the most common species, the removal of only one species had notable effects in the residual portion of the models. This effect is more pronounced in communities in which a few species dominate the community, such as the fish community in our study (Fig. 3), since the removal of the most abundant species steeply reduces the percentage of environmental and spatial variation explained.

Therefore, when analyzing general patterns of community distribution in spatial or environmental gradients using variation partitioning, the exclusion of rare species does not promote significant residual increase. Due to potential analytical problems associated with rare species in community analyses, for example, the great number of zeros in species matrices[11] and having more species than samples[27], it is recommended to exclude rare species before performing data analyses. Our results suggest this is possible without harming the visualization of general patterns of communities along spatial and environmental gradients[28].

In tropical countries, where species diversity is high[29], big knowledge shortfalls regarding biodiversity distribution[30] (known as the Wallacean shortfall[31]) exist. Additional financial crisis may reduce research investments[32], so it is common for the biological sample to be from one sole sampling/field trip[13,26,33,34]. These sampling strategies can result in a low representation of rare species, and, consequently, can affect the detection of diversity patterns[35]. However, despite these possible limitations, our results demonstrate that even if a considerable portion of rare species is not captured this does not make it impossible to evaluate the general distributional patterns of communities in spatial and environmental gradients.

Nonetheless, for conservation[24] or to assess the functional role of species in ecosystems[36,37], the exclusion of rare species can bias the results. Functionally, rare species contribute considerably to patterns of richness, specialization, and functional originality of communities[36,37]. Conservation assessment at the community level
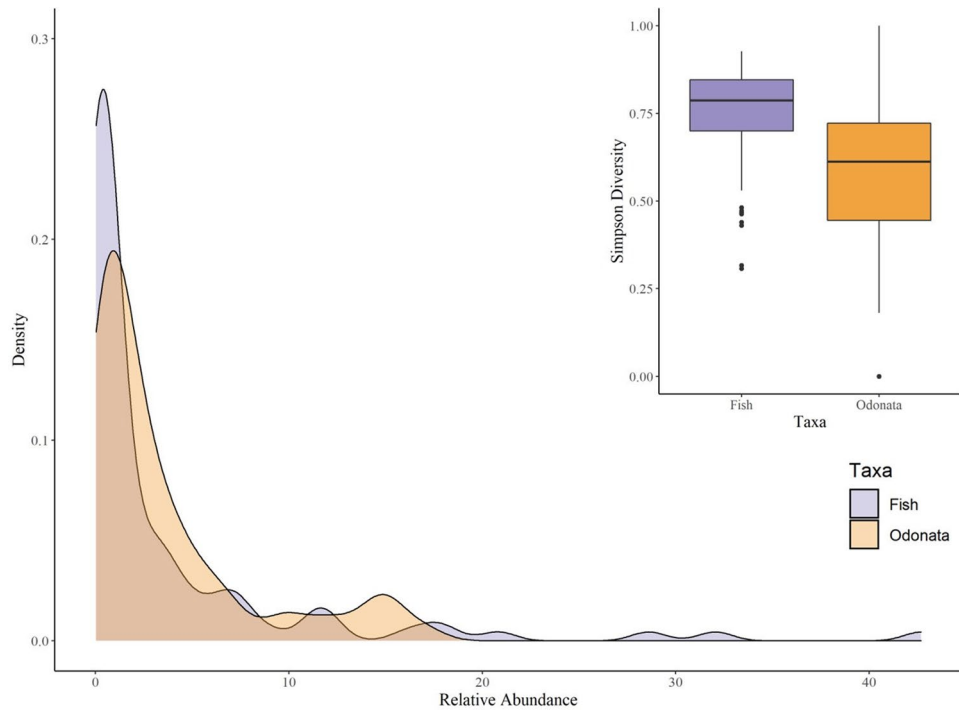
Figure 3. Abundance curves of fish and Odonata communities observed. The boxplot represents the variation of dominance represented by the Simpson index in the subsamples.
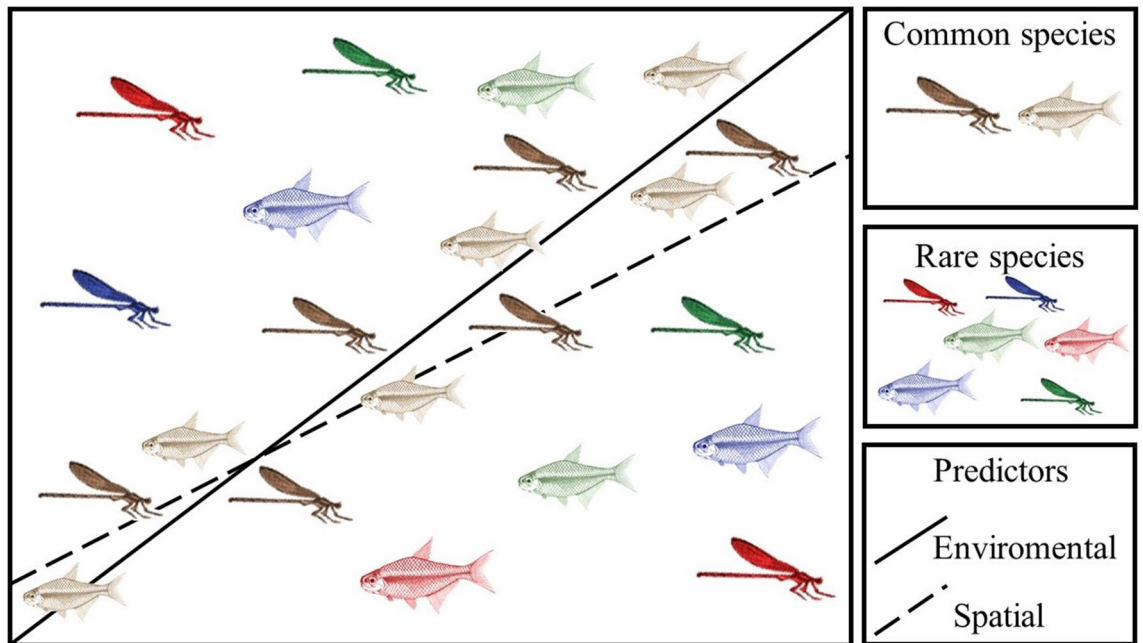

Figure 4. Graphic summary of the main results.

can disregard rare and important taxa, such as endemic species[38]. Thus, conservation measures are commonly evaluated and proposed at the population level[39].

From our findings, we conclude that common species are important, and rare species contribute little for understanding the relationships of communities with spatial and environmental gradients using variation partitioning (Fig. 4). However, when considering other objectives such as conservation or the analysis of functional patterns, information about rare species continues to be essential. We further emphasize that studies with sole sample events often may not be effective in detecting rare species but are efficient in identifying general

community distribution patterns along environmental and spatial gradients. Therefore, when necessary, the exclusion of rare species does not harm the interpretation of community patterns in those circumstances.

Material and methods

Study area.   We sampled Zygoptera and Fish assemblages of small streams (up to 4 m width and a 0.8 m mean depth) in Eastern Amazon (Brazil). We sampled fish in 109 streams and Odonates in 141 streams. The sampling sites are in areas where the natural forest exists but has been suffering with a continuous process of landscape fragmentation due to the advancement of agrosystems (e.g.: agriculture, pasture and logging) in natural areas[13,26]. The study area consists of three macro climatic zones according to the Köppen classification[40]: The Am zone (North), the Aw zone (South), and the Af zone (East). The Am is a tropical rainy climate (also called tropical monsoon climate), Aw is a tropical climate with distinct dry and rainy seasons, and Af is a tropical rainforest climate (also called equatorial climate)[40].

All methods were carried out in accordance with relevant guidelines and regulations. For Odonata samples the surveys were authorized by the Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA, Licence No. 1993421). For Fish samples the Ethics Committee of the Federal University of Pará approved the surveys (CEUA nº 8293020418), and they were conducted under the permission nº 4681-1 granted by the Instituto Chico Mendes de Conservação da Biodiversidade (IcmBIO; Ministério do Meio Ambiente).

Zygoptera sample.   We sampled adults Zygoptera once on each stream. Sampling happened between 2009 and 2018 always between the months of July and December, when the highest diversity of aquatic insects is collected in the Amazon. In each stream, we demarcated a linear transect of 150 m and captured all the Zygoptera adult specimens observed in the span of one hour. We used an entomological net of 40 cm diameter and 65 cm width. To minimize bias related to the different types of Zygoptera thermoregulation (thermal conformers, heliotherms and endotherms), the collection of specimens occurred between 10 a.m. and 2 p.m., when sunlight reaches the stream channel[14,41].

The specimens were stored according to the protocol proposed by Ref.[42]. We identified the material using taxonomic keys and specialized guidebooks e.g.: Refs.[42–44], and, when necessary, we sent the material to specialists. The specimens are deposited in the collection of the Zoology Museum of the Federal University of Pará, located in Belém, state of Pará, Brazil.

Fish sample.   The fish were collected in 109 first and third-order streams, along 150 m transects, in the dry season, between July and December of 2012–2015. We selected this period to avoid seasonal variation in the fish assemblage structure[45,46] and to increase sampling efficiency, which is easier under low streamflow.

The fish were captured using circular 55-cm-diameter dip nets with a 3 mm metallic mesh. This type of dip net is considered an efficient method to collect specimens, allowing the capture of fish in small stream microhabitats, including complex habitats (e.g. riparian vegetation, leaf banks and the inferior portion of wood trunks and branches[47]). The 150 m transect was divided into 10 sections. Each section was sampled by 2 collectors for 18 min, and the total sampling time was 3 h/stream. The specimens were euthanized using anesthetics (Eugenol, following the Brazilian Civil Law nº 11.794/2008), fixed in a 10% formalin solution, and transferred to 70% ethanol after 48 h. In the laboratory, the specimens were identified using taxonomic keys[48,49] and with the aid of specialists. All specimens were deposited in the ichthyology collection of the Emílio Goeldi Museum (MPEG), Belém, State of Pará, Brazil.

Environmental predictors.   We used environmental variables measured by the protocol of environmental assessment developed by the United States Environmental Protection Agency (US-EPA)[50,51]. We took measurements and observed the characteristics of the habitat in the same 150 m stretch where Zygoptera and fish sampling took place. In total, we obtained 186 measurements: 27 are related to channel morphology and hydraulics, 26 are related to substrate, 28 to organic debris, 60 are related to wood, 16 to the characteristics of the riparian forest, and 29 refer to human influence[13]. We took a series of steps to reduce the number of variables. Initially, we removed variables that had a low (< 40%) coefficient of variation (CV) and the ones that had high values (> 80%). We removed them because predictors that have little variation are not representative in the environmental gradients, and samples with too many zeroes (high coefficient of variation) indicate a sampling problem[13,26].

Considering this selection criterion for the 186 initial environmental variables, there were 139 remaining predictors for Zygoptera, and 51 for Fish. This difference is because some of the collection sites were distinct for fish and Zygoptera. After that, we used those environmental matrices to perform a model selection (forward stepwise) with species composition (separately for Zygoptera and fish). We used this method of selection to eliminate predictors that did not associate with the species matrices and to avoid inflated residual values because of the use of spurious variables[4]. To avoid multicollinearity in further analyses, we performed Principal Component Analyses (PCA) with the original predictors and used the ordination axes as the predictive variables. To avoid information loss, we retained the axes until the sum of explanation (eigenvalue) was equal to or higher than 90%[52]. For Zygoptera and fish, seven axes were necessary for > 90% in PCA (Supplementary Information).

Spatial predictors.   It is expected for communities that are spatially closer to have similar species composition[53]. This happens because there is a higher likelihood of a migratory species flow between environmental patches that are spatially closer[3]. Consequently, when we analyze biological communities through space it is necessary to understand how much of the similarity or dissimilarity among species relates only with the spatial distribution of the sampling sites. To analyze the importance of space, we used spatial filters[6]. For this approach, we used latitude and longitude data of each sampling site to calculate Principal Coordinates of
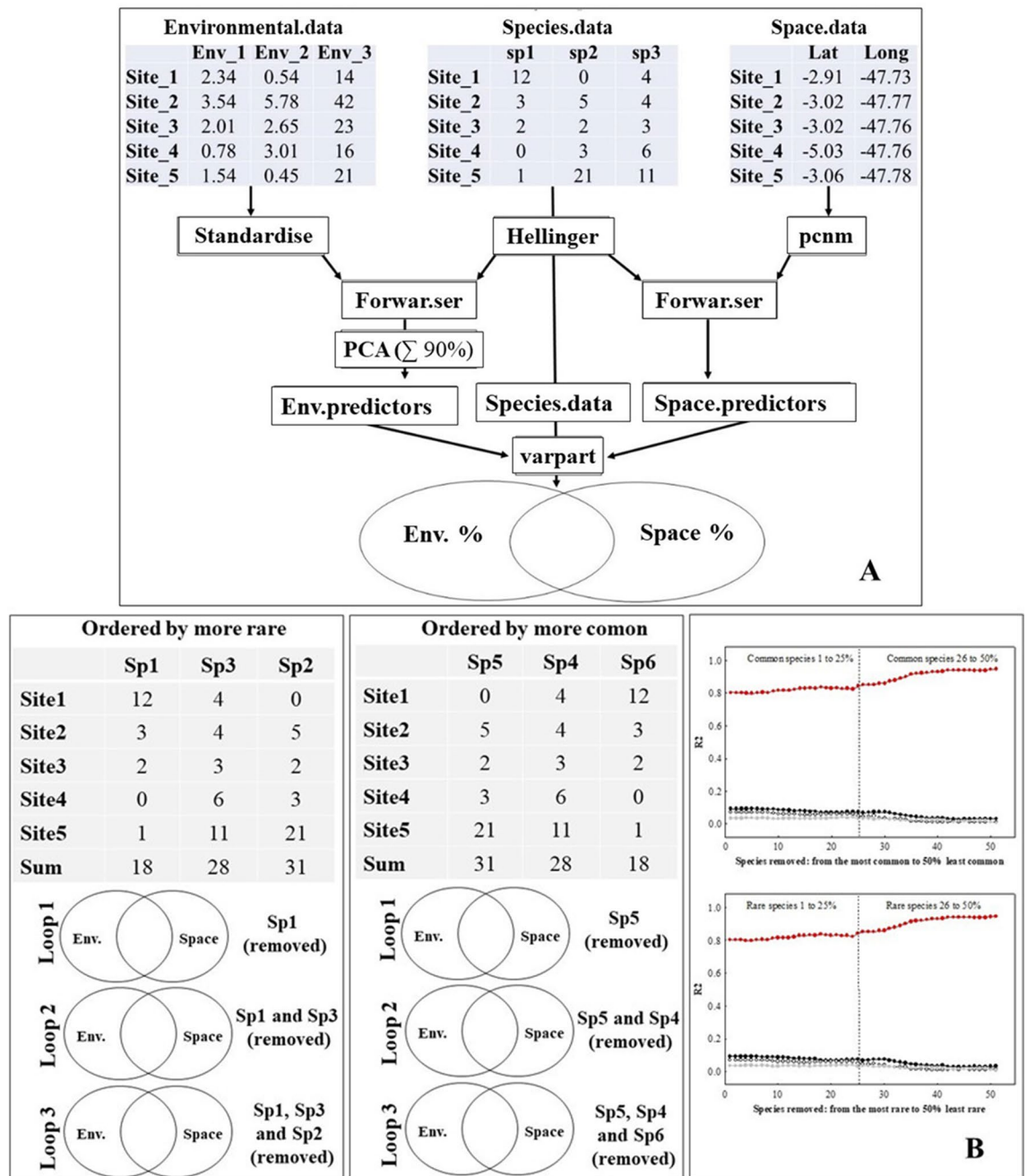
Figure 5. (**A**) The analytic procedure of variation partitioning using environmental and spatial predictors and species composition. The graphic model illustrates the abundance-based species composition data. (**B**) The analytic procedure used to analyze the effects of environmental predictors (black circles), spatial predictors (empty circles), the interaction term between environmental and spatial predictors (grey circles), and the residual portion (red circles), while gradually removing the rarest and most common species. *Env* environmental.

Neighbour Matrices (PCNM). Then, we used the set of vectors as the spatial predictors. To select which vectors are important predictors for the communities, we performed a forward selection model using the PCNMs and the community matrices of Zygoptera and Fish, respectively (Fig. 5A). We used two sets of spatial predictors (PCNMs), one for each taxonomic group. Three PCNMs were selected as spatial predictors for Zygoptera and two for Fish.

Data analysis. To test the hypothesis that the residual portion of the model would decrease as rare species were removed, we performed the models for each taxonomic group (Zygoptera and Fish). For each taxonomic group, we used the abundance matrix to perform variation partitioning (varpart function from the vegan package in R) using the whole community and the environmental and spatial predictors (Fig. 5A). This was con-

6

sidered the reference value. Then, we built a function where in each loop one species was removed from the community matrix and then the variation partitioning was performed again. The procedure of species removal was executed from the rarest species to the most common species and from the most common species to the rarest species. The species were removed until the resulting richness was equal to 50% of the original community richness. To facilitate interpretation of the results, we created a correlogram using the proportion of variation explained by the environmental, spatial, interaction term, and residual components of each routine step and of each dataset, with the number of species removed from the analysis (Fig. 5B). We tested our hypothesis that residual variance decreases when rare species are removed by contrasting the difference in residual variation before and after removing the rare species ($\Delta R^2 = R^2$Before-R^2After) against a null model in which we took caution to perform a stratified sample of our original community, in order to create subsamples that resemble the distribution of relative abundances in the original community. Species from our subsamples were removed randomly and we calculated the difference in residual variation before and after removing 50% of the community (ΔR^2). We repeated this procedure 999 times to create a distribution of values for ΔR^2. We then contrasted our values obtained when removing rare fish and Odonata species against this distribution to obtain a measure of significance for the effect of removing rare species on the proportion of residual variation.

References

1. Hutchinson, G. E. Homage to santa rosalia or why are there so many kinds of animals?. *Am. Nat.* **93**, 145–159 (1959).
2. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton University Press, Princeton, 2001).
3. Leibold, M. A. *et al.* The metacommunity concept: A framework for multi-scale community ecology. *Ecol. Lett.* **7**, 601–613 (2004).
4. Legendre, P. & Legendre, L. *Numerical Ecology* (Elsevier, Amsterdam, 2012).
5. Cottenie, K. Integrating environmental and spatial processes in ecological community dynamics. *Ecol. Lett.* **8**, 1175–1182 (2005).
6. Dray, S., Legendre, P. & Peres-Neto, P. R. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecol. Model.* **196**, 483–493 (2006).
7. Leibold, M. A. & Chase, J. M. *Metacommunity Ecology* (Princeton University Press, Princeton, 2018).
8. Grinnell, J. Field tests of theories concerning distributional control. *Am. Nat.* **51**, 115–128 (1917).
9. Elton, C. Competition and the structure of ecological communities. *J. Anim. Ecol.* **15**, 54–68 (1946).
10. Griffith, D. A. & Peres-Neto, P. R. Spatial modeling in ecology: The flexibility of eigenfunction spatial analyses. *Ecology* **87**, 2603–2613 (2006).
11. Peres-Neto, P. R., Legendre, P., Dray, S. & Borcard, D. Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology* **87**, 2614–2625 (2006).
12. Oksanen, J. *et al.* Vegan: community ecology package. R package version 2.4–5. https://cran.r-project.org/web/packages/vegan/index.html (2017).
13. Montag, L. F. A. *et al.* Contrasting associations between habitat conditions and stream aquatic biodiversity in a forest reserve and its surrounding area in the Eastern Amazon. *Hydrobiologia* **826**, 263–277 (2019).
14. Juen, L. & De Marco, P. Odonate biodiversity in terra-firme forest streamlets in Central Amazonia: On the relative effects of neutral and niche drivers at small geographical extents. *Insect Conserv. Divers.* **4**, 265–274 (2011).
15. Brasil, L. S. *et al.* Spatial, biogeographic and environmental predictors of diversity in Amazonian Zygoptera. *Insect Conserv. Divers.* **11**, 174–184 (2018).
16. Dambros, C. S., Morais, J. W., Azevedo, R. A. & Gotelli, N. J. Isolation by distance, not rivers, control the distribution of termite species in the Amazonian rain forest. *Ecography* **40**, 1242–1250 (2017).
17. Hepp, L. U., Landeiro, V. L. & Melo, A. S. Experimental assessment of the effects of environmental factors and longitudinal position on alpha and beta diversities of aquatic insects in a neotropical stream. *Int. Rev. Hydrobiol.* **97**, 157–167 (2012).
18. Siqueira, T. *et al.* Common and rare species respond to similar niche processes in macroinvertebratemetacommunities. *Ecography* **35**, 183–192 (2012).
19. Alahuhta, J. & Heino, J. Spatial extent, regional specificity and metacommunity structuring in lake macrophytes. *J. Biogeogr.* **40**, 1572–1582 (2013).
20. Heino, J. & Alahuhta, J. Elements of regional beetle faunas: Faunal variation and compositional breakpoints along climate, land cover and geographical gradients. *J. Anim. Ecol.* **84**, 427–441 (2015).
21. Algarte, V. M., Rodrigues, L., Landeiro, V. L., Siqueira, T. & Bini, L. M. Variance partitioning of deconstructed periphyton communities: Does the use of biological traits matter?. *Hydrobiologia* **722**, 279–290 (2014).
22. Brasil, L. S., Juen, L., Giehl, N. F. S. & Cabette, H. S. R. Effect of environmental and temporal factors on patterns of rarity of ephemeroptera in stream of the braziliancerrado. *Neotrop. Entomol.* **46**, 29–35 (2017).
23. Gaston, K. J. Valuing common species. *Science* **327**, 154–155 (2010).
24. Gaston, K. J. The importance of being rare. *Ecology* **487**, 46–47 (2012).
25. Lários, M. C. *et al.* Evidence of cross-taxon congruence in Neotropical wetlands: Importance of environmental and spatial factors. *Glob. Ecol. Conserv.* **12**, 108–118 (2017).
26. Juen, L. *et al.* Effects of oil palm plantations on the habitat structure and biota of streams in eastern Amazon. *River Res. Appl.* **32**, 2081–2094 (2016).
27. Legendre, P. & Anderson, M. J. Distance-based redundancy analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* **69**, 1–24 (1999).
28. Borcard, D., Gillet, F. & Legendre, P. *Numerical Ecology with R* (Springer, New York, 2018).
29. Barlow, J. *et al.* The future of hyperdiverse tropical ecosystems. *Nature* **559**, 517–526 (2018).
30. Bini, L. M., Diniz-Filho, J. A. F., Rangel, T. F., Bastos, R. P. & Pinto, M. P. Challenging Wallacean and Linnean shortfalls: Knowledge gradients and conservation planning in a biodiversity hotspot. *Divers. Distrib.* **12**, 475–482 (2006).
31. Whittaker, R. J. *et al.* Conservation biogeography: Assessment and prospect. *Divers. Distrib.* **11**, 3–23 (2005).
32. Crouzeilles, R., Feltran-Barbieri, R., Ferreira, M. S. & Strassburg, B. B. Hard times for the Brazilian environment. *Nature ecology & evolution.* **1**, 1213–1213 (2017).
33. Vieira, T. B. *et al.* A multiple hypothesis approach to explain species richness patterns in neotropical stream-dweller fish communities. *PLoS ONE* **13**, 1–17 (2018).

34. Brasil, L. S. *et al.* Net primary productivity and seasonality of temperature and precipitation are predictors of the species richness of the Damselflies in the Amazon. *Basic Appl. Ecol.* **35**, 45–53 (2019).
35. Kéry, M. & Schmid, H. Monitoring programs need to take into account imperfect species detectability. *Basic Appl. Ecol.* **5**, 65–73 (2004).
36. Leitão, R. P. *et al.* Rare species contribute disproportionately to the functional structure of species assemblages. *Proc. R. Soc. B Biol. Sci.* **2838**, 20160084 (2016).
37. Pereira, D. F. G., Oliveira-Junior, J. M. B. & Juen, L. Environmental changes promote larger species of Odonata (Insecta) in Amazonian streams. *Ecol. Ind.* **98**, 179–192 (2019).
38. Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**(6772), 853 (2000).
39. Rodrigues, M. E. *et al.* Nonlinear responses in damselfly community along a gradient of habitat loss in a savanna landscape. *Biol. Conserv.* **194**, 113–120 (2016).
40. Peel, M. C., Finlayson, B. L. & McMahon, T. A. Updated world map of the Köppen–Geiger climate classification. *Hydrol. Earth Syst. Sci. Discuss.* **4**, 439–473 (2007).
41. Brasil, L. S. *et al.* Does the damming of streams in the southern Amazon basin affect dragonfly and damselfly assemblages (Odonata: Insecta)? A preliminary study. *Int. J. Odonatol.* **17**, 187–197 (2014).
42. Lencioni, F. A. A. *The Damselflies of Brazil: An Illustrated Guide the Non Coenagrionidae Families* (All Print Editora, São Paulo, 2005).
43. Lencioni, F. A. A. *The Damselflies of Brazil: An Illustrated Guide—Coenagrionidae* (All Print Editora, São Paulo, 2006).
44. Garrison, N. & Ellenrieder, J. A. L. *Louton Damselfly Genera of the New World: An Illustrated and Annotated Key to the Zygoptera University Press* (Johns Hopkins, Baltimore, 2010).
45. Frissell, C. R., Liss, W. J., Warren, C. E. & Hurley, M. D. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. *Environ. Manag.* **10**, 199–214 (1986).
46. Espírito-Santo, H. M. V., Magnusson, W. E., Zuanon, J., Mendonça, F. P. & Landeiro, V. L. Seasonal variation in the composition of fish assemblages in small Amazonian forest streams: Evidence for predictable changes. *Freshw. Biol.* **54**, 536–548 (2009).
47. Uieda, V. S. & Castro, R. M. C. *Coleta e fixação de peixes de riachos: Ecologia de peixes de riacho* (OecologiaAustralis, Rio de Janeiro, 1999).
48. Planquette, P., Keith, P. & Bail, P. Y. L. *Atlas des poissons d'eau douce de Guyane* (Service du patrimoine naturel, Paris, 1996).
49. Albert JS. *Species Diversity and Phylogenetic Systematics of American Knifefishes (Gymnotiformes, Teleostei).* (Miscellaneous Publications of the Museum of Zoology of the University of Michigan, Michigan, 2001).
50. Kaufmann, P. R., Levine, P., Robison, E. G., Seeliger, C. & Peck, D. V. *Quantifying Physical Habitat in Wadeable Streams* (U.S. Environmental Protection Agency, Washington, 1999).
51. Peck, D. V. *et al. Invironmental Monitoring and Assessment Program-Surface Waters: Western Pilot Study Field Operations Manual for Wadeable Streams* (U.S. Environmental ProtectionAgency, Washington, 2006).
52. De Marco, P. & Nobrega, C. C. Evaluating collinearity effects on species distribution models: An approach based on virtual species simulation. *PLoS ONE* **13**, 1–25 (2018).
53. Legendre, P. Spatial autocorrelation: Trouble or new paradigm?. *Ecology* **74**, 1659–1673 (1993).

Author contributions

L. S.B. wrote the main manuscript text, prepared figures and reviewed the manuscript. T. B. V. wrote the main manuscript text, prepared figures and reviewed the manuscript. R. C. B. wrote the main manuscript text, prepared figures and reviewed the manuscript. L. F. A. M. wrote the main manuscript text, prepared figures and reviewed the manuscript. L. J. wrote the main manuscript text, prepared figures and reviewed the manuscript. A. F. A. A. wrote the main manuscript text, prepared figures and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-76833-5.

Correspondence and requests for materials should be addressed to L.S.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.