Highlights

We assessed congruence patterns of macroinvertebrates and fish and their drivers

We evaluate community-environmental relationships and their drivers

Low-resolution data serve as surrogates of high-resolution data

Considering cost-effectiveness, we recommend the use of coarser taxonomic resolution

The shortcuts proposed here support large-scale biomonitoring programs in Brazil.

Cost-effective alternatives to facilitate biomonitoring and bioassessment of Neotropical streams

Francisco Valente-Neto[1*], Renato B. Dala-Corte[2,3], Almir M. Cunico[4], André L. B. Magalhães[5], Bruno S. Godoy[6], Cecília G. Leal[7], Diego M. P. Castro[8], Diego R. Macedo[9], Dilermando P. Lima-Junior[10], Éder A. Gubiani[11], Fabio de O. Roque[12,13], Fabrício B. Teresa[14], Fagner Junior Machado de Oliveira[10], Fernando G. Becker[15], Gabriel L. Brejão[16], Janaina G. de Brito[17], Jansen Zuanon[18,19], Jean R. S. Vitule[20], Karina Dias-Silva[21], Lilian Casatti[22], Luciano Benedito de Lima[10], Luciano F. A. Montag[23], Marcos Callisto[8], Mireile R. dos Santos[24], Neusa Hamada[25], Paulo A. Z. Pamplin[26], Paulo S. Pompeu[27], Rafael P. Leitão[28], Renata Ruaro[29], Sheyla R. M. Couceiro[30], Vinícius Abilhoa[31], Yzel R. Súarez[32], Renato T. Martins[25]

[1] Laboratório de Interações Multitróficas e Biodiversidade, Departamento de Biologia Animal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas-SP, Brazil

*fvalenteneto@gmail.com

[2] Departamento de Ecologia, Universidade Federal de Goiás, Goiânia-GO, Brazil.

[3] Princeton School of Public and International Affairs, Princeton University, Princeton-NJ, USA

[4] Laboratório de Ecologia, Pesca e Ictiologia, Departamento de Biodiversidade, Setor Palotina, Universidade Federal do Paraná, Palotina-PR, Brazil

[5] Programa de Pós-Graduação em Ecologia de Biomas Tropicais, Departamento de Biodiversidade, Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto (UFOP), Ouro Preto-MG, Brazil

[6] Núcleo de Ecologia Aquática e Pesca da Amazônia, Universidade Federal do Pará, Belém-PA, Brazil

[7] Lancaster University, Lancaster, UK

[8] Laboratório de Ecologia de Bentos, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil

[9] Departamento de Geografia, Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil

[10] Laboratório de Ecologia e Conservação de Ecossistemas Aquáticos, Universidade Federal de Mato Grosso, Pontal do Araguaia-MT, Brazil

[11] Laboratório de Ictiologia e Estatística Pesqueira, Instituto Neotropical de Pesquisas Ambientais, Grupo de Pesquisas em Recursos Pesqueiros e Limnologia, Programa de Pós-Graduação em Conservação e Manejo de Recursos Naturais, Programa de Pós-Graduação em Recursos Pesqueiros e Engenharia de Pesca, Universidade Estadual do Oeste do Paraná, Toledo-PR, Brazil.

[12] Universidade Federal de Mato Grosso do Sul, Campo Grande-MS, Brazil

[13] Knowledge Center on Biodiversity, Belo Horizonte-MG, Brazil

[14] Laboratório de Biogeografia e Ecologia Aquática, Universidade Estadual de Goiás, Anápolis-GO, Brazil

[15] Laboratório de Ecologia de Paisagens, Departamento de Ecologia, Universidade Federal do Rio Grande do Sul, Porto Alegre-RS, Brazil

[16] Universidade Estadual Paulista (UNESP), Instituto de Biociências, Rio Claro-SP, Brazil

[17] Secretaria de Educação do Estado de Mato Grosso, Diretoria Regional de Educação, Barra do Garças-MT, Brazil

[18] Instituto Nacional de Pesquisas da Amazônia, Manaus-AM, Brazil

[19] Universidade Santa Cecília, Santos-SP, Brazil

[20] Laboratório de Ecologia e Conservação, Setor de Tecnologia, Departamento de Engenharia Ambiental, Universidade Federal do Paraná, Curitiba-PR, Brazil

[21] Programa de Pós-Graduação em Ecologia, Universidade Federal do Pará, Belém-PA, Brazil

[22] Universidade Estadual Paulista (UNESP), Departamento de Ciências Biológicas, São José do Rio Preto-SP, Brazil

[23] Laboratório de Ecologia e Conservação, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém-PA, Brazil

[24] Departamento de Ciências e Biologia, Colégio Pedro II, Campus São Cristóvão II, Rio de Janeiro-RJ, Brazil

[25] Coordenação de Biodiversidade, Instituto Nacional de Pesquisas da Amazônia, Manaus-AM, Brazil

[26] Instituto de Ciência e Tecnologia, Universidade Federal de Alfenas, Poços de Caldas-MG, Brazil

[27] Departamento de Ecologia e Conservação, Universidade Federal de Lavras, Lavras-MG, Brazil

[28] Laboratório de Ecologia de Peixes, Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brazil.

[29] Laboratório de Biomonitoramento e Ecologia Aplicada, Programa de Pós-graduação em Ciência e Tecnologia Ambiental, Universidade Tecnológica Federal do Paraná, Curitiba-PR, Brazil.

[30] Laboratório de Ecologia e Taxonomia de Invertebrados Aquáticos, Universidade Federal do Oeste do Pará, Santarém-PA, Brazil.

[31] Museu de História Natural Capão da Imbuia, Prefeitura de Curitiba, Curitiba-PR, Brazil.

[32] Laboratório de Ecologia, Centro de Estudos em Recursos Naturais, Universidade Estadual de Mato Grosso do Sul, Dourados-MS, Brazil

Abstract

A reliable response of bioindicators to environmental variation is a cornerstone for effective bioassessment and biomonitoring. Fish and aquatic macroinvertebrates are widely used as bioindicators of different human impacts in freshwater ecosystems, but the cost-effectiveness of their usage can be improved through the use of surrogates. We investigated congruence patterns between using different taxonomic and numeric resolutions for aquatic macroinvertebrates and fish to assess community-environment relationships. We also tested whether dataset characteristics (e.g., area sampled, species pool) could explain the variation in the effectiveness of using different taxonomic and numerical resolutions. We used a Brazilian nationwide database encompassing multiple datasets with a gradient of riparian deforestation each. Our findings suggest that families and genera can effectively represent macroinvertebrate genera and fish species, respectively, when using community matrices for assessing community-environment relationships, with an acceptable loss of information. EPT (Ephemeroptera, Plecoptera, and Trichoptera) and Characiformes or Siluriformes may be used as a surrogate, in some cases, for the entire assemblages of macroinvertebrates and fish, respectively, but their use may result in higher loss of information. Presence-absence data also presented a minimal loss of information compared to abundance data, for both macroinvertebrates and fish. The variation in congruence levels among macroinvertebrate datasets was less predicted by dataset characteristics than fish. Across distinct resolutions, on average, 10% and 19% of the variation in community composition of macroinvertebrates and fish, respectively, was explained by broad-scale environmental variables, and the effect size was negatively affected by the dataset's sample size and spatial extent for fish. Whereas identification at species (fish) and genus (macroinvertebrates) level and quantification of all individuals still provide the best scenario, we provide evidence that coarser taxonomical resolution and presence-absence data can be used as cost-effective alternatives to facilitate biomonitoring and bioassessment of freshwaters in the Neotropical region impacted by deforestation.

Keywords: Amazon, Atlantic Forest, Aquatic insects, Correlation, Fish, Surrogates.

1. Introduction

Freshwater ecosystems are species-rich and provide multiple ecosystem services to human society, such as water for drinking and irrigation, food provision and recreation (Lynch et al., 2023). These ecosystems have been widely modified by human activities for a long time, resulting in negative effects on biological communities (Reid et al., 2019). Biological monitoring and assessment are essential for managing and controlling water quality, which rely on the detection of strong associations between metrics used to describe bioindicators and environmental variables related to anthropogenic impacts (Bonada et al., 2006). Bioassessment and biomonitoring have a long history in the evaluation of freshwater environments (Doledec and Statzner, 2010; Ruaro and Gubiani, 2013) and have been included in a recent global agenda for advancing freshwater biodiversity research (Maasri et al., 2022). There is an urgent demand to define indicators and monitor ecosystems from the United Nations 2030 agenda for sustainable development and the 2030 targets of the Convention on Biological Diversity (Leadley et al., 2022). However, our capacity to efficiently assess and monitor freshwaters is still challenging, mainly in highly diverse tropical regions, due to multiple and interacting human impacts, high costs, lack of basic knowledge on taxonomy and species distribution, and uncertainty and variability in biological responses (Barlow et al., 2018; Dala-Corte et al., 2020; Firmiano et al., 2021; Siqueira et al., 2020; Sousa et al., 2023).

Riparian loss has led to changes in community composition and pronounced declines in biodiversity (Albert et al., 2021; Dornelas et al., 2014; Jähnig et al., 2021; Pelicice et al., 2017). Aquatic macroinvertebrates and fish are regarded as reliable indicators of the effects of riparian vegetation loss and are commonly used in bioassessment and monitoring studies (Dala-Corte et al., 2020; Feio et al., 2023, 2021; Ruaro and Gubiani, 2013; Valente-Neto et al., 2021). Aquatic macroinvertebrates are generally assumed to be more sensitive to the effects of riparian vegetation loss than fish (e.g., Dala-Corte et al., 2020; Martins et al., 2022; Valente-Neto et al., 2021), but no consensus has been reached. In addition to reliable bioindicator response, low costs for sampling, sorting, and identifying the biological group are critical factors for improving the cost-effectiveness of bioassessment and biomonitoring (Bonada et al., 2006; Ruaro et al., 2024; Valente-Neto et al., 2021).

The use of different taxonomic (Lopes et al., 2021; Martins et al., 2022; Melo, 2005; Santos et al., 2022) and numerical resolutions (Lopes et al., 2021; Martins et al., 2022; Melo, 2005; Ribas and Padial, 2015; Santos et al., 2022; Valente-Neto et al., 2018), and the use of biodiversity surrogates (Faquim et al., 2021; Roque et al., 2017; Siqueira et al., 2012a; Valente-Neto et al., 2018) can be cost-effective measures in bioassessment and biomonitoring. Taxonomic resolution can be defined as the level of detail at which organisms are identified, while numeric resolution is the detail in representing numeric values, i.e., abundance or presence-absence data. The existence of strong relationship (correlations greater than 0.7 – Heino, 2010) between taxonomic (e.g., species vs. family) and numerical resolution (abundance vs. presence-absence), or taxonomic groups (e.g., Ephemeroptera, Plecoptera, and Trichoptera – EPT, hereafter – vs. macroinvertebrates) indicate that one taxonomic level, numerical resolution or taxonomic group can be use as surrogate in some cases when a detailed assessment is unfeasible. Such strong associations are assumed to indicate minimal loss of crucial information, ensuring reliable biomonitoring (Heino, 2010).

The detection of significant and highly correlated relationships is required for using simplifications in biomonitoring programs, and indicate that different taxonomic and numerical resolutions respond similarly to anthropogenic impacts (Heino, 2010). A coarser taxonomic level (e.g., family) is easier and faster to identify compared to a refined taxonomic level (e.g., species or genera). Additionally, presence-absence data do not require counting all individuals sampled, speeding up data acquisition, or even reducing the need for sampling large number of animals in the field. Therefore, these approaches save time and funds in biomonitoring efforts. However, studies assessing these simplifications for biomonitoring impacts of riparian vegetation loss on streams are limited to local or regional scales (Brito et al., 2018; Landeiro et al., 2012; Ribas and Padial, 2015; Santos et al., 2022; Valente-Neto et al., 2018). Evaluating a nationwide database that encompasses multiple biomes would enable the identification of robust patterns and broader generalizations. This is particularly important in Brazil, one of the most biologically diverse countries, including six terrestrial biomes, two of which are biodiversity hotspots (Atlantic Forest and Cerrado).

To advance in the use of congruence patterns in bioassessment and biomonitoring, it is important to understand the factors contributing to wide variations in correlations between taxonomic and numerical resolutions (Lopes et al., 2021; Ribas and Padial, 2015) and community-environment relationships (Heino et al., 2015). Dataset characteristics significantly influence species composition patterns (e.g., matrix fill) and the processes driving them. For example, increasing the regional species pool, particularly through the inclusion of rare species, adds complexity to the ecological data and reduces the clarity of congruence patterns and the strength of associations between species composition and environmental variables (Heino et al., 2015; Podani and Schmera, 2011; Siqueira et al., 2012b). Additionally, sample size, spatial extent, and environmental variability are positively correlated, with larger sample size and broader spatial extents generally enhancing community-environment relationship (Leibold and Chase, 2018). The range of environmental gradients assessed also affects congruence, with wider gradients better capturing species responses compared to narrower ones (Viana and Chase, 2019). Despite other studies have synthesized cross-taxon congruence in terrestrial (Westgate et al., 2014) and marine (Mellin et al., 2011) ecosystems, a synthesis on the potential drivers of congruence and variation across freshwater ecosystem datasets is still lacking. Such an understanding is needed for improving the use of surrogates in monitoring and assessing rivers and streams, especially when evaluating human impacts.

Our study had three main objectives. First, we investigated the congruence patterns in using different taxonomic and numerical resolutions for aquatic macroinvertebrates and fish of Neotropical streams. Considering that species within taxonomic levels often share similar ecological preferences (Warwick, 1993), we hypothesized that coarse taxonomic resolution is congruent with refined taxonomic resolution. Specifically, we expected stronger congruence between closely related taxonomic levels (e.g., higher correlation between genus and family than between genus and order). Given the high species turnover in the neotropics, we expected a high congruence between abundance and presence-absence data. Second, we assessed the relationship between community composition and environmental variables to understand how much of the variation in composition is explained by environmental variables. We expected that environmental variables partially explain a portion of the variation in aquatic biodiversity composition, with the refined taxonomic resolution and abundance data showing greater explanatory power. Third, we evaluated how the datasets characteristics (e.g., area sampled, taxa rarity and dominance, and regional taxa pool) could influence the variation in the congruence patterns and community-environment relationships. We hypothesized that these characteristics help explain the variation observed between datasets. To investigate these objectives, we used a comprehensive database of aquatic macroinvertebrates and fish sampled along a gradient of riparian vegetation loss in four biomes of Brazil, including three of the largest biomes (Amazon, Atlantic Forest, and Cerrado) and the comparatively smaller Pampa (Dala-Corte et al., 2020).

2. Methods

2.1 Database

We used a recent database compiled by Dala-Corte et al. (2020) comprising aquatic macroinvertebrate and fish communities sampled in streams across Brazil. This database consists of datasets specifically designed to capture variation in riparian vegetation loss. Streams included in the datasets are subject to the effects of riparian vegetation loss within forest-agriculture landscapes, with other anthropogenic impacts less evident. From the database, we selected 10 aquatic macroinvertebrate and 17 fish datasets following the criteria described below, comprising 700 and 1,290 stream communities, respectively (Fig. S1). We included datasets that i) had abundance data; ii) had at least 20 sampling sites; iii) sampled small streams (10-m wide or less) and iv) assessed the entire community composition, i.e., all taxonomic groups belonging to macroinvertebrates and fish. Additional specific criteria were used according to the taxonomic group. For macroinvertebrates, we included datasets that v) met the criterion of identifying at least 50% of individuals, as well as all the EPT taxa, to the genus level, with the remaining macroinvertebrate taxa identified at least to the family level. The genus level represents the most refined taxonomic resolution commonly achievable in Neotropical freshwater studies because reliable identification to species level often requires the adult stage (Hamada et al., 2018). Some datasets used family level because of restricted taxonomic knowledge for some groups (Hamada et al., 2018) and/or because the refined identification was focused on a subset of macroinvertebrates only (e.g., EPT, EPT + Coleoptera). On average 78% of macroinvertebrate taxa were identified to genus level in all the datasets compiled (minimum = 53%; maximum = 100%). The database used here does not allow testing cross-taxon congruence (EPT vs. fish), because each dataset collected exclusively one group. For fish, we included datasets that vii) identified all taxa to species level. All the fish species were included in the analysis, and we did not exclude non-native species.

The following biomes were represented in the database used here: Macroinvertebrates: Amazon (four datasets), Cerrado (three), Atlantic Forest (three); Fish: Amazon (four datasets), Cerrado (seven), Atlantic Forest (five), Pampa (one) – see Table S1 for the list of datasets included.

2.2 Congruence between taxonomic and numerical resolutions

We used the Mantel correlation to investigate the congruence among community matrices reorganized into different numeric and taxonomic resolutions for both macroinvertebrates and fish (Fig. 1). Numeric resolution comprised the community matrix filled either with abundance data (i.e., number of individuals per taxa) or with presence-absence (i.e., 0-1) data. Correlations between the numeric resolution matrices were only performed comparing presence-absence data to fish abundance matrices identified at species level and to macroinvertebrate abundance matrices identified at genus level. Taxonomic resolution comprised matrices filled with taxon individuals (i.e., abundance data) clustered into different taxonomic levels. For macroinvertebrates, genus was clustered into family, order, or EPT (Ephemeroptera, Plecoptera, and Trichoptera) matrices. EPT is a bioindicator group commonly used in stream assessment and biomonitoring (Barbour et al., 1999; Buss et al., 2015), although other groups may better indicate anthropogenic impacts in certain contexts (Serra et al., 2017). For EPT subset, we tested correlations for genus as taxonomic resolution only. For fish, the species were grouped into genus, family, order, or were subset into Characiformes or Siluriformes. The last two orders were used because they are the most common fish taxa in the Neotropical region (Reis et al., 2016), and were frequently used as indicators of stream system conditions, especially related to landscape alterations (e.g., Carvalho et al., 2017). Given that we had 10 aquatic macroinvertebrate and 17 fish datasets, we calculated 10 and 17 Mantel correlations between all the matrices reorganized into different resolutions. Then, the values were averaged to obtain a single overall correlation for each comparison, but the variation was used in subsequent analyses. We used the mantel function from the vegan package within the R environment.

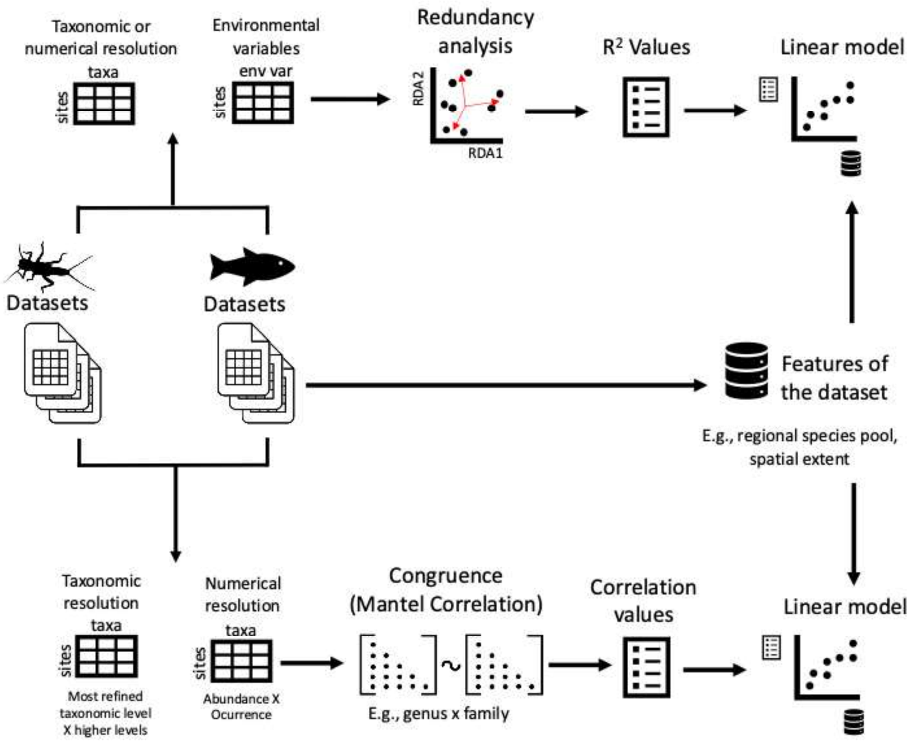


Fig. 1. Diagram depicting the statistical analyses employed in this study. We used multiple datasets of aquatic macroinvertebrates and fish to assess congruence patterns between different taxonomic resolutions and between abundance and occurrence data (first aim). We also investigate community-environment relationship (second aim) and dataset characteristics that could account for the variation in the effectiveness of using different taxonomic and numerical resolutions, and the community-environment relationships (third aim). To address the first aim, we used the Mantel test. The analysis was conducted separately for each dataset and for each comparison of composition matrices (e.g., macroinvertebrates genus x family), resulting in a correlation value for each comparison. Since multiple datasets were analyzed, the outcome was a vector containing correlation values for each composition matrix comparison. For understanding community-environment relationship, we employed redundancy analysis (RDA). The composition matrix (e.g., fish species) was the response variable, and predictors were composed by landscape and climatic variables. This analysis resulted in a vector with values of variation explained by the RDA models (adjusted R^2) for each composition matrix. Regarding the third aim, we fitted multiple linear models using either the correlation values or the RDA explanation variance for different taxonomic and numeric resolutions as the response variables

and the dataset features as predictors. We assessed the importance of these dataset features in the models using AICc and variable importance.

2.3 Community composition analysis

Subsequently, we used redundancy analysis (RDA, hereafter) to evaluate how environmental variables (predictors) explain the variation in community composition (response variables) (Fig. 1). RDA is a constrained ordination and a multivariate extension of multiple regression, making it well-suited for analyzing community-environment relationships. We previously transformed the abundance data using Hellinger transformation, which is suited to species abundance data as it gives low weights to variables with low counts and zero-inflated (Legendre and Gallagher, 2001). To assess the consistence of community-environment relationships across datasets, we used landscape and climate environmental variables as explanatory variables, and each matrix derived from the different numeric and taxonomic resolution matrices as response variables. Local environmental factors, such as water chemistry and instream habitat conditions were not available for all datasets and so we included only broad scale environmental variables that could be used across all datasets. We included 10 exploratory variables in the RDA models, (Table 1; Table S1), and selected them based on their ecological relevance to stream biodiversity (Table 1) and to ensure comparability across all datasets. In addition, to avoid model inflation, we excluded climate variables with correlation values higher than 0.8. Overall, a wide range of values were present in the landscape and climate predictors within each dataset (Table S1). For instance, the percentage of native vegetation cover within a 500-m radius buffer around sampled sites varied significantly, ranging from between deforested (0-17%) to fully forested streams (~100%) (Table S1). We used the rda function from the vegan package within the R environment.

Table 1. Description and source of environmental variables used in the community-environment relationship analysis (RDA, redundancy analysis models).

Variable	Description	Ecological effects in streams	Source
Altitude	Site altitude	Environmental (e.g., temperature and oxygen) and flow conditions (Jacobsen, 2008). Community structure and richness (Jacobsen, 2008).	SRTM (Shuttle Radar Topographic Mission)
Slope	Surface terrain slope measured as the ratio of the height difference between two points and their horizontal distance using altitude data	Environmental and flow conditions (e.g., current velocity; sediment deposition and substrate stability) (Allan and Castillo, 2007).	SRTM (Shuttle Radar Topographic Mission)
Density	Drainage network density, i.e., stream length per area	Runoff, connectivity between streams, aquatic insects distribution and aquatic subsidies (Kopp and Allen, 2019).	HydroSHEDS project (Lehner et al., 2008)
Hand50	Vertical distance of the water table calculated as the HAND (Height Above the Nearest Drainage), indicating the availability of groundwater	Flow conditions (flood, discharge and streamflow) (Nobre et al., 2011).	AMBDATA (https://www.dpi.inpe.br/Ambdata/download.php)
SA2001Tree	Tree cover percent, indicating forest canopy cover derived from MODIS (MODerate-resolution Imaging Spectroradiometer) for the sampling site at 1 km^2 resolution	Organic matter input (e.g., leaf and wood), light penetration, stabilization of stream bed and banks (Allan, 2004). Community structure (Allan, 2004).	MODIS (MODerate-resolution Imaging Spectroradiometer) global data acquired between 2000 and 2001, at 1 km pixel size, also obtained from AMBDATA.

PercMA500m	Percentage of native vegetation cover within a 500-m radius buffer around sampled sites	Organic matter input (e.g., leaf and wood), filters pollutants, stabilizes banks, and maintains water quality. Community structure (Allan, 2004; Dala-Corte et al., 2020).	MapBiomas (http://mapbiomas.org), which was the scale at which macroinvertebrates and fish were most responsive in Dala-Corte et al. (2020b)
BIO1	Annual mean temperature	Physiology, metabolism, size and community structure (Bonacina et al., 2023).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO2	Mean diurnal temperature range	Large diurnal temperature fluctuations can stress aquatic organisms, influence respiration rates, and alter predator-prey interactions, affecting community stability (Bonacina et al., 2023).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO12	Annual precipitation	Drives periodic changes in streamflow, impacting life cycles, migration patterns, and resilience of aquatic communities to environmental fluctuations (Allan and Castillo, 2007; Carvallo et al., 2022).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO15	Precipitation seasonality (coefficient of variation)	Drives periodic changes in streamflow, impacting life cycles, migration patterns, and resilience of aquatic communities to environmental fluctuations (Allan and Castillo, 2007; Carvallo et al., 2022).	WorldClim - Global Climate Data (https://www.worldclim.org/)

2.4 Variation between datasets

To assess the sources of variation observed among datasets in terms of the correlations of distinct taxonomic and numeric resolutions, as well as in the variation explained by the RDA models, we considered six features describing the datasets: i) number of sites sampled in the dataset (Samp.Size); ii) extension of the area sampled calculated as the convex hull (Area in km^2); iii) variation in the environmental variables of the sites sampled, calculated as the summed coefficient of variation (CV) for all environmental variables used in the RDA models (Env.CV); iv) species pool size of the dataset, expressed as species richness for fish and family richness for macroinvertebrates (Nspp or Nfam); v) proportion of rare species in the dataset represented by singletons and doubletons (Rares); vi) species dominance (Domin) was calculated using the Simpson index (D) with the vegan package (Oksanen et al., 2022). Since the vegan diversity function computes 1 - D, we adjusted the result to obtain D. In general, these six features varied significantly for both macroinvertebrates and fish datasets (Table S2; Table S3). For example, the number of sampled sites ranged from 35 to 110 for the macroinvertebrate dataset (Table S2) and from 24 to 232 for the fish dataset (Table S3).

The variation among datasets was investigated using both the results from the congruence values expressed as correlations (subsection 2.2) and adjusted R^2 from RDA models (subsection 2.3) as effect sizes. Thus, we fitted multiple linear models using as response variables either the correlation values of the distinct taxonomic and numeric resolutions or the RDA explanation variance for the distinct datasets (Fig. 1). The dataset's features described above were used as predictors in these models. First, we evaluated the variance inflation factor (vif) of the models with the car package (Fox and Weisberg, 2019) to remove multicollinearity in the models, ensuring that selected variables contribute independently to explain variation in correlations and adjusted R^2 from RDA models. Models for fish had no variables with high vif (all vif values were <5), while for macroinvertebrates the proportion of rare species had a vif greater than 5 and was therefore excluded from the analysis. We then assessed the importance of these dataset features in the models using Akaike Information Criterion for small sample size (AICc) and variable importance. For this, we used the dredge function of the MuMIn package (Bartoń, 2022) to generate all possible combinations of predictors, allowing us to estimate the most plausible models. We then calculated the averaged effect size (standardized beta coefficient), which represents the weighted mean effect size of each predictor across all models in which it appeared, with weights determined by the Akaike Information Criterion (AIC). Additionally, we assessed predictor importance by summing Akaike weights for each predictor across all models. Models with ΔAICc < 2.0 and greater values of weight were considered the most plausible. We also obtained adjusted R^2 of the models as another measure of model goodness-of-fit.

3. Results

3.1 Taxonomic and numerical correlations

For aquatic macroinvertebrates, the congruence between genus and family was high (mean correlation = 0.88, range = 0.71 – 0.99; Table S4; Fig. 2A). Family and order (mean correlation = 0.76, range = 0.60 – 0.90) were more strongly correlated than genus and order (mean correlation = 0.66, range = 0.54 – 0.78). Mean correlation between macroinvertebrate genera and EPT subset was high, but varied greatly between datasets (mean correlation = 0.72, range = 0.40 – 0.90) (Table S4; Fig. 2A).

Abundance and presence-absence (PA, hereafter) data, identified at the genus level, were also highly correlated (mean correlation = 0.83, range = 0.67 – 0.95).

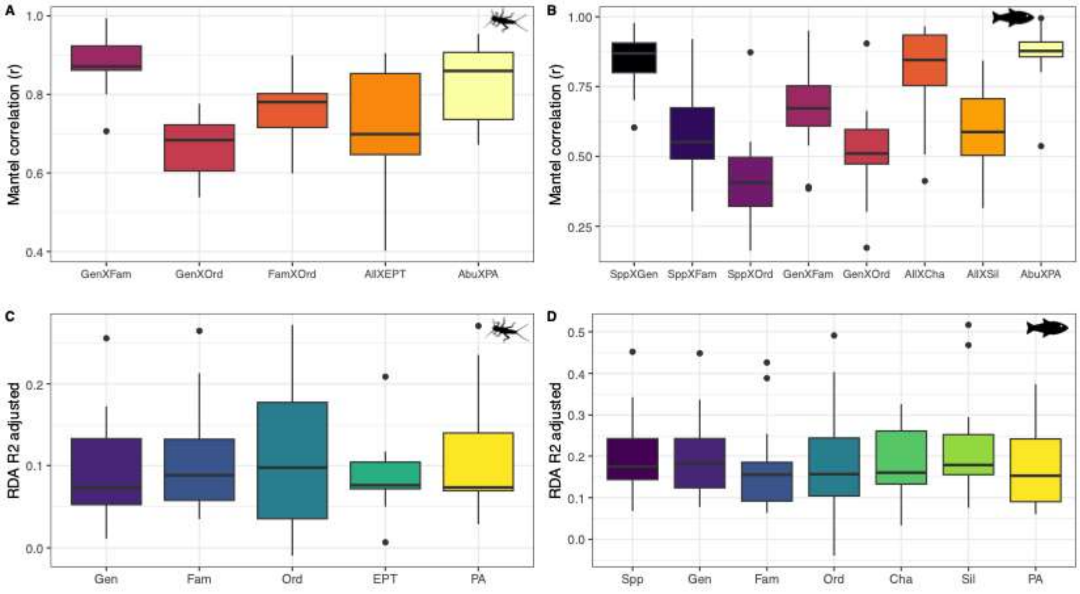


Fig. 2. Variation in Mantel correlations values across taxonomic and numeric resolutions for macroinvertebrates (A) and fish (B), as well as the community-environment relationship captured by adjusted R^2 of redundancy analysis models (RDA) for stream macroinvertebrates (C) and fish (D). Box-plots show the median (horizontal black line), interquartile range (colored rectangles), minimum (line below the interquartile range), maximum (line above the interquartile range) and outliers (values greater or lower than 1.5 times the interquartile range). Spp = Species; Gen = Genus; Fam = Family; Ord = Order; Abu = Abundance data; PA = presence-absence data; EPT = Ephemeroptera, Plecoptera, Trichoptera subset; Cha = Characiformes; Sil = Siluriformes.

For fish, the correlation between species and genus was high (mean correlation = 0.85, range = 0.60 – 0.98), but there was a low mean correlation between species and family (mean correlation = 0.58, range = 0.30 – 0.92), and between species and order (mean correlation = 0.43, range = 0.16 – 0.87; Fig. 2B; Table S5). Genus and family (mean correlation = 0.68, range = 0.38 – 0.95) and genus and order (mean correlation = 0.52, range = 0.17 – 0.90) had also low mean correlations. Total species in the fish assemblage were more correlated with Characiformes (mean correlation = 0.81, range = 0.41 – 0.97) than with Siluriformes species (mean correlation = 0.59, range = 0.32 – 0.84). Fish species abundance and PA data were usually highly correlated (mean correlation = 0.87, range = 0.54 – 1.00).

3.2 Community-environment relationships

Overall, the relationship between macroinvertebrate composition and environmental variables was similar regardless of the different taxonomic and numeric resolutions tested (mean adjusted R^2 = 9.0-11.6%; Table 2; Fig. 2C; Table S6). Environmental variables related to deforestation were the mains drivers of the macroinvertebrate community-environment relationship (Fig. S2). For fish communities, the amount of variation explained by the environmental variables did not differ clearly between the taxonomical and numerical resolutions (mean R^2 = 17.1 – 21.99%; Table 2; Fig. 2D; Table S7). Climate, landscape, and environmental variables related to deforestation were the mains drivers of the fish community-environment relationship (Fig. S3).

Table 2. Redundancy Analysis (RDA) of environmental variables explaining stream macroinvertebrate or fish communities identified at species, genus, family or order, all using abundance data, and presence-absence data (PA). Mean adjusted R^2 is the mean explanation obtained from 17 datasets for fish and 10 datasets for macroinvertebrates from Brazilian Amazon, Atlantic Forest, Cerrado, and Pampa biomes. Max = Maximum; Min = Minimum; EPT = Ephemeroptera, Plecoptera, and Trichoptera.

Response variables	Mean adj-R^2 (%)	Min adj-R^2	Max adj-R^2
Macroinvertebrate genus	10.0	1.1	25.5
Macroinvertebrate family	11.2	3.5	26.4
Macroinvertebrate order	11.0	0	27.2
Macroinvertebrate PA	11.6	3.0	27.0
EPT	9.0	0.7	20.9
Fish species	20.0	7.0	45.2
Fish genus	20.1	7.76	44.8
Fish family	17.1	6.3	42.6
Fish order	18.4	0	49.1
Fish PA	17.8	6.0	37.4
Characiformes	17.5	3.3	32.6
Siluriformes	21.9	7.6	51.6

3.3 Predictors of variation in taxonomic and numerical correlations

For macroinvertebrates, the variation in the correlation levels across different taxonomical and numerical resolutions was poorly explained by the predictors, with most models not differing from the null expectation (Fig. 3A; Table S8). However, the variation in correlations between genus and family, as well as between abundance and PA data was explained by species dominance in the community (Domin) within each dataset (Fig. 3A; Table S9). The effect was positive for genus vs. family correlations (average coefficient = 0.74; summed weight = 0.84) and negative for abundance vs. PA data (average coefficient = -0.87; summed weight = 0.99) (Fig. 3A; Table S9).

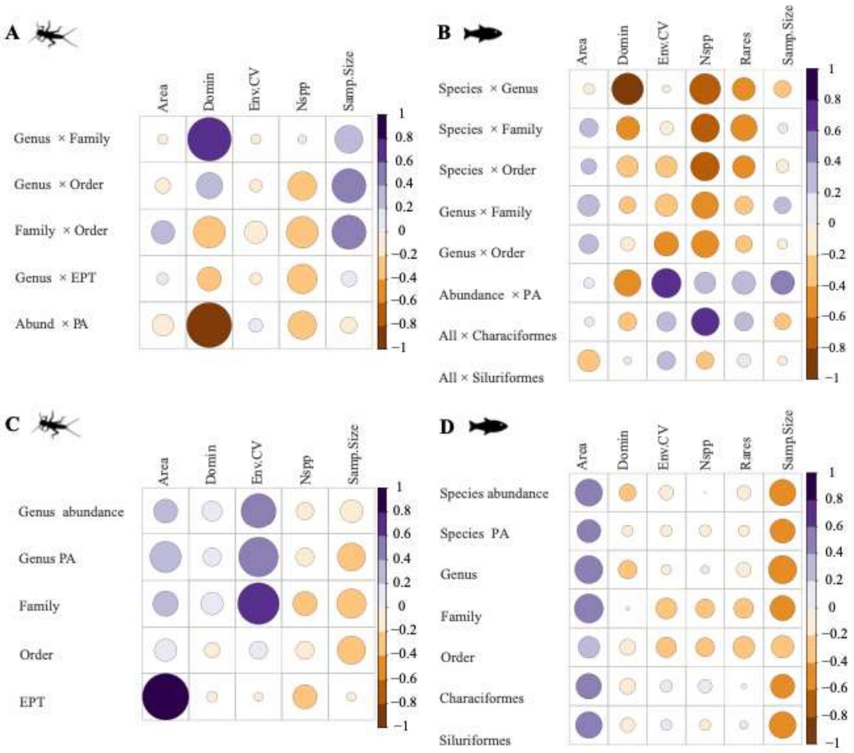


Fig 3. Effect size and direction (average coefficients) for predictors (dataset characteristics) of the variation in congruence patterns for A) macroinvertebrates and B) fish, and in adjusted R^2 of redundancy analysis for C) macroinvertebrates and D) fish. The circle size represents the strength of the correlation, while the color (purple-brown gradient) indicates the direction of the correlation. Samp.Size = number of sites sampled in the dataset; Area = extension of the area sampled calculated as the convex hull; Env.Cv = variation in the environmental variables of the sites sampled calculated as the summed coefficient of variation calculated for all environmental variables used in the RDA models; Nspp = species pool size of the dataset, expressed as total number of families; Rare = proportion of rare species in the assemblages represented by singletons and doubletons; Domin = species dominance in terms of abundance in the assemblage calculated as the Simpson index.

For fish datasets, the variation in the correlation levels was better explained by models including different predictors, although some models were not well supported (Fig. 3B; Table S10; Table S11). Domin was negatively related to matrix correlation between species vs. genus. Likewise, an overall negative effect was observed for rarity (Rare) and species pool size (Nspp) within each dataset on the matrix correlation between most taxonomic levels (Fig. 3B; Table S11). A different result was observed for species abundances vs. PA data, where increased Samp.Size (i.e., number of sites sampled) and environmental variation (Env.CV) were related to higher correlations. Also, for fish species vs. Characiformes matrices, Nspp was positively related to correlation values (Fig. 3B; Table S11).

3.4 Predictors of variation in community-environment relationship

For macroinvertebrates, the variation in the community-environment relationships using different taxonomical and numerical resolutions was poorly explained by predictors and most models did not differ from the null expectation (Fig. 3C; Table S12). The variation in the EPT community-environment relationship was explained by the extension of the area sampled (Area) (Fig. 3C; Tables S12; S13). For fish, Area and Samp.Size were selected in the most plausible models to explain variation in the community-environment relationship effect sizes (Table S14). Increased number of sampling sites and decreased extension of sampled area were associated with a low and high percentage of variance explained by environmental factors on fish community composition, respectively, and these effects were consistent across taxonomical and numerical resolutions, except order level (Fig. 3D; Table S15).

4. Discussion

4.1 Cost-effective alternatives for taxonomic and numeric resolutions

Overall, our results support that bioassessment studies and biomonitoring programs in Neotropical streams focused on monitoring the effects of riparian deforestation can be simplified by using different taxonomical and numerical resolutions, according to our first hypothesis. This was supported for both aquatic macroinvertebrates and fish, which are commonly used in bioassessment and biomonitoring approaches (Dala-Corte et al., 2020; Feio et al., 2023, 2021; Melo, 2005; Valente-Neto et al., 2021). Our second hypothesis was partially supported, as around one fifth and one tenth of the variation in the composition of macroinvertebrates and fish, respectively, was explained by environmental variables, and there was no increased explanation gain when tested finer taxonomical or numerical resolutions. Additionally, datasets characteristics explained the variability in correlation values between taxonomic and numerical resolutions, and in community-environment relationships, mainly for fish datasets, partially supporting our third hypothesis. We discuss below the implementation of cost-effective strategies in bioassessment and biomonitoring, considering the economic and personnel constraints, and the implications of the drivers of variation in correlation values.

In general, we found that taxonomic resolutions coarser than genus for aquatic macroinvertebrates and coarser than species for fish can be used with an acceptable loss of information to represent overall community patterns in the megadiverse Neotropical region. For macroinvertebrates, the family level can be used as a surrogate for the genus level due to the high correlation values between them (88%, on average) and similar community-environment relationship (10% and 11.2% on average for genus and family, respectively). Likewise, for fish, genus and species were highly correlated (85% on average) and showed similar associations with environmental variables (~20% for both species and genus). These results agree with previous studies assessing the congruence of distinct taxonomic resolutions for macroinvertebrates and fish based on local or regional data (e.g., Brito et al., 2018; Faquim et al., 2021; Martins et al., 2022; Ribas and Padial, 2015). In addition, a recent meta-analysis showed that coarser taxonomic levels, such as genus and family, yielded more responsive bioassessment tools when using community-level data (Ruaro et al., 2024). Employing coarser taxonomic resolution can significantly reduce time and costs when studying the entire community or an assemblage, as their determination is more readily accomplished than refined taxonomic resolution (Koch et al., 2021; Williams and Gaston, 1994) (Fig. 4). This optimization is particularly beneficial for biomonitoring programs and bioassessment initiatives aiming to detect, for instance, the impacts of riparian vegetation loss on community diversity. In highly biodiverse regions, such as the Neotropics, the use of higher taxonomical resolution could be important in some cases due to the knowledge gap regarding described species (referred to as the taxonomic gap, or Linnean shortfall) and the shortage of taxonomists (taxonomic impediment) (Raposo et al., 2021). Such optimization can enhance and expedite data acquisition, thereby aiding decision-making processes regarding conservation and restoration efforts in megadiverse tropical regions like Brazil, which face threats from intensive and rapid land use changes (Pelicice et al., 2017; Souza et al., 2020; Strassburg et al., 2017).

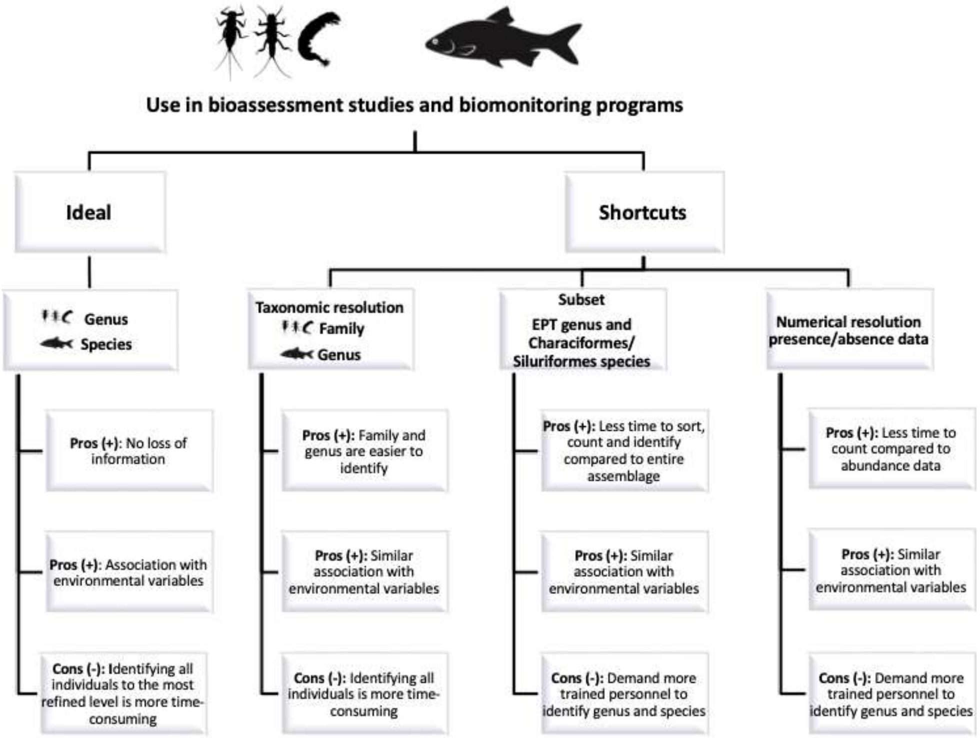


Fig. 4. Decision tree on the use of aquatic macroinvertebrates and fish in bioassessment studies and biomonitoring programs to detect the effects of riparian vegetation loss. The ideal scenario considers sufficient funds and taxonomic expertise to identify all organisms of each taxon to the most refined taxonomic resolution (macroinvertebrate genus and fish species). The simplifications that we recommend included taxonomic resolution, subsets of each taxonomic group, and numerical resolution. The adoption of one strategy would depend on the pros and cons of each simplification, which is mainly related to time and personnel trained to identify individuals.

We also found support for using presence-absence data for both aquatic macroinvertebrates and fish due to the high correlation between abundance and presence-absence data in the community-level matrices (83% and 87% on average for macroinvertebrates and fish, respectively) and the similar explanatory power of environmental variables on community composition (~12% for macroinvertebrates and ~18% for fish presence-absence on average). These results provide further support to previous studies in tropical regions that suggested the effectiveness of using presence-absence data for both groups in biomonitoring and bioassessment of aquatic systems (Brito et al., 2018; Faquim et al., 2021; Martins et al., 2022; Ribas and Padial, 2015). The similar explanatory power of environmental variables on abundance and presence-absence community composition differs from studies conducted in high latitude regions, where environmental variables have better explained abundance-weighed community composition than presence-absence data (Heino, 2014). This difference may be related to the higher species richness and spatial species turnover encountered in the tropics (Soininen et al., 2018). These processes contribute to a complex composition matrix, where species composition varies significantly between sites, reducing the explanatory power of environmental variables.

Presence-absence data are easier to process than abundance data, especially for abundant and diverse groups such as macroinvertebrates and fish. Presence-absence data only requires recording the presence of species, simplifying the counting process (Fig. 4). Although estimates of the monetary costs involved in collecting, counting, and identifying abundance data of aquatic macroinvertebrates and fish exist, this information is scarce for presence-absence data. For example, the average monitoring cost per genus of EPT (including collecting, sorting, and identification) is estimated in US$94, while for each fish species, it is estimated in US$48 for streams located in the Cerrado and Atlantic Forest transition in Brazil (Valente-Neto et al., 2021), and these costs should be lower to get presence-absence data. It is important to note that presence-absence data provide less detailed ecological information compared to abundance data, as they treat rare and common species equally. However, in terms of cost-effectiveness, presence-absence data require less processing time. The time and resources saved by using this simpler resolution can be redirected to increase sampling coverage in bioassessment studies and biomonitoring programs (Joseph et al., 2006). The use of presence-absence data can also reduce the need to collect unnecessarily large numbers of individuals in the field, particularly for fish, when species can be identified and counted on-site.

On average, the amount of variation in community composition explained by environmental variables was higher for fish than for macroinvertebrates, and this pattern remained consistent regardless of the taxonomic and numerical resolution. Some studies have shown that fish are more consistently related to environmental variables than aquatic macroinvertebrates (Herlihy et al., 2020; Picard et al., 2022), while others have indicated the inverse trend, including earlier detection of the effects of native riparian vegetation loss on macroinvertebrates (Dala-Corte et al., 2020; Valente-Neto et al., 2021). On average, environmental variables explained approximately 10% of the variation in macroinvertebrates composition and 19% in fish composition. This finding is not surprising, given that we used only landscape and climate variables as predictors due to the lack of comparable local environmental data across datasets. In general, the explanatory power of landscape and climate variables for aquatic fauna is relatively low (Heino et al., 2008; but see Junqueira et al., 2016; Montag et al., 2019; Roque et al., 2010). The varying responses of freshwater species to deforestation, as indicated by the percentage of native vegetation cover (Dala-Corte et al., 2020; Valente-Neto et al., 2021), can obscure the detection of clear community-environment relationships.

Local environmental variables, such as water dissolved oxygen and pH, are known to significantly influence aquatic macroinvertebrate biodiversity (Allan and Castillo, 2007; Tonkin et al., 2016) and may enhance the predictive power of environmental variables for macroinvertebrate and fish composition (e.g., Tonkin et al., 2016). However, some studies found that local environmental variables have limited explanatory power in predicting stream macroinvertebrates (Heino et al., 2015, 2008; Mykrä et al., 2008) and fish composition (Leal et al., 2018), even using global and standardized datasets (Heino et al., 2015). Additionally, studies incorporating climate, landscape and local environmental variables have still reported weak community-environment relationships (Heino et al., 2008; Siqueira et al., 2012b) or context dependency in biodiversity patterns, where the variation explained by environmental factors depends on the dataset analyzed (Tonkin et al., 2016). Beyond environmental selection, other processes such as ecological drift and dispersal also shape community patterns (Leibold and Chase, 2018). Some studies suggest that Neotropical stream communities are more strongly influenced by ecological drift compared to temperate streams due to their smaller community sizes (Saito et al., 2021; Siqueira et al., 2020). Community size, defined as the number of individuals sampled, affects the strength of ecological drift, which is more pronounced in smaller communities. This is because a lower number of individuals increases their susceptibility to random births and deaths (Orrock and Watling, 2010). Despite these various potential explanations for the unexplained variation in composition, it is important to note that we selected the most relevant environmental variables related to deforestation, specifically the percentage of native vegetation cover and tree cover percent, both of which are known to influence species occurrence, persistence community structure and forest resource availability (Fahrig, 2013, 2003). Thus, the inclusion of additional variables would likely not improve the composition variation related to deforestation. As a result, trends observed across taxonomic and numerical resolution in this study are expected to remain consistent, regardless of additional predictors or an increase in the overall community-environment relationship variation.

The mean correlation between the EPT subset and the entire macroinvertebrate community in the datasets was moderately high ($r = 0.72$), albeit lower than correlations reported in other studies, which found correlations greater than 0.90 (see Brito et al., 2018; Martins et al., 2022). Environmental variables explained a similar percentage of the variance in community composition when analyzing the EPT subset (mean R^2 across datasets = 9%) compared to the entire macroinvertebrate assemblage (mean R^2 across datasets = 10%). This result indicates that EPT are not more sensitive to deforestation, as represented by the percentage of native vegetation cover and tree cover percent, than other subsets of aquatic macroinvertebrates (Siqueira et al., 2012a). Therefore, while our findings support the widespread use of taxonomic EPT metrics (e.g., abundance, richness, and community composition) in studies assessing stream integrity (Couceiro et al., 2012; Suriano et al., 2011) (Fig. 4), other subsets of aquatic macroinvertebrates may be equally useful.

Similar to EPT and macroinvertebrate families, Characiformes fish species appear to be a good surrogate for entire stream fish assemblages, given their high correlation with fish species (mean correlation = 81%) and similar percentage of variance explained by environmental variables (17.5%) (Fig. 4). Our result are consistent with those recorded for Amazonian streams along a deforestation gradient (Martins et al., 2022). Moreover, families within the Characiformes (such as Lebiasinidae, Crenuchidae, and Characidae) have been identified as effective surrogates for entire stream fish assemblages in Amazon basins (Santos et al., 2022).

In general, Characiformes species dominate fish communities in Neotropical streams, displaying diverse habitat preferences and responses to environmental changes (Brejão et al., 2018). In addition, although Siluriformes were not strongly correlated with entire fish assemblage (mean correlation = 59%), environmental variables similarly accounted for variation in both Siluriformes and Characiformes (21.9% and 17.5%). One possible reason for this result is that Siluriformes encompass a great proportion of benthic species highly reliant on substrate heterogeneity, which may be adversely affected early by siltation of the stream bottom due to land use changes and riparian vegetation loss (Dala-Corte et al., 2016). Therefore, although Characiformes better represented the variance of the entire fish assemblage, Siluriformes may provide important insights into the impacts of riparian deforestation (Fig. 4).

4.2 What explains the variation across datasets?

Spatial extent and grain size have been reported to affect different types of congruence (cross-taxon, higher levels or subset) in terrestrial and marine ecosystems (Mellin et al., 2011; Westgate et al., 2014). Using predictors to elucidate the variation in congruence values is helpful in discerning scenarios where surrogates are more appropriate (Mellin et al., 2011). The variation in most congruence levels among macroinvertebrate datasets was not effectively explained by dataset characteristics, with exceptions noted for comparisons between genus and family, and between abundance and presence-absence. One possible explanation is the lower variation in correlation values for macroinvertebrates compared to fish (Table S2), which may hinder the detection of significant effect. On the other hand, the variation in congruence levels among fish datasets was effectively explained by dataset characteristics. Datasets of fish assemblages exhibiting high dominance, low rarity, and low species pool size may employ coarser taxonomic resolutions (e.g., replacing species by genus level) in a bioassessment and biomonitoring context with minimal information loss. These results suggest that the variation in congruence levels among macroinvertebrate datasets is less predictable than among fish datasets, at least for the database available and studied here.

The variation in community-environment relationships was high among datasets (around 4-fold for macroinvertebrates and 6-fold for fish) (Fig. 2). Understanding the factors that determine our capacity to explain freshwater communities is crucial for the effective use of bioindicators. For macroinvertebrates, the variation in community-environment relationships was poorly explained by datasets characteristics. Similar to congruence patterns, the lower variation in adjusted R^2 of community-environment relationships can explain this pattern (Table S6). For fish datasets, sample size and spatial extent was consistently found to affect our capacity to explain community composition using environmental variables, and this pattern was consistent across the different taxonomical and numerical resolutions evaluated. Increasing the spatial extent enhances the adjusted R^2 of community composition explained by environmental variables, aligning with findings from other studies (Leibold and Chase, 2018). On the other hand, increasing sample size may paradoxically decrease our ability to explain community variation through landscape and climate predictors in freshwater studies. Interestingly, environmental variability was not selected as important predictors of the community-environment relationship for fish. This suggest that increasing the number of sampling sites may add more data variation independent of the environmental variability. One possible explanation is that increasing the sample size may increase spatial turnover in species composition in highly diverse tropics (Soininen et al., 2018), leading to varied and contrasting responses among assemblages. However, further studies are needed to investigate this hypothesis. The database used here was designed to capture variation in riparian vegetation loss. Therefore, increasing sample size may not necessarily increase environmental variability. For example, if a dataset encompasses 20 streams spanning a continuous gradient of riparian vegetation loss (ranging from 0 to 100%, in 5% increments), adding more sites might not necessarily enhance the variation of riparian vegetation loss or the associated environmental variability. Nevertheless, expanding the number of sampling sites would undoubtedly enhance confidence in the obtained results, which is crucial for accurately assessing human impacts. Therefore, our results do not indicate that stream assessment and biomonitoring studies should decrease sample size to improve the community-environment relationship for fish. Instead, a rigorous study design remains crucial for an accurate environmental diagnosis.

Despite using a comprehensive database encompassing multiple biomes in a highly diverse Neotropical country, our study is subject to certain limitations, and potential caveats should be considered. Firstly, the available taxonomic knowledge and expertise in the Neotropical region limit our ability to identify aquatic macroinvertebrates beyond the genus level (Hamada et al., 2018). Consequently, we are unable to test the congruence between macroinvertebrate genus and species or determine whether genus can be used as surrogate for species. Therefore, our inference regarding the taxonomic resolution of macroinvertebrates is constrained to the genus level. Secondly, our study reveals several simplifications that can be employed to achieve the goals of bioassessment and biomonitoring. However, the implications and recommendations derived from our findings are limited to studies aimed at detecting and monitoring the effects of riparian vegetation loss on streams and using community-level data. The most refined taxonomic resolution of aquatic biodiversity is crucial in various contexts, including monitoring temporal trends in abundance and identifying biogeographical and macroecological patterns. Thirdly, our analyses incorporated both native and non-native fish species, thereby potentially influencing fish composition. While invasive species are known to impact riverine macroinvertebrate biomonitoring scores and indices that incorporate abundance weightings (Mathers et al., 2016), the influence of non-native species on fish congruence patterns remains unclear and needs to be better explored, considering the negative impacts in aquatic megadiverse ecoregions (Vitule and Pelicice, 2023). Even if we excluded non-native fish species from the analyses, the effects of their presence on native species would persist. Therefore, future studies should investigate this issue to provide further insights.

5. Conclusions

Ideally, with sufficient funds and taxonomic expertise, identifying organisms to the species level is the best choice for detecting and monitoring anthropogenic impacts, but this is rarely achievable (Jones, 2008). Therefore, our study provides clear and practical recommendations for stream assessment and biomonitoring the effects of riparian deforestation of Neotropical freshwater systems (Fig. 4). The congruences evaluated here provide good alternatives, suggesting that coarser taxonomical resolution and presence-absence data can serve as surrogates for refined taxonomical resolution and abundance data, respectively, for biomonitoring programs and bioassessment using community-level metrics of aquatic macroinvertebrates and fish. In addition, EPT and Characiformes or Siluriformes subsets may be good substitutes for macroinvertebrates and fish assemblages in some cases, but using these surrogates may result in some loss of information compared to using macroinvertebrate family or fish genera identification level. The predictors used in our analysis could only explain part of the variation in the congruence values for fish and macroinvertebrates. Our results indicate that, with a rigorous study design, both macroinvertebrates and fish were partially explained by environmental variables, but the magnitude of the effects of these variables on community composition was negatively affected by sample size and spatial extent for fish. Therefore, the choice of taxonomical resolution (macroinvertebrate family and fish genera), taxonomic subsets (e.g., EPT, Characiformes/Siluriformes) or numerical resolution (abundance or presence-absence) should depend on the budget and availability of trained personnel for sampling, sorting, and identifying individuals.

CRediT authorship contribution statement

Francisco Valente-Neto: Conceptualization, Writing – original draft, Methodology, Investigation. **Renato B. Dala-Corte:** Conceptualization, Writing – original draft, Methodology, Data curation, Formal Analysis, Investigation.

Almir M. Cunico: Writing – review & editing, Investigation. **André L. B. Magalhães**: Writing – review & editing, Investigation. **Bruno S. Godoy**: Writing – review & editing, Investigation. **Cecília G. Leal**: Writing – review & editing, Investigation**. Diego M. P. Castro**: Writing – review & editing, Investigation. **Diego R. Macedo**: Writing – review & editing, Investigation. **Dilermando P. Lima-Junior**: Writing – review & editing, Investigation**. Éder A. Gubiani**: Writing – review & editing, Investigation. **Fabio de O. Roque**: Writing – review & editing, Investigation**.** **Fabrício B. Teresa**: Writing – review & editing, Investigation. **Fagner Junior Machado de Oliveira**: Writing – review & editing, Investigation. **Fernando G.**

Becker: Writing – review & editing, Investigation. **Gabriel L. Brejão**: Writing – review & editing, Investigation. **Janaina G. de Brito**: Writing – review & editing, Investigation. **Jansen Zuanon**: Writing – review & editing, Investigation. **Jean R. S. Vitule**: Writing – review & editing, Investigation. **Karina Dias-Silva**: Writing – review & editing, Investigation. **Lilian Casatti**: Writing – review & editing, Investigation. **Luciano Benedito de Lima**: Writing – review & editing, Investigation. **Luciano F. A. Montag**: Writing – review & editing, Investigation. **Marcos Callisto**: Writing – review & editing, Investigation. **Mireile R. dos Santos**: Writing – review & editing, Investigation. **Neusa Hamada**: Writing – review & editing, Investigation. **Paulo A. Z. Pamplin**: Writing – review & editing, Investigation. **Paulo S. Pompeu**: Writing – review & editing, Investigation. **Rafael P. Leitão**: Writing – review & editing, Investigation. **Renata Ruaro**: Writing – review & editing, Investigation. **Sheyla R. M. Couceiro**: Writing – review & editing, Investigation. **Vinícius Abilhoa**: Writing – review & editing, Investigation. **Yzel R. Súarez**: Writing – review & editing, Investigation. **Renato T. Martins**: Conceptualization, Writing – original draft, Methodology, Investigation, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Myckey Gonçalves for producing the map showing the distribution of sampling points. This study was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Coordenação de

Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) and Amazonas State Research Foundation. This research was supported by National Institutes for Science and Technology (INCT) in Ecology, Evolution and Biodiversity Conservation (EECBio) funded by CNPq (grant 465610/2014-5), FAPEG (grant 201810267000023); and INCT ADAPTA II funded by CNPq (465540/2014–7) and FAPEAM (0621187/ 2017). FVN was supported by grant #2021/13299–0, São Paulo Research Foundation (FAPESP).). RTM received visiting researcher fellowships from CNPq (processes: 380592/2022–3 and 382557/2023-9). CGL is funded by the UKRI Future Leaders Fellowship (MR/W011085/1). DMPC is grateful to the National Council for Scientific and Technological Development (CNPq) for post-doctoral (PDJ) scholarship (process 151169/2022-5). DPLM was supported by grants #155509/2015 and #000073/2023, Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT). GLB was supported by FAPESP (processes: 12/21916-0 and 18/11954-9). DRM, FBT, JRSV, KDS, LC, LFAM, MC, NH, PSP and RPL are grateful to the National Council for Scientific and Technological Development (CNPq) for research productivity scholarship (processes: 311002-2023-4, 312844/2021-2, 310471/2023-0, 311550/2023-1, 304403/2021-0, 302881/2022–0, 304060/2020-8, 308970/2019–5, 302328/2022–0, 314464/2023-9, respectively). FBT was supported by National Council for Scientific and Technological Development (CNPq) (process 405706/2022-7). FGB was supported by Programa de Pesquisa em Biodiversidade (PPBio) and CNPq (process 457503/2012-2). GLB was supported by FAPESP (processes 12/21916-0 and 18/11954-9. JZ was supported by CNPq and FAPEAM for long-term support for research in the Amazon. LBL was supported by Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT/CNPq, Process

FAPEMAT-PRO 000073/2023) MC was supported by Companhia Energética de Minas Gerais (CEMIG) through Programa Cemig-Peixe Vivo, Programa de Pesquisa e Desenvolvimento Aneel-Cemig (GT-487, GT-550, GT-599, GT-611), and Cemig/Fapemig (APQ-01961-15, APQ-00261-22, and CRA 3147); Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG (APQ-01432-17 and APQ-00261-22). PAZP was supported by FAPEMIG (process #CRA-APQ-01467-11) and CAPES (process #152/2013). PSP was supported by Companhia Energética de Minas Gerais (CEMIG) through Programa Cemig-Peixe Vivo, Programa de Pesquisa e Desenvolvimento Aneel-Cemig andFundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG. RPL was supported by Companhia Energética de Minas Gerais - CEMIG (P&D Aneel/Cemig - GT599), Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG (APQ-01611-17). YRS was supported by Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT) and CNPq.

References

Albert, J.S., Destouni, G., Duke-Sylvester, S.M., Magurran, A.E., Oberdorff, T., Reis, R.E., Winemiller, K.O., Ripple, W.J., 2021. Scientists' warning to humanity on the freshwater biodiversity crisis. Ambio 50, 85–94. https://doi.org/10.1007/s13280-020-01318-8

Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257–284. https://doi.org/10.1146/annurev.ecolsys.35.120202.110122

Allan, J.D., Castillo, M.M., 2007. Stream ecology: structure and function of running waters. Springer Netherlands. https://doi.org/10.2307/2265800

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid Bioassessment Protocols For Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. Washington, D.C.

Barlow, J., França, F., Gardner, T.A., Hicks, C.C., Lennox, G.D., Berenguer, E., Castello, L., Economo, E.P., Ferreira, J., Guénard, B., Gontijo Leal, C., Isaac, V., Lees, A.C., Parr, C.L., Wilson, S.K., Young, P.J., Graham, N.A.J., 2018. The future of hyperdiverse tropical ecosystems. Nature 559, 517–526. https://doi.org/10.1038/s41586-018-0301-1

Bartoń, K., 2022. MuMIn: Multi-Model Inference. Package version 1.47.1.

Bonacina, L., Fasano, F., Mezzanotte, V., Fornaroli, R., 2023. Effects of water temperature on freshwater macroinvertebrates: a systematic review. Biol. Rev. 98, 191–221. https://doi.org/10.1111/brv.12903

Bonada, N., Prat, N., Resh, V.H., Statzner, B., 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. Annu. Rev.

Entomol. 51, 495–523. https://doi.org/10.1146/annurev.ento.51.110104.151124

Brejão, G.L., Hoeinghaus, D.J., Pérez-Mayorga, M.A., Ferraz, S.F.B., Casatti, L., 2018. Threshold responses of Amazonian stream fishes to timing and extent of deforestation. Conserv. Biol. 32, 860–871. https://doi.org/10.1111/cobi.13061

Brito, J.G., Martins, R.T., Oliveira, V.C., Hamada, N., Nessimian, J.L., Hughes, R.M., Ferraz, S.F.B., de Paula, F.R., 2018. Biological indicators of diversity in tropical streams: Congruence in the similarity of invertebrate assemblages. Ecol. Indic. 85, 85–92. https://doi.org/10.1016/j.ecolind.2017.09.001

Buss, D.F., Carlisle, D.M., Chon, T.-S., Culp, J., Harding, J.S., Keizer-Vlek, H.E., Robinson, W.A., Strachan, S., Thirion, C., Hughes, R.M., 2015. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. Environ. Monit. Assess. 187, 4132. https://doi.org/10.1007/s10661-014-4132-8

Carvalho, D.R. de, Leal, C.G., Junqueira, N.T., de Castro, M.A., Fagundes, D.C., Alves, C.B.M., Hughes, R.M., Pompeu, P.S., 2017. A fish-based multimetric index for Brazilian savanna streams. Ecol. Indic. 77, 386–396. https://doi.org/10.1016/j.ecolind.2017.02.032

Carvallo, F.R., Strickland, B.A., Kinard, S.K., Reese, B.K., Hogan, J.D., Patrick, C.J., 2022. Structure and functional composition of macroinvertebrate communities in coastal plain streams across a precipitation gradient. Freshw. Biol. 67, 1725–1738. https://doi.org/10.1111/fwb.13968

Couceiro, S.R.M., Hamada, N., Forsberg, B.R., Pimentel, T.P., Luz, S.L.B., 2012. A macroinvertebrate multimetric index to evaluate the biological condition of streams in the Central Amazon region of Brazil. Ecol. Indic. 18, 118–125. https://doi.org/10.1016/j.ecolind.2011.11.001

Dala-Corte, R.B., Giam, X., Olden, J.D., Becker, F.G., Guimarães, T. de F., Melo, A.S., 2016. Revealing the pathways by which agricultural land-use affects stream fish communities in South Brazilian grasslands. Freshw. Biol. 61, 1921–1934. https://doi.org/10.1111/fwb.12825

Dala-Corte, R.B., Melo, A.S., Siqueira, T., Bini, L.M., Martins, R.T., Cunico, A.M., Pes, A.M., Magalhães, A.L.B., Godoy, B.S., Leal, C.G., Monteiro-Júnior, C.S., Stenert, C., Castro, D.M.P., Macedo, D.R., Lima, D., Gubiani, É.A., Massariol, F.C., Teresa, F.B., Becker, F.G., Souza, F.N., Valente-Neto, F., de Souza, F.L., Salles, F.F., Brejão, G.L., Brito, J.G., Vitule, J.R.S., Simião-Ferreira, J., Dias-Silva, K., Albuquerque, L., Juen, L., Maltchik, L., Casatti, L., Montag, L., Rodrigues, M.E., Callisto, M., Nogueira, M.A.M., dos Santos, M.R., Hamada, N., Pamplin, P.A.Z., Pompeu, P.S., Leitão, R.P., Ruaro, R., Mariano, R., Couceiro, S.R.M., Abilhoa, V., Oliveira, V.C., Shimano, Y., Moretto, Y., Suarez, Y.R., de O. Roque, F., 2020. Thresholds of freshwater biodiversity in response to riparian vegetation loss in the Neotropical region. J. Appl. Ecol. 57, 1365-2664.13657. https://doi.org/10.1111/1365-2664.13657

Doledec, S., Statzner, B., 2010. Responses of freshwater biota to human disturbances : contribution of J-NABS to developments in ecological integrity assessments. J. North Am. Benthol. Soc. 29, 286–311. https://doi.org/10.1899/08-090.1

Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., Magurran, A.E., 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science (80-.). 344, 296–299. https://doi.org/10.1126/science.1248484

Fahrig, L., 2013. Rethinking patch size and isolation effects: the habitat amount hypothesis. J. Biogeogr. 40, 1649–1663. https://doi.org/10.1111/jbi.12130

Fahrig, L., 2003. Effects of Habitat Fragmentation on Biodiversity. Annu. Rev. Ecol. Evol. Syst. 34, 487–515. https://doi.org/10.1146/annurev.ecolsys.34.011802.132419

Faquim, R.C.P., Machado, K.B., Teresa, F.B., Oliveira, P.H.F. de, Granjeiro, G.F., Galli Vieira, L.C., Nabout, J.C., 2021. Shortcuts for biomonitoring programs of stream ecosystems: Evaluating the taxonomic, numeric, and cross-taxa congruence in phytoplankton, periphyton, zooplankton, and fish assemblages. PLoS One 16, e0258342. https://doi.org/10.1371/journal.pone.0258342

Feio, M.J., Hughes, R.M., Callisto, M., Nichols, S.J., Odume, O.N., Quintella, B.R., Kuemmerlen, M., Aguiar, F.C., Almeida, S.F.P., Alonso-EguíaLis, P., Arimoro, F.O., Dyer, F.J., Harding, J.S., Jang, S., Kaufmann, P.R., Lee, S., Li, J., Macedo, D.R., Mendes, A., Mercado-Silva, N., Monk, W., Nakamura, K., Ndiritu, G.G., Ogden, R., Peat, M., Reynoldson, T.B., Rios-Touma, B., Segurado, P., Yates, A.G., 2021. The biological assessment and rehabilitation of the world's rivers: an overview. Water 13, 371. https://doi.org/10.3390/w13030371

Feio, M.J., Hughes, R.M., Serra, S.R.Q., Nichols, S.J., Kefford, B.J., Lintermans, M., Robinson, W., Odume, O.N., Callisto, M., Macedo, D.R., Harding, J.S., Yates, A.G., Monk, W., Nakamura, K., Mori, T., Sueyoshi, M., Mercado-Silva, N., Chen, K., Baek, M.J., Bae, Y.J., Tachamo-Shah, R.D., Shah, D.N., Campbell, I., Moya, N., Arimoro, F.O., Keke, U.N., Martins, R.T., Alves, C.B.M., Pompeu, P.S., Sharma, S., 2023. Fish and macroinvertebrate assemblages reveal extensive degradation of the world's rivers. Glob. Chang. Biol. 29, 355–374. https://doi.org/10.1111/gcb.16439

Firmiano, K.R., Castro, D.M.P., Linares, M.S., Callisto, M., 2021. Functional responses of aquatic invertebrates to anthropogenic stressors in riparian zones of

Neotropical savanna streams. Sci. Total Environ. 753, 141865. https://doi.org/10.1016/j.scitotenv.2020.141865

Fox, J., Weisberg, S., 2019. An R Companion to Applied Regression, Third Edit. ed. Sage, Thousand Oaks CA.

Godoy, B.S., Faria, A.P.J., Juen, L., Sara, L., Oliveira, L.G., 2019. Taxonomic sufficiency and effects of environmental and spatial drivers on aquatic insect community. Ecol. Indic. 107. https://doi.org/10.1016/j.ecolind.2019.105624

Hamada, N., Thorp, J.H., Rogers, D.C., 2018. Keys to neotropical Hexapoda. In Thorp and Covich's Freshwater invertebrates, Fourth. ed. Academic Press.

Heino, J., 2014. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: Are the inferences transferable among regions? Ecol. Indic. 36, 186–194. https://doi.org/10.1016/j.ecolind.2013.07.022

Heino, J., 2010. Are indicator groups and cross-taxon congruence useful for predicting biodiversity in aquatic ecosystems? Ecol. Indic. 10, 112–117. https://doi.org/10.1016/j.ecolind.2009.04.013

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., Md Rawi, C.S., Rodrigues, M.E., Roque, F. de O., Sandin, L., Schmera, D., Sgarbi, L.F., Simaika, J.P., Siqueira, T., Thompson, R.M., 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. https://doi.org/10.1002/ece3.1439

Heino, J., Mykrä, H., Kotanen, J., 2008. Weak relationships between landscape characteristics and multiple facets of stream macroinvertebrate biodiversity in a boreal drainage basin. Landsc. Ecol. 23, 417–426. https://doi.org/10.1007/s10980-008-9199-6

Herlihy, A.T., Sifneos, J.C., Hughes, R.M., Peck, D. V., Mitchell, R.M., 2020. The relation of lotic fish and benthic macroinvertebrate condition indices to environmental factors across the conterminous USA. Ecol. Indic. 112, 105958. https://doi.org/10.1016/j.ecolind.2019.105958

Jacobsen, D., 2008. Tropical high-altitude streams, in: Dudgeon, D. (Ed.), Tropical Stream Ecology. Academic Press, San Diego, pp. 219–256.

Jähnig, S.C., Baranov, V., Altermatt, F., Cranston, P., Friedrichs-Manthey, M., Geist, J., He, F., Heino, J., Hering, D., Hölker, F., Jourdan, J., Kalinkat, G., Kiesel, J., Leese, F., Maasri, A., Monaghan, M.T., Schäfer, R.B., Tockner, K., Tonkin, J.D., Domisch, S., 2021. Revisiting global trends in freshwater insect biodiversity. WIREs Water 8, 1–5. https://doi.org/10.1002/wat2.1506

Jones, F.C., 2008. Taxonomic sufficiency: The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. Environ. Rev. 16, 45–69. https://doi.org/10.1139/A07-010

Joseph, L.N., Field, S.A., Wilcox, C., Possingham, H.P., 2006. Presence-absence versus abundance data for monitoring threatened species. Conserv. Biol. 20, 1679–1687. https://doi.org/10.1111/j.1523-1739.2006.00529.x

Junqueira, N.T., Macedo, D.R., de Souza, R.C.R., Hughes, R.M., Callisto, M., Pompeu, P.S., 2016. Influence of environmental variables on stream fish fauna at multiple spatial scales. Neotrop. Ichthyol. 14. https://doi.org/10.1590/1982-0224-20150116

Koch, E.B.A., Nascimento, I.C., Majer, J.D., Delabie, J.H.C., 2021. Seeking surrogates for taxonomic and functional evaluations of leaf-litter ant faunas. Ecol. Indic. 122, 107305. https://doi.org/10.1016/j.ecolind.2020.107305

Kopp, D.A., Allen, D.C., 2019. Stream network geometry and the spatial influence of aquatic insect subsidies across the contiguous United States. Ecosphere 10. https://doi.org/10.1002/ecs2.2926

Landeiro, V.L., Bini, L.M., Costa, F.R.C., Franklin, E., Nogueira, A., De Souza, J.L.P., Moraes, J., Magnusson, W.E., 2012. How far can we go in simplifying biomonitoring assessments? An integrated analysis of taxonomic surrogacy, taxonomic sufficiency and numerical resolution in a megadiverse region. Ecol. Indic. 23, 366–373. https://doi.org/10.1016/j.ecolind.2012.04.023

Leadley, P., Gonzalez, A., Obura, D., Krug, C.B., Londoño-Murcia, M.C., Millette, K.L., Radulovici, A., Rankovic, A., Shannon, L.J., Archer, E., Armah, F.A., Bax, N., Chaudhari, K., Costello, M.J., Dávalos, L.M., Roque, F. de O., DeClerck, F., Dee, L.E., Essl, F., Ferrier, S., Genovesi, P., Guariguata, M.R., Hashimoto, S., Ifejika Speranza, C., Isbell, F., Kok, M., Lavery, S.D., Leclère, D., Loyola, R., Lwasa, S., McGeoch, M., Mori, A.S., Nicholson, E., Ochoa, J.M., Öllerer, K., Polasky, S., Rondinini, C., Schroer, S., Selomane, O., Shen, X., Strassburg, B., Sumaila, U.R., Tittensor, D.P., Turak, E., Urbina, L., Vallejos, M., Vázquez-Domínguez, E., Verburg, P.H., Visconti, P., Woodley, S., Xu, J., 2022. Achieving global biodiversity goals by 2050 requires urgent and integrated actions. One Earth 5, 597–603. https://doi.org/10.1016/j.oneear.2022.05.009

Leal, C.G., Barlow, J., Gardner, T.A., Hughes, R.M., Leitão, R.P., Mac Nally, R., Kaufmann, P.R., Ferraz, S.F.B., Zuanon, J., de Paula, F.R., Ferreira, J., Thomson, J.R., Lennox, G.D., Dary, E.P., Röpke, C.P., Pompeu, P.S., 2018. Is environmental legislation conserving tropical stream faunas? A large-scale assessment of local, riparian and catchment-scale influences on Amazonian fish. J. Appl. Ecol. 55, 1312–1326. https://doi.org/10.1111/1365-2664.13028

Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordination of species data 2001, 271–280. https://doi.org/10.1007/s004420100716

Lehner, B., Verdin, K., Jarvis, A., 2008. New global hydrography derived from spaceborne elevation data. Eos (Washington. DC). 89, 93–94. https://doi.org/10.1029/2008EO100001

Leibold, M.A., Chase, J.M., 2018. Metacommunity ecology. Princeton University Press, Princeton.

Lopes, V.G., Castelo Branco, C.W., Kozlowsky-Suzuki, B., Bini, L.M., 2021. The reliability of low taxonomic and numerical resolutions for biodiversity monitoring is site specific and dependent on the statistical method. Ecol. Indic. 129, 107999. https://doi.org/10.1016/j.ecolind.2021.107999

Lynch, A.J., Cooke, S.J., Arthington, A.H., Baigun, C., Bossenbroek, L., Dickens, C., Harrison, I., Kimirei, I., Langhans, S.D., Murchie, K.J., Olden, J.D., Ormerod, S.J., Owuor, M., Raghavan, R., Samways, M.J., Schinegger, R., Sharma, S., Tachamo-Shah, R.D., Tickner, D., Tweddle, D., Young, N., Jähnig, S.C., 2023. People need freshwater biodiversity. WIREs Water 10, 1–31. https://doi.org/10.1002/wat2.1633

Maasri, A., Jähnig, S.C., Adamescu, M.C., Adrian, R., Baigun, C., Baird, D.J., Batista-Morales, A., Bonada, N., Brown, L.E., Cai, Q., Campos-Silva, J. V., Clausnitzer, V., Contreras-MacBeath, T., Cooke, S.J., Datry, T., Delacámara, G., De Meester, L., Dijkstra, K.B., Do, V.T., Domisch, S., Dudgeon, D., Erös, T., Freitag, H., Freyhof, J., Friedrich, J., Friedrichs-Manthey, M., Geist, J., Gessner, M.O., Goethals, P., Gollock, M., Gordon, C., Grossart, H., Gulemvuga, G., Gutiérrez-Fonseca, P.E., Haase, P., Hering, D., Hahn, H.J., Hawkins, C.P., He, F., Heino, J., Hermoso, V., Hogan, Z., Hölker, F., Jeschke, J.M., Jiang, M., Johnson, R.K., Kalinkat, G., Karimov, B.K., Kasangaki, A., Kimirei, I.A., Kohlmann, B., Kuemmerlen, M., Kuiper, J.J., Kupilas, B., Langhans, S.D., Lansdown, R., Leese, F., Magbanua, F.S., Matsuzaki, S.S., Monaghan, M.T., Mumladze, L., Muzon, J., Mvogo Ndongo, P.A., Nejstgaard, J.C., Nikitina, O., Ochs, C., Odume, O.N., Opperman, J.J., Patricio, H., Pauls, S.U., Raghavan, R., Ramírez, A., Rashni, B., Ross-Gillespie, V., Samways, M.J., Schäfer, R.B., Schmidt-Kloiber, A., Seehausen, O., Shah, D.N., Sharma, S., Soininen, J., Sommerwerk, N., Stockwell, J.D., Suhling, F., Tachamo Shah, R.D., Tharme, R.E., Thorp, J.H., Tickner, D., Tockner, K., Tonkin, J.D., Valle, M., Vitule, J., Volk, M., Wang, D., Wolter, C., Worischka, S., 2022. A global agenda for advancing freshwater biodiversity research. Ecol. Lett. 25, 255–263. https://doi.org/10.1111/ele.13931

Martins, R.T., Brito, J., Dias-Silva, K., Leal, C.G., Leitão, R.P., Oliveira, V.C., Oliveira-Júnior, J.M.B., de Paula, F.R., Roque, F.O., Hamada, N., Juen, L., Nessimian, J.L., Pompeu, P.S., Hughes, R.M., 2022. Congruence and responsiveness in the taxonomic compositions of Amazonian aquatic macroinvertebrate and fish assemblages. Hydrobiologia 849, 2281–2298. https://doi.org/10.1007/s10750-022-04867-z

Mathers, K.L., Chadd, R.P., Extence, C.A., Rice, S.P., Wood, P.J., 2016. The implications of an invasive species on the reliability of macroinvertebrate biomonitoring tools used in freshwater ecological assessments. Ecol. Indic. 63, 23–28. https://doi.org/10.1016/j.ecolind.2015.11.051

Mellin, C., Delean, S., Caley, J., Edgar, G., Meekan, M., Pitcher, R., Przeslawski, R., Williams, A., Bradshaw, C., 2011. Effectiveness of biological surrogates for predicting patterns of marine biodiversity: A global meta-analysis. PLoS One 6. https://doi.org/10.1371/journal.pone.0020141

Melo, A.S., 2005. Effects of taxonomic and numeric resolution on the ability to detect ecological patterns at a local scale using stream macroinvertebrates. Arch. für Hydrobiol. 164, 309–323. https://doi.org/10.1127/0003-9136/2005/0164-0309

Montag, L.F.A., Winemiller, K.O., Keppeler, F.W., Leão, H., Benone, N.L., Torres, N.R., Prudente, B.S., Begot, T.O., Bower, L.M., Saenz, D.E., Lopez-Delgado, E.O., Quintana, Y., Hoeinghaus, D.J., Juen, L., 2019. Land cover, riparian zones and instream habitat influence stream fish assemblages in the eastern Amazon. Ecol. Freshw. Fish 28, 317–329. https://doi.org/10.1111/eff.12455

Mykrä, H., Aroviita, J., Hämäläinen, H., Kotanen, J., Vuori, K.M., Muotka, T., 2008. Assessing stream condition using macroinvertebrates and macrophytes: Concordance of community responses to human impact. Fundam. Appl. Limnol. 172, 191–203. https://doi.org/10.1127/1863-9135/2008/0172-0191

Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A., Waterloo, M., Saleska, S., 2011. Height Above the Nearest Drainage - a hydrologically relevant new terrain model. J. Hydrol. 404, 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Oksanen, J., Simpson, G., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O’Haram, R., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., Barbour, M., M, B., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M., Lahti, L., McGlinn, D., Ouellette, M., E., R.C., Smith, T., Stier, A., Ter Braak, C., Weedon, J., 2022. Vegan: community ecology package. R package version 2.6-2.

Orrock, J.L., Watling, J.I., 2010. Local Community size mediates ecological drift and competition in metacommunities. Proc. R. Soc. B Biol. Sci. 277, 2185–2191. https://doi.org/10.1098/rspb.2009.2344

Pelicice, F.M., Azevedo-Santos, V.M., Vitule, J.R.S., Orsi, M.L., Lima Junior, D.P., Magalhães, A.L.B., Pompeu, P.S., Petrere, M., Agostinho, A.A., 2017. Neotropical freshwater fishes imperilled by unsustainable policies. Fish Fish. 18, 1119–1133. https://doi.org/10.1111/faf.12228

Picard, C., Floury, M., Seyedhashemi, H., Morel, M., Pella, H., Lamouroux, N., Buisson, L., Moatar, F., Maire, A., 2022. Direct habitat descriptors improve the understanding of the organization of fish and macroinvertebrate communities across a large catchment. PLoS One 17, 1–22. https://doi.org/10.1371/journal.pone.0274167

Podani, J., Schmera, D., 2011. A new conceptual and methodological framework for exploring and explaining pattern in presence - absence data. Oikos 120, 1625–1638. https://doi.org/10.1111/j.1600-0706.2011.19451.x

Raposo, M.A., Kirwan, G.M., Lourenço, A.C.C., Sobral, G., Bockmann, F.A., Stopiglia, R., 2021. On the notions of taxonomic 'impediment', 'gap', 'inflation' and 'anarchy', and their effects on the field of conservation. Syst. Biodivers. 19, 296–311. https://doi.org/10.1080/14772000.2020.1829157

Reid, A.J., Carlson, A.K., Creed, I.F., Eliason, E.J., Gell, P.A., Johnson, P.T.J., Kidd, K.A., MacCormack, T.J., Olden, J.D., Ormerod, S.J., Smol, J.P., Taylor, W.W., Tockner, K., Vermaire, J.C., Dudgeon, D., Cooke, S.J., 2019. Emerging threats and persistent conservation challenges for freshwater biodiversity. Biol. Rev. 94, 849–873. https://doi.org/10.1111/brv.12480

Reis, R.E., Albert, J.S., Di Dario, F., Mincarone, M.M., Petry, P., Rocha, L.A., 2016. Fish biodiversity and conservation in South America. J. Fish Biol. 89, 12–47. https://doi.org/10.1111/jfb.13016

Ribas, L.G. dos S., Padial, A.A., 2015. The use of coarser data is an effective strategy for biological assessments. Hydrobiologia 747, 83–95. https://doi.org/10.1007/s10750-014-2128-6

Roque, F.O., Corrêa, E.C., Valente-Neto, F., Stefan, G., Schulz, G., Barbosa Souza, P.R., Motta, C.M., Oliveira Bavutti, L.L., Colzani, E., Demétrio, M.F., Escarpinati, S.C., Silvestre, R., Vaz-de-Mello, F.Z., Siqueira, T., Ochoa Quintero, J.M., 2017. Idiosyncratic responses of aquatic and terrestrial insects to different levels of environmental integrity in riparian zones in a karst tropical dry forest region. Austral Entomol. 56, 459–465. https://doi.org/10.1111/aen.12259

Roque, F.O., Siqueira, T., Bini, L.M., Ribeiro, M.C., Tambosi, L.R., Ciocheti, G., Trivinho-Strixino, S., 2010. Untangling associations between chironomid taxa in Neotropical streams using local and landscape filters. Freshw. Biol. 55, 847–865. https://doi.org/10.1111/j.1365-2427.2009.02314.x

Ruaro, R., Gubiani, É.A., 2013. A scientometric assessment of 30 years of the Index of Biotic Integrity in aquatic ecosystems: applications and main flaws. Ecol. Indic. 29, 105–110. https://doi.org/10.1016/j.ecolind.2012.12.016

Ruaro, R., Gubiani, É.A., Padial, A.A., Karr, J.R., Hughes, R.M., Mormul, R.P., 2024. Responses of multimetric indices to disturbance are affected by index construction features. Environ. Rev. 32, 278–293. https://doi.org/10.1139/er-2023-0098

Saito, V.S., Stoppa, N.E., Shimabukuro, E.M., Cañedo-Argüelles, M., Bonada, N., Siqueira, T., 2021. Stochastic colonisation dynamics can be a major driver of temporal β diversity in Atlantic Forest coastal stream communities. Freshw. Biol. 66, 1560–1570. https://doi.org/10.1111/fwb.13738

Santos, L.L. dos, Benone, N.L., Brasil, L.S., Pires, T.H.S., Begot, T.O., Dantas, D.D.F., Montag, L.F. de A., 2022. The use of taxonomic families as biological surrogates of the diversity of the Amazonian stream fish. Ecol. Indic. 141, 109094. https://doi.org/10.1016/j.ecolind.2022.109094

Serra, S.R.Q., Graça, M.A.S., Dolédec, S., Feio, M.J., 2017. Chironomidae traits and life history strategies as indicators of anthropogenic disturbance. Environ. Monit. Assess. 189. https://doi.org/10.1007/s10661-017-6027-y

Siqueira, T., Bini, L.M., Roque, F.O., Cottenie, K., 2012a. A metacommunity framework for enhancing the effectiveness of biological monitoring strategies. PLoS One 7. https://doi.org/10.1371/journal.pone.0043626

Siqueira, T., Bini, L.M., Roque, F.O., Marques Couceiro, S.R., Trivinho-Strixino, S., Cottenie, K., 2012b. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. Ecography (Cop.). 35, 183–192. https://doi.org/10.1111/j.1600-0587.2011.06875.x

Siqueira, T., Saito, V.S., Bini, L.M., Melo, A.S., Petsch, D.K., Landeiro, V.L., Tolonen, K.T., Jyrkänkallio-Mikkola, J., Soininen, J., Heino, J., 2020. Community size can affect the signals of ecological drift and niche selection on biodiversity. Ecology 101, 1–10. https://doi.org/10.1002/ecy.3014

Soininen, J., Heino, J., Wang, J., 2018. A meta-analysis of nestedness and turnover components of beta diversity across organisms and ecosystems. Glob. Ecol. Biogeogr. 27, 96–109. https://doi.org/10.1111/geb.12660

Sousa, V., Dala-Corte, R.B., Benedito, E., Brejão, G.L., Carvalho, F.R., Casatti, L., Cetra, M., Pompeu, P.S., Súarez, Y.R., Tejerina-Garro, F.L., Borges, P.P., Teresa, F.B., 2023. Factors affecting the transferability of bioindicators based on stream fish assemblages. Sci. Total Environ. 881, 163417. https://doi.org/10.1016/j.scitotenv.2023.163417

Souza, C.M., Shimbo, J.Z., Rosa, M.R., Parente, L.L., Alencar, A.A., Rudorff, B.F.T., Hasenack, H., Matsumoto, M., Ferreira, L.G., Souza-Filho, P.W.M., de Oliveira, S.W., Rocha, W.F., Fonseca, A. V., Marques, C.B., Diniz, C.G., Costa, D., Monteiro, D., Rosa, E.R., Vélez-Martin, E., Weber, E.J., Lenti, F.E.B., Paternost, F.F., Pareyn, F.G.C., Siqueira, J. V., Viera, J.L., Neto, L.C.F., Saraiva, M.M., Sales, M.H., Salgado, M.P.G., Vasconcelos, R., Galano, S., Mesquita, V. V., Azevedo, T., 2020. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. Remote Sens. 12. https://doi.org/10.3390/RS12172735

Strassburg, B.B.N., Brooks, T., Feltran-Barbieri, R., Iribarrem, A., Crouzeilles, R., Loyola, R., Latawiec, A.E., Oliveira Filho, F.J.B., De Scaramuzza, C.A.M., Scarano, F.R., Soares-Filho, B., Balmford, A., 2017. Moment of truth for the Cerrado hotspot. Nat. Ecol. Evol. 1, 1–3. https://doi.org/10.1038/s41559-017-0099

Suriano, M.T., Fonseca-Gessner, A.A., Roque, F.O., Froehlich, C.G., 2011. Choice of macroinvertebrate metrics to evaluate stream conditions in Atlantic Forest, Brazil. Environ. Monit. Assess. 175, 87–101. https://doi.org/10.1007/s10661-010-1495-3

Tonkin, J.D., Heino, J., Sundermann, A., Haase, P., Jähnig, S.C., 2016. Context dependency in biodiversity patterns of central German stream metacommunities.

Freshw. Biol. 61, 607–620. https://doi.org/10.1111/fwb.12728

Valente-Neto, F., Martínez, B.T., Hughes, R.M., Ferreira, A., Severo-Neto, F., Souza, F.L. de, Souza, R.P., Escarpinati, S.C., Roque, F. de O., 2021. Incorporating costs, thresholds and spatial extents for selecting stream bioindicators in an ecotone between two Brazilian biodiversity hotspots. Ecol. Indic. 127. https://doi.org/10.1016/j.ecolind.2021.107761

Valente-Neto, F., Rodrigues, M.E., Roque, F. de O., 2018. Selecting indicators based on biodiversity surrogacy and environmental response in a riverine network: Bringing operationality to biomonitoring. Ecol. Indic. 94, 198–206. https://doi.org/10.1016/j.ecolind.2018.06.066

Viana, D.S., Chase, J.M., 2019. Spatial scale modulates the inference of metacommunity assembly processes. Ecology 100, 1–9. https://doi.org/10.1002/ecy.2576

Vitule, J.R.S., Pelicice, F.M., 2023. Care needed when evaluating the contributions of non-native species. Trends Ecol. Evol. 38, 499–500. https://doi.org/10.1016/j.tree.2023.01.005

Warwick, R.M., 1993. Environmental impact studies on marine communities: Pragmatical considerations. Aust. J. Ecol. 18, 63–80. https://doi.org/10.1111/j.1442-9993.1993.tb00435.x

Westgate, M.J., Barton, P.S., Lane, P.W., Lindenmayer, D.B., 2014. Global meta-analysis reveals low consistency of biodiversity congruence relationships. Nat. Commun. 5, 1–8. https://doi.org/10.1038/ncomms4899

Williams, P.H., Gaston, K.J., 1994. Measuring more of biodiversity: Can higher-taxon richness predict wholesale species richness? Biol. Conserv. 67, 211–217. https://doi.org/10.1016/0006-3207(94)90612-2

Table 1. Description and source of environmental variables used in the community-environmental relationship analysis (RDA, redundancy analysis models).

Variable	Description	Ecological effects in streams	Source
Altitude	Site altitude	Environmental (e.g., temperature and oxygen) and flow conditions (Jacobsen, 2008). Community structure and richness (Jacobsen, 2008).	SRTM (Shuttle Radar Topographic Mission)
Slope	Surface terrain slope measured as the ratio of the height difference between two points and their horizontal distance using altitude data	Environmental and flow conditions (e.g., current velocity; sediment deposition and substrate stability) (Allan and Castillo, 2007).	SRTM (Shuttle Radar Topographic Mission)
Density	Drainage network density, i.e., stream length per area	Runoff, connectivity between streams, aquatic insects distribution and aquatic subsidies (Kopp and Allen, 2019).	HydroSHEDS project (Lehner et al., 2008)
Hand50	Vertical distance of the water table calculated as the HAND (Height Above the Nearest Drainage), indicating the availability of groundwater	Flow conditions (flood, discharge and streamflow) (Nobre et al., 2011).	AMBDATA (https://www.dpi.inpe.br/Ambdata/download.php)
SA2001Tree	Tree cover percent, indicating forest canopy cover derived from MODIS (MODerate-resolution Imaging Spectroradiometer) for the sampling site at 1km^2 resolution	Organic matter input (e.g., leaf and wood), light penetration, stabilization of stream bed and banks (Allan, 2004). Community structure (Allan, 2004).	MODIS (MODerate-resolution Imaging Spectroradiometer) global data acquired between 2000 and 2001, at 1 km pixel size, also obtained from AMBDATA.

PercMA500m	Percentage of native vegetation cover within a 500-m radius buffer around sampled sites	Organic matter input (e.g., leaf and wood), filters pollutants, stabilizes banks, and maintains water quality. Community structure (Allan, 2004; Dala-Corte et al., 2020).	MapBiomas (http://mapbiomas.org), which was the scale at which macroinvertebrates and fish were most responsive in Dala-Corte et al. (2020b)
BIO1	Annual mean temperature	Physiology, metabolism, size and community structure (Bonacina et al., 2023).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO2	Mean diurnal temperature range	Large diurnal temperature fluctuations can stress aquatic organisms, influence respiration rates, and alter predator-prey interactions, affecting community stability (Bonacina et al., 2023).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO12	Annual precipitation	Drives periodic changes in streamflow, impacting life cycles, migration patterns, and resilience of aquatic communities to environmental fluctuations (Allan and Castillo, 2007; Carvallo et al., 2022).	WorldClim - Global Climate Data (https://www.worldclim.org/)
BIO15	Precipitation seasonality (coefficient of variation)	Drives periodic changes in streamflow, impacting life cycles, migration patterns, and resilience of aquatic communities to environmental fluctuations (Allan and Castillo, 2007; Carvallo et al., 2022).	WorldClim - Global Climate Data (https://www.worldclim.org/)

Table 2. Redundancy Analysis (RDA) of environmental variables explaining stream macroinvertebrate or fish communities identified at species, genera, family or order, all using abundance data, and presence-absence data (PA). Mean adjusted R^2 is the mean explanation obtained from 17 datasets for fish and 10 datasets for macroinvertebrates from Brazilian Amazon, Atlantic Forest, Cerrado, and Pampa biomes. Max = Maximum; Min = Minimum; EPT = Ephemeroptera, Plecoptera, and Trichoptera.

Response variables	Mean adj-R^2 (%)	Min adj-R^2	Max adj-R^2
Macroinvertebrate genera	10.0	1.1	25.5
Macroinvertebrate family	11.2	3.5	26.4
Macroinvertebrate order	11.0	0	27.2
Macroinvertebrate PA	11.6	3.0	27.0
EPT	9.0	0.7	20.9
Fish species	20.0	7.0	45.2
Fish genera	20.1	7.76	44.8
Fish family	17.1	6.3	42.6
Fish order	18.4	0	49.1
Fish PA	17.8	6.0	37.4
Characiformes	17.5	3.3	32.6
Siluriformes	21.9	7.6	51.6

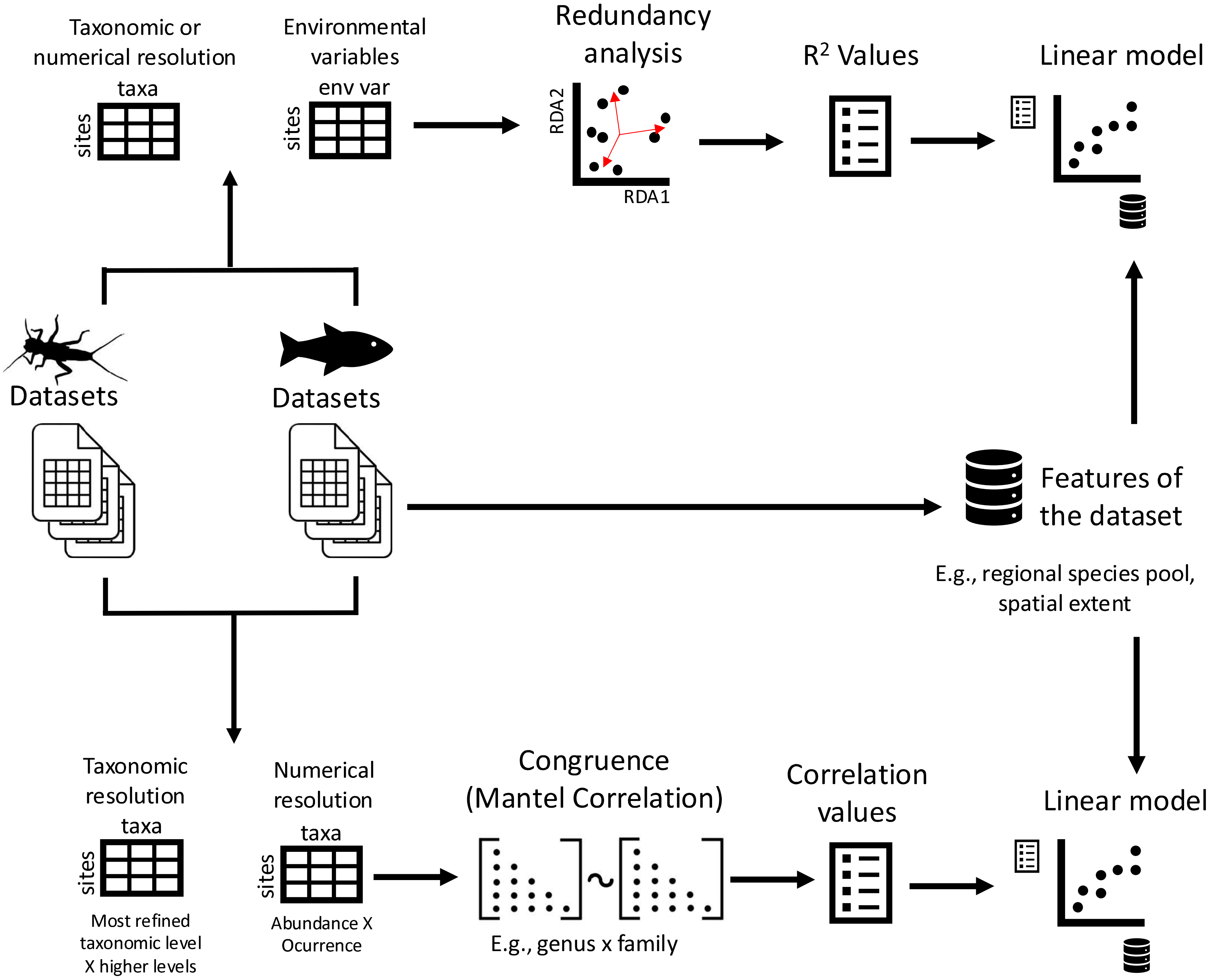

Taxonomic or numerical resolution
taxa
sites
Environmental variables
env var
sites
Redundancy analysis
RDA2
RDA1
R^2 Values
Linear model
Datasets
Datasets
Features of the dataset
E.g., regional species pool, spatial extent
Taxonomic resolution
taxa
sites
Most refined taxonomic level X higher levels
Numerical resolution
taxa
sites
Abundance X Ocurrence
Congruence (Mantel Correlation)
~
E.g., genus x family
Correlation values
Linear model

Figure 2

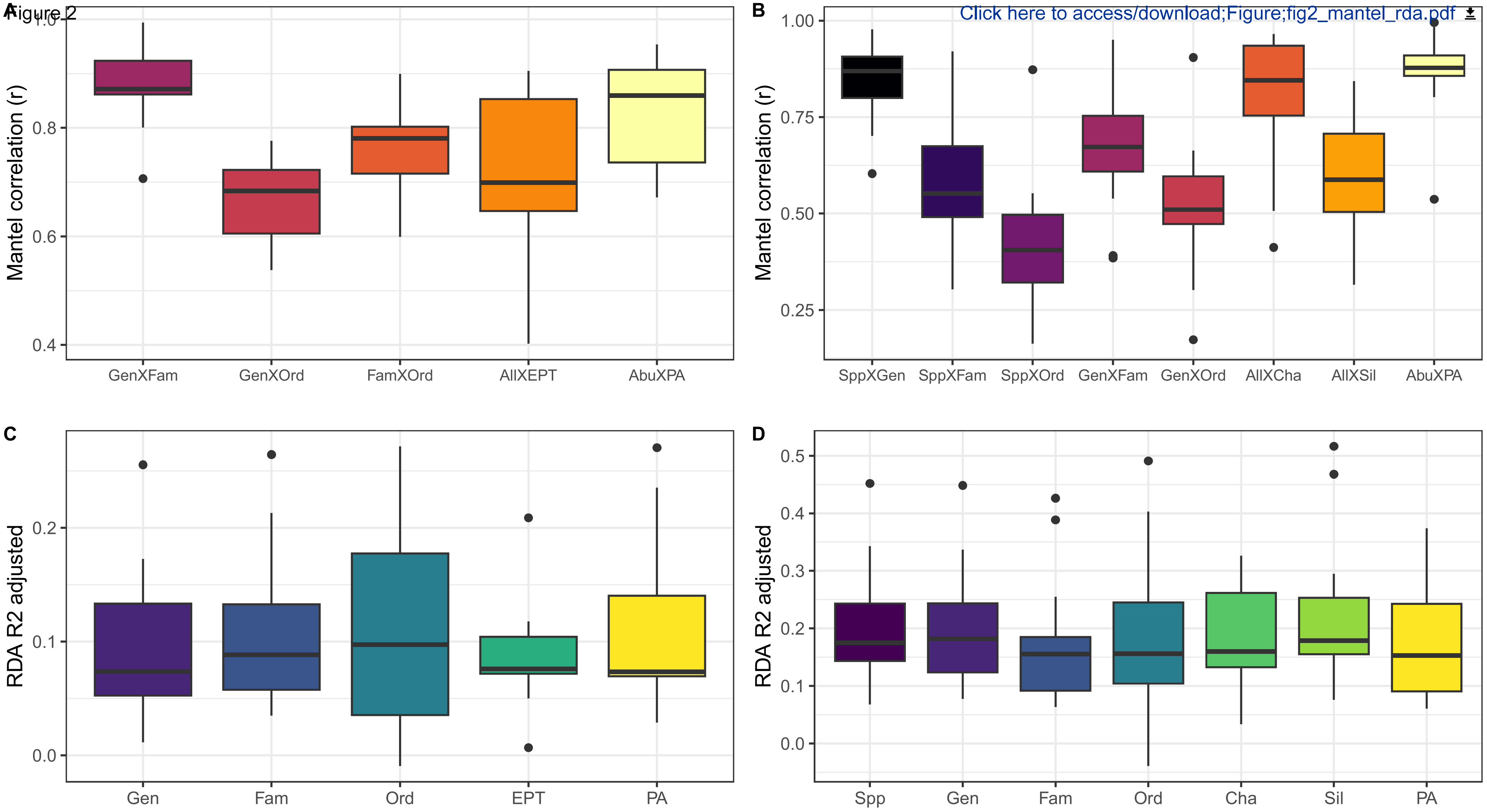

A
Mantel correlation (r)
1.0
0.8
0.6
0.4
GenXFam
GenXOrd
FamXOrd
AllXEPT
AbuXPA
B
Mantel correlation (r)
1.00
0.75
0.50
0.25
SppXGen
SppXFam
SppXOrd
GenXFam
GenXOrd
AllXCha
AllXSil
AbuXPA
C
RDA R2 adjusted
0.2
0.1
0.0
Gen
Fam
Ord
EPT
PA
D
RDA R2 adjusted
0.5
0.4
0.3
0.2
0.1
0.0
Spp
Gen
Fam
Ord
Cha
Sil
PA

Figure 3

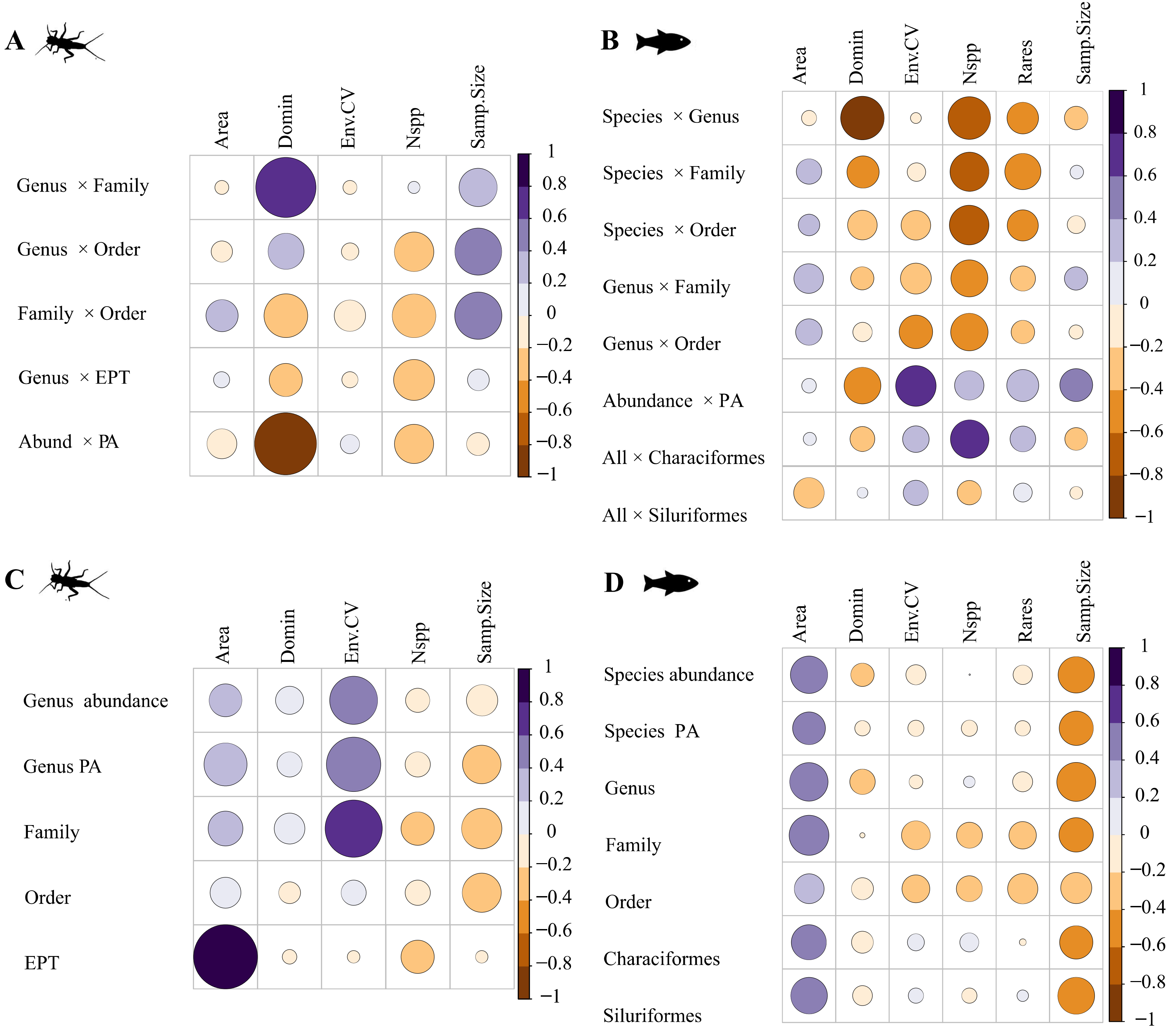
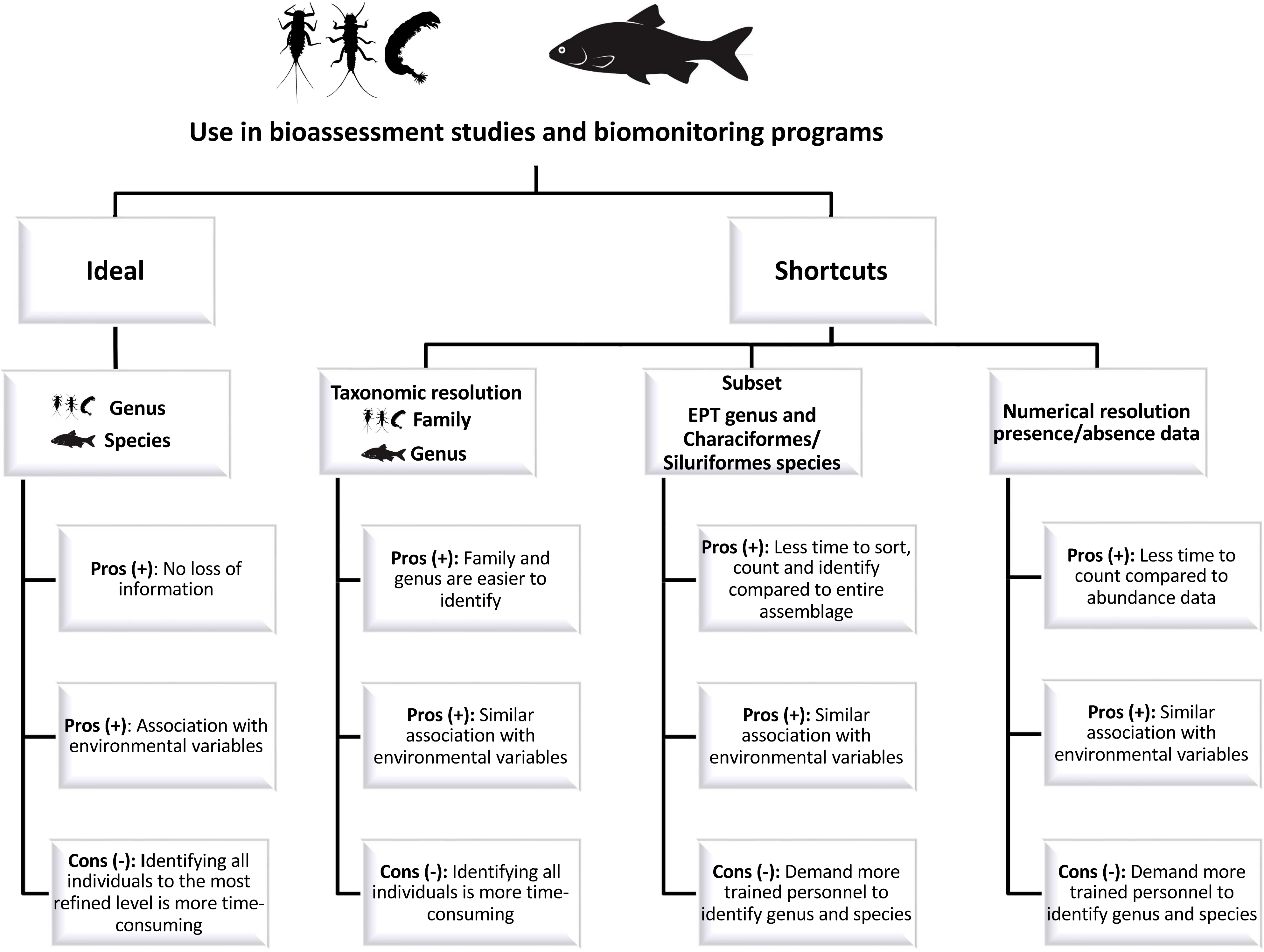

A
Area
Domin
Env.CV
Nspp
Samp.Size
Genus × Family
Genus × Order
Family × Order
Genus × EPT
Abund × PA
1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
B
Area
Domin
Env.CV
Nspp
Rares
Samp.Size
Species × Genus
Species × Family
Species × Order
Genus × Family
Genus × Order
Abundance × PA
All × Characiformes
All × Siluriformes
C
Genus abundance
Genus PA
Family
Order
EPT
D
Species abundance
Species PA
Genus
Family
Order
Characiformes
Siluriformes

Figure 4


Use in bioassessment studies and biomonitoring programs
Ideal
Shortcuts
Genus
Species
Taxonomic resolution
Family
Genus
Subset
EPT genus and Characiformes/ Siluriformes species
Numerical resolution presence/absence data
Pros (+): No loss of information
Pros (+): Association with environmental variables
Cons (-): Identifying all individuals to the most refined level is more time-consuming
Pros (+): Family and genus are easier to identify
Pros (+): Similar association with environmental variables
Cons (-): Identifying all individuals is more time-consuming
Pros (+): Less time to sort, count and identify compared to entire assemblage
Pros (+): Similar association with environmental variables
Cons (-): Demand more trained personnel to identify genus and species
Pros (+): Less time to count compared to abundance data
Pros (+): Similar association with environmental variables
Cons (-): Demand more trained personnel to identify genus and species

upplementary Material

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: