Environ Monit Assess (2016) 188:340
DOI 10.1007/s10661-016-5326-z

The role of physical habitat and sampling effort on estimates of benthic macroinvertebrate taxonomic richness at basin and site scales

Déborah R. O. Silva · Raphael Ligeiro · Robert M. Hughes · Marcos Callisto

Received: 2 December 2015 / Accepted: 25 April 2016
© Springer International Publishing Switzerland 2016

Abstract Taxonomic richness is one of the most important measures of biological diversity in ecological studies, including those with stream macroinvertebrates. However, it is impractical to measure the true richness of any site directly by sampling. Our objective was to evaluate the effect of sampling effort on estimates of macroinvertebrate family and Ephemeroptera, Plecoptera, and Trichoptera (EPT) genera richness at two scales: basin and stream site. In addition, we tried to determine which environmental factors at the site scale most influenced the amount of sampling effort needed. We sampled 39 sites in the Cerrado biome (neotropical savanna). In each site, we obtained 11 equidistant samples of the benthic assemblage and multiple physical habitat measurements. The observed basin-scale richness achieved a consistent estimation from Chao 1, Jack 1, and Jack 2 richness estimators. However, at the site scale, there was a constant increase in the observed number of taxa with increased number of samples. Models that best explained the slope of site-scale sampling curves (representing the necessity of greater sampling effort) included metrics that describe habitat heterogeneity, habitat structure, anthropogenic disturbance, and water quality, for both macroinvertebrate family and EPT genera richness. Our results demonstrate the importance of considering basin- and site-scale sampling effort in ecological surveys and that taxa accumulation curves and richness estimators are good tools for assessing sampling efficiency. The physical habitat explained a significant amount of the sampling effort needed. Therefore, future studies should explore the possible implications of physical habitat characteristics when developing sampling objectives, study designs, and calculating the needed sampling effort.

Keywords Sampling efficiency · Biodiversity · Richness estimators · Physical habitat

D. R. O. Silva (✉) · M. Callisto
Universidade Federal de Minas Gerais, Departamento de Biologia Geral, Instituto de Ciências Biológicas, Av. Antônio Carlos 6627, CP 486, CEP 30161-970 Belo Horizonte, Minas Gerais, Brazil
e-mail: deborah.ufmg@gmail.com

R. Ligeiro
Universidade Federal do Pará, Departamento de Biologia Geral, Instituto de Ciências Biológicas, Rua Augusto Corrêa, 01, CEP 66075-110 Belém, Pará, Brazil

R. M. Hughes
Amnis Opes Institute and Department of Fisheries and Wildlife, Oregon State University, Nash Hall, 97331-4501 Corvallis, OR, USA

Introduction

At the assemblage level, the most straightforward and frequently used measure of biodiversity is taxonomic richness, i.e., the number of taxa found in a given locality (Gotelli and Colwell 2001; Brown et al. 2007). Being an intuitive measure of assemblage diversity, taxonomic richness is well understood by researchers, managers, and the general public (Basualdo 2011). It is an important metric for composing biological indices (Klemm et al. 2003; Macedo et al. 2016), ecological

Published online: 10 May 2016

models (Arita and Vázquez-Domínguez 2008), macroecological and biogeographic studies (Oberdorff et al. 1995; Gaston 2000; Heino 2011), and for developing conservation strategies (Meir et al. 2004; Freemark et al. 2006). Patterns that determine biodiversity aspects are better studied in terrestrial than in aquatic ecosystems (Heino 2002). Rivers and streams exhibit a particular distribution pattern of biological diversity among taxonomic groups and among regions (Ligeiro et al. 2010; Hughes et al. 2012; McGarvey and Terra 2015). Those variations are associated with differences in physical habitat characteristics, water quality, colonization history, and frequency and magnitude of disturbances (Clarke et al. 2008; Ligeiro et al. 2010; McGarvey and Terra 2015).

Because the number of species in a given assemblage is limited, total species richness theoretically could be determined for any assemblage (Walther and Morand 1998). However, conducting a complete species inventory often is limited by overwhelming costs to obtain the data (Bartsch et al. 1998; Melo and Froehlich 2001) and there is also a lack of adequate taxonomic keys for most assemblages in many regions (Chao 2005; Hughes and Peck 2008). Especially for aquatic invertebrate assemblages, the measurement of taxonomic richness is still an obstacle in ecological studies (Allan and Flecker 1993). This situation is worse in tropical ecosystems than in temperate waters (Heino 2002).

The number of taxa generally increases asymptotically with increased number of samples or individuals counted, and species accumulation curves (or collection curves) are simple ways to assess how species richness varies with sampling effort (Cao et al. 2001; Gotelli and Colwell 2001; Hughes et al. 2012). When the curve stabilizes, i.e., when no significant increase in richness is observed with an increase in the sampling effort, almost all species will have been sampled. Thus, such curves allow estimates of the minimum number of samples or individuals required for adequate characterization of an assemblage (Bartsch et al. 1998; Bady et al. 2005). This dependency of species richness in relation to sampling effort has hindered proper comparisons among studies and identification of broad global patterns of species distributions. Stout and Vandermeer (1975), comparing streams in different latitudes, suggested that tropical streams harbor greater species richness of aquatic insects relative to temperate streams; however, because of small sample sizes, observed richness in the tropics is usually underestimated.

Species accumulation curves are directly influenced by the capture of rare species, because the greater the number of rare species, the steeper the observed accumulation curve (Melo and Froehlich 2001). The definition of rare species will depend mostly on the spatial and temporal scale of the study (Mao and Colwell 2005). Species that occur in low numbers or in low frequencies in sample units play an important role for detecting ecological changes (Leitão et al. 2016) and are critical for bioassesment studies (Cao et al. 1998). Because of the constant capture of rare taxa, it is rarely possible to observe stabilization in accumulation curves of benthic macroinvertebrates in tropical streams (Melo and Froehlich 2001; Schneck and Melo 2010). Given the impossibility of collecting all species from a site, richness estimators are used to fill the gaps left by incomplete inventories (Walther and Morand 1998). Among the ways to estimate richness, nonparametric estimators are widely employed (Gotelli and Colwell 2010). These estimators use the number of rare species that could not be detected consistently in multiple samples to try to predict the total number of species (Gotelli and Colwell 2010).

Understanding macroinvertebrate diversity is especially important for Cerrado (neotropical savanna) streams. Macroinvertebrates are considered a good ecological indicator, being sensitive to environmental changes in the physical, chemical, and biological components of aquatic ecosystems (Bonada et al. 2006). The Cerrado is the second largest neotropical biome (Wantzen 2003), originally covering about 20 % of Brazilian territory. It is a global terrestrial biodiversity hotspot (Myers et al. 2000) and is highly threatened by rapidly expanding anthropogenic activities (Wantzen et al. 2006). Different watershed land uses and cover can alter riparian zone structure and flow regimes, increase loadings of sediments, nutrients, and other pollutants, and reduce the availability of habitats for aquatic biota (Wang et al. 1997; Allan 2004). Because of its close proximity to aquatic organisms, the variety and variability of habitats have a strong influence on aquatic biodiversity (Allan 2004). Others have demonstrated how physical habitat complexity can explain stream macroinvertebrate assemblage composition and structure (Allan 2004; Ferreira et al. 2014; Macedo et al. 2014). However, less attention has been given to learning how those environmental characteristics can influence the amount of sampling effort necessary to characterize stream sites in an adequate manner.

Considering the importance of both within- and among-site heterogeneity for assessing macroinvertebrates in Cerrado streams (Ligeiro et al. 2010; Silva et al. 2014), the first objective of our study was to evaluate the sampling efficiency of taxonomic richness obtained in benthic macroinvertebrate samples at site and basin spatial scales. We were particularly interested in analyzing sampling effort in terms of the number of samples obtained (subsamples at the local scale and stream sites at the basin scale), in this way evaluating the patterns of areal richness, also known as richness density (sensu Hulbert 1971; Ligeiro et al. 2013a). This operational aspect of taxonomic richness is very often employed in ecological studies and corresponds simply to the number of species found in a certain area of habitat sampled, independently of the number of individuals counted. Richness density is related to numerous ecological phenomena including habitat heterogeneity, ecosystem productivity, isolation, climatic stability, and disturbance effects (Oberdorff et al. 1995; Gotelli and Colwell 2001; Ligeiro et al. 2013b). Our second objective was to determine which habitat characteristics are associated with increased site-scale sampling effort. Because species-level taxonomic keys are poorly developed for neotropical benthic macroinvertebrates, to answer both questions we assessed the effects at two taxonomic levels: macroinvertebrate family richness and Ephemeroptera, Plecoptera, and Trichoptera (EPT) genera richness.

Materials and methods

We conducted our study in wadeable Cerrado biome streams in the Upper São Francisco Basin (Minas Gerais State, southeastern Brazil) above Três Marias reservoir (Fig. 1). The basin climate is defined by a rainy season from October to March and a dry season from April to September, with mean annual precipitation

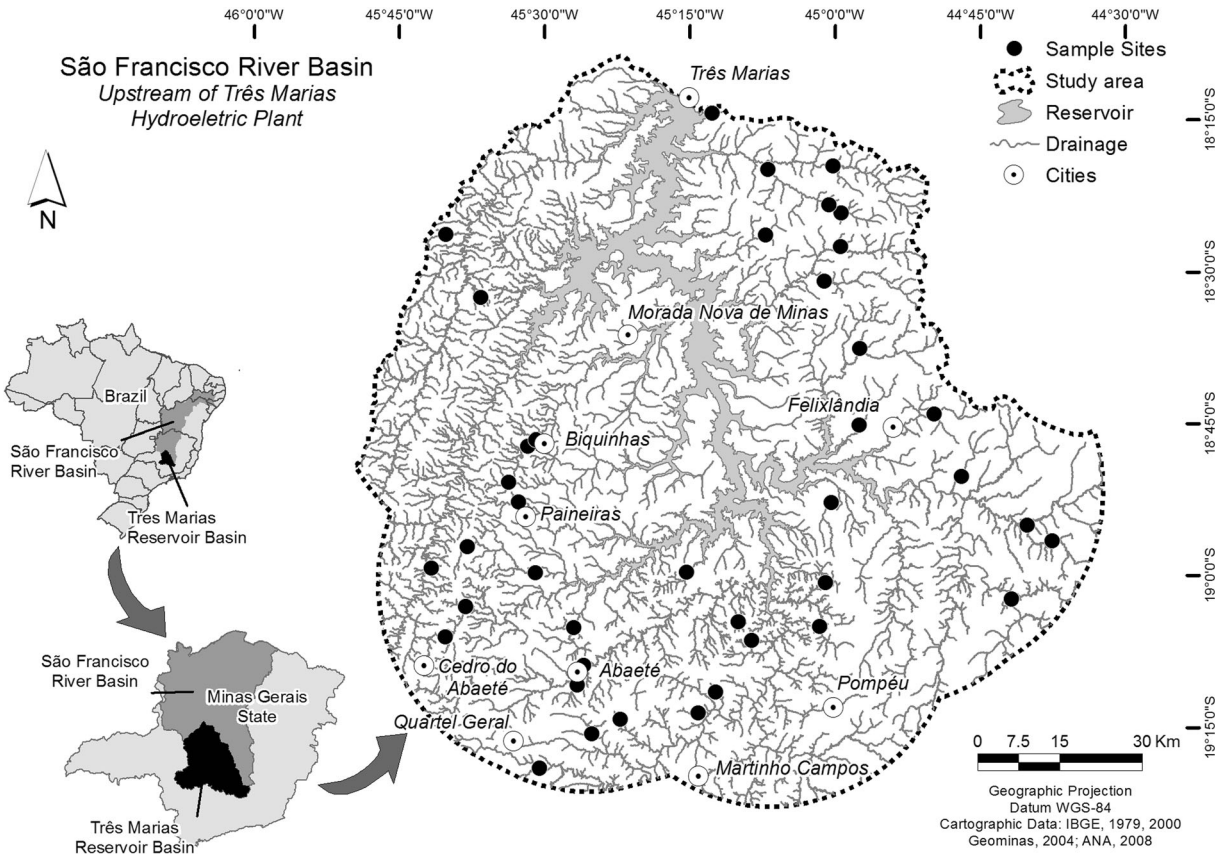

Fig. 1 Location of Três Marias Reservoir and stream sites sampled

ranging from 1200 to 1800 mm and mean annual temperature of 23 °C (Brasil 1992).

Field sampling was conducted in the end of the dry season (September 2010) once in each of 39 stream sites ranging from first to third order (1:100,000 scale maps) randomly selected by using the spatially balanced unequal probability survey design described by Olsen and Peck (2008). The length of each site was equal to 40 times the mean wetted width, with a minimum length of 150 m, to encompass the habitat complexity of the site, and 11 equidistant subsamples of benthic macroinvertebrates were collected following Peck et al. (2006). The subsamples were obtained systematically through a zigzag sampling pattern along the site, ensuring the sampling of multiple types of habitats and microhabitats. We used a D-net (30-cm mouth width, 0.09-m^2 area, 500-μm mesh) to collect macroinvertebrates. The samples were fixed with 10 % formalin and taken to the laboratory, where all macroinvertebrates were sorted and identified through use of a stereomicroscope (×100 magnification). Most invertebrates were identified to family level through use of identification keys (Pérez 1988; Merritt and Cummins 1996; Fernández and Domínguez 2001; Costa et al. 2006; Mugnai et al. 2010). Some noninsect invertebrates, like Hydracarina and Oligochaeta, were not identified to family. These groups accounted for very few of the total number of individuals collected (fewer than 5 %) and therefore were included in the family-level analyses. We identified EPT individuals to genus, allowing us to interpret results at two taxonomic levels: total macroinvertebrate families and EPT genera.

Also following Peck et al. (2006), we recorded many measurements of physical habitat characteristics in the sites, including stream channel morphology (e.g., slope, sinuosity, depth, wetted and bankfull width, incision, bank angle), habitat features (substrate size, flow types, presence of wood in the channel), riparian structure (canopy cover, vegetation type), and human alterations in the channel and riparian zones (e.g., presence of buildings, pasture, crops, roads, trash). Physical habitat metrics were then calculated according to Kaufmann et al. (1999). Some water quality characteristics were also measured once in each site: pH, electrical conductivity, total dissolved solids, dissolved oxygen, turbidity, total alkalinity, total nitrogen, and total phosphorus.

To estimate basin-scale sampling effort for the 39 sites, we evaluated the performance of three nonparametric richness estimators: the incidence-based first-order Jackknife and second-order Jackknife (hereafter Jack 1 and Jack 2, Burnham and Overton 1978) and the abundance-based Chao 1 (Chao 1984). The Jack 1 and Jack 2 richness estimators use counts of taxa that are present only in one (uniques) or two (duplicates) samples. The Chao 1 estimator bases its calculation on the number of taxa represented by just one or two individuals (singletons and doubletons). Those estimators have fundamentally different conceptual bases, but all three account for the presence of rare species in the samples. They were previously reported as producing good and reliable estimation of true taxa richness (Melo and Froehlich 2001; Martínez-Sanz et al. 2010; Basualdo 2011). We constructed curves of observed richness and estimated richness (Jack 1, Jack 2, and Chao 1) considering both total families and EPT genera richness using EstimateS 8.0 (Colwell 2006). One thousand randomizations were employed in the analyses.

To evaluate site-scale sampling effort, we constructed taxa accumulation curves for each of the 39 sites sampled, given the observed richness found in 1 to 11 subsamples. We represented the variation of the observed richness among the sites for each sampling effort by using boxplots. To assess the influence of physical habitat metrics on site-scale sampling effort, we first computed the observed richness at each subsample as a proportion of the total richness observed in each site, in this way avoiding biased slope calculations resulting from small or high richness values. Then, we obtained the logarithmic equation of the accumulation curve of each site:

$$y = a^{*}\ln(x) + b$$

where y is the expected richness, x is the sampling effort (number of subsamples collected), and b is a fitting constant. We used the slope of the curves (a value) to describe the sampling efficiency of the site. The greater the slope, the greater was the amount of sampling effort necessary to represent local richness. The sampling curves of all sites fitted well to the logarithmic equation (R^2 values > 0.95). We replicated this procedure for both total families and EPT genera datasets, also using EstimateS 8.0 (Colwell 2006), with 1000 randomizations in all cases.

We then performed multiple regression analysis (best-subset procedure) to determine the degree to which our predictor variables (metrics of disturbance, physical habitat structure, and water chemistry; Table 1)

Table 1 Candidate predictor metrics used in the best-subset multiple linear regression analysis

Metric categories and names	Code
Land use	
Percentage of agriculture	%agr
Percentage of pasture	%pas
Riparian human disturbance (proximity weighted index)	
Buildings	w1h_bldg
Pavement	w1h_pvmt
Trash and landfill	w1h_ldfl
Pasture	w1h_pstr
Logging	w1h_log
Channel morphology	
Mean (width/depth)	xwd_rat
Mean slope	xslope
Mean depth	xdepth
Mean bankfull width	xbkf_w
Mean width x mean depth	xwxd
Channel sinuosity (m/m)	sinu
Bankfull (width/depth)	bkf_wdrat
Mean water volume/m^2	v1w_msq
Bed substrate	
Substrate % cobbles (diameter 64–250 mm)	pct_cb
Percentage of coarse substrate (>16 mm)	pct_bigr
Percentage of fines (silt and clay)	pct_fn
Mean substrate embeddedness	xembed
Coarse litter (%)	pct_bf
Flow type	
Percentage of pools	pct_pool
Percentage of fast water	pct_fast
Shelter	
Natural fish cover in the stream (all)	xfc_nat
Brush and small debris fish cover (areal proportion)	xfc_brs
Undercut banks (areal proportion)	xfc_ucb
Anthropogenic fish cover (%)	pfc_ant
Riparian	
Riparian canopy (>5 m high) cover	xc
Riparian canopy (>5 m high) presence	xpcan
Total riparian cover (all vegetation layers)	xcmg
Mean mid-channel canopy density	xcdenmid
Water chemistry	
Negative log hydrogen ion concentration	pH
Material in suspension (NTU)	turb
Dissolved oxygen (mg/L)	do
Total nitrogen (mg/L)	ntotal

explained the variation in the slope values. In other words, we also evaluated the degree that habitat metrics influenced the amount of sampling effort required. We considered models generated with up to four explanatory variables, equivalent to 10 % of the number of sites analyzed, in this way avoiding model overfitting. Assumptions of homocedasticity and normality of errors were respected. We excluded correlated metrics that indicated high redundancy ($p \geq |0.8|$). Corrected Akaike information criteria (AICc) were used to select the best models. The analyses were performed in Systat 13 (Systat Software Inc 2009).

Results

We collected a total of 69,726 individuals, belonging to 80 taxa. Of those, 15,137 organisms were EPT, distributed in 65 genera. Family and EPT richness varied from 14 to 42 taxa and 1 to 36 genera per site, respectively.

At the basin scale, family and EPT patterns differed slightly. We observed that family-richness sampling effort achieved a consistent representation of the total richness because the curves of observed richness and all three estimators approached asymptotes (Fig. 2a). Regarding EPT genera richness, we observed a continuous increase in richness with increased number of sampled sites and proportionately less consistency of the estimators (Fig. 2b).

At the site scale, the average number of taxa observed increased continuously with increasing number of subsamples collected per site, both for total families and EPT genera (Fig. 3a, b). The wide range of richness values in each sampling effort reveals different richness patterns among sites. The variation in the slopes of the accumulation curves were moderately but significantly correlated with site-scale physical and water chemistry metrics ($R^2 = 0.369$ and 0.317, for macroinvertebrate families and EPT genera, respectively, Table 2). For the sampling effort of families, the significant metrics were dissolved oxygen, bankfull width/depth, mean natural fish cover (all natural instream habitat cover index), and mean bankfull width. Similar results were observed for EPT genera: dissolved oxygen, bankfull width/depth, mean brush and small debris fish cover (fraction of reach area covered by brush and small debris), and proximity of logging (Table 2).

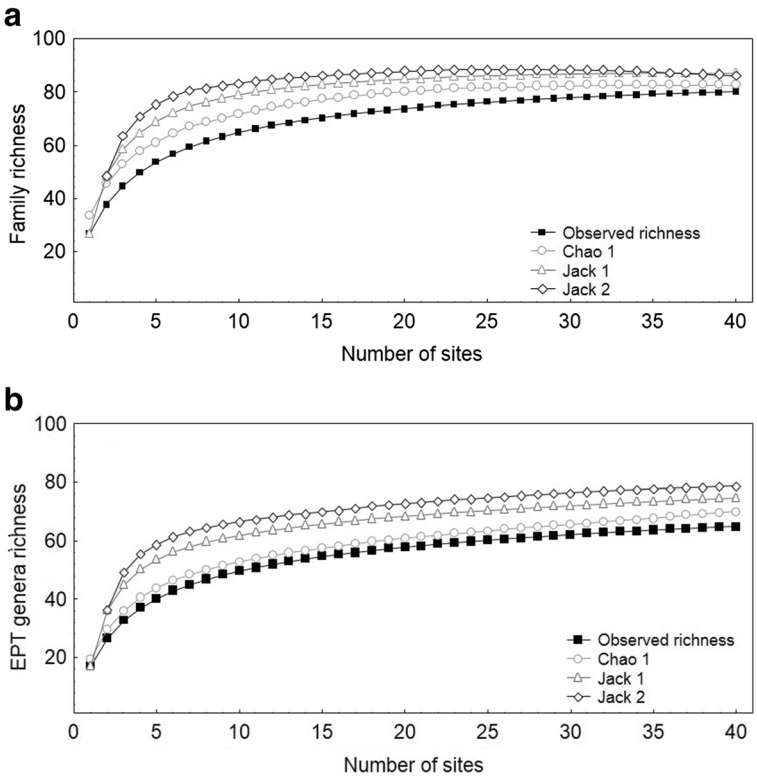

Fig. 2 Observed and estimated richness as a function of the number of sites sampled in the Upper São Francisco Basin: **a** macroinvertebrate families; **b** EPT genera

Discussion

Sampling sufficiency for estimates of taxonomic richness is an important issue in ecological studies because too little sampling greatly underestimates richness and excessive sampling is very expensive (Hughes and Peck 2008). In addition, sampling sufficiency may vary with the scale of the study (basin versus site) as well as with environmental factors influencing sampling effort.

The observed basin-scale richness curves approached stabilization and consistent estimations by Chao 1, Jack 1, and Jack 2. Although nonparametric methods of estimation are limited by their dependence on the number of rare taxa in an assemblage, they provide better estimates of true taxa richness than observed richness (Hughes et al. 2002; Melo 2004; dos Anjos and Zuanon 2007) allowing comparisons of different regions. This is particularly the case with tropical streams, which usually support many rare macroinvertebrate (Melo and Froehlich 2001; Oliveira et al. 2011).

Considering site-scale sampling effort, we did not observe curve stabilization; taxa richness increased continuously with increased numbers of subsamples for both taxonomic levels, reflecting that rare taxa continue to be added to the samples. Different stream habitats often exhibit different assemblage compositions, each one yielding its own rare species (Stout and Vandermeer 1975; Melo and Froehlich 2001). Therefore, our sampling design, collecting at many types of habitats in each site, may have contributed to the nonstabilization of the accumulation curves. Different results could have been obtained if we had focused on a single habitat type. However, multiple habitat sampling is preferable to targeting a single habitat type when the aim of the study is to obtain a comprehensive taxa list for assessing biodiversity and overall site condition (Gerth and Herlihy 2006; Hughes and Peck 2008; Buss et al. 2015). Other authors have observed the same results concerning site-scale sampling effort. Bady et al. (2005) found that ten subsamples represented only 50 % of the estimated richness in European stream sites. Li et al. (2001), studying headwater streams in Oregon (USA), also noted a constant increase in richness even after 50 subsamples were taken from each site. Sampling 11 subsamples at each of 20 random sites in seven separate Oregon and Washington rivers, Hughes et al. (2012) found that much greater sampling effort was required to estimate the true richness of riverine

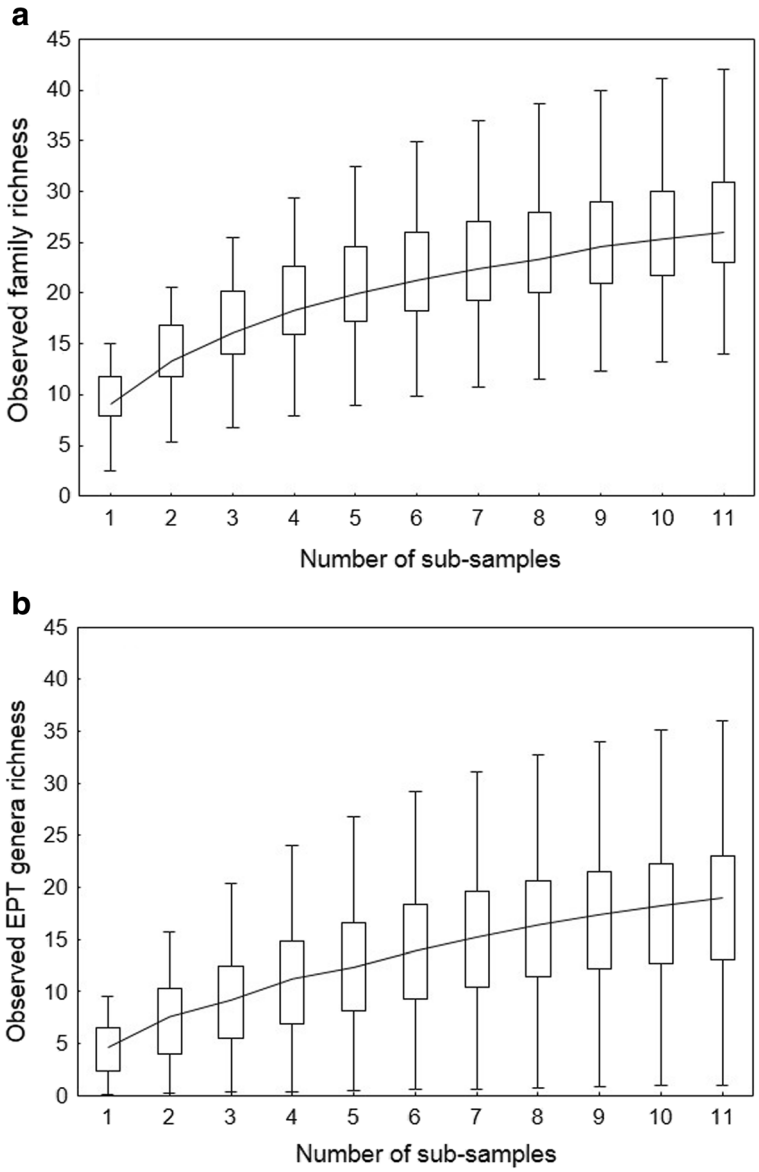

Fig. 3 Cumulative taxa richness as a function of the average number of subsamples for 39 sites in the Upper São Francisco Basin (*black line*). The *boxes* show the interquartile ranges; *whiskers* show the minimum and maximum values: **a** macroinvertebrate families; **b** EPT genera

benthic macroinvertebrate taxa. Li et al. (2014) reported that 20 subsamples were needed to reach asymptotes in sand-bottomed northeastern China streams. However, 11 subsamples per site are considered sufficient in biomonitoring studies for assessing stream ecological condition through use of multimetric indices in temperate streams (Peck et al. 2006; Hughes and Peck 2008). For the tropical region, this number is superior to the three subsamples commonly used in bioassesment studies (Moreno et al. 2010; Ferreira et al. 2011; Feio et al. 2013).

Observed taxonomic richness is highly biased toward the amount of sampling effort; however, fewer subsamples may be sufficient for estimating diversity via Simpson and Shannon-Wienner indices (Magurran 2004). In addition, fewer samples are needed to represent assemblage composition via similarity indices (Cao et al. 2001; Cao et al. 2002; Schneck and Melo 2010) and to represent ecological condition via multimetric indices (Stoddard et al. 2008; Oliveira et al. 2010; Oliveira et al. 2011; Chen et al. 2014) and via predictive models (Hawkins et al. 2000; Moya et al. 2011). Thus,

Table 2 Multiple linear regression models for habitat metrics explaining the slope of richness accumulation curves generated for macroinvertebrate families and EPT genera (*df* = 4,34)

	F value	*p* value	R^2	AIC		Beta	Std. error of beta	*p* value
					Intercept			<0.001
Families	4.97	0.003	0.369	−176.2	xfc_nat	0.337	0.148	0.030
					xbkf_w	−0.616	0.190	0.003
					bkf_wdrat	0.629	0.180	0.001
					do	0.302	0.157	0.063
					Intercept			<0.001
EPT genera	3.95	0.010	0.317	−156.9	w1h_log	−0.405	0.150	0.011
					bkf_wdrat	0.261	0.161	0.114
					xfc_brs	0.480	0.171	0.008
					do	0.339	0.162	0.044

See Table 1 for full metric names

adequate sample size will depend on the effect size of the study (Schneck and Melo 2010). Biomonitoring programs and rapid bioassessment can use small samples sizes to detect human impact, resulting also in reduced costs and laboratory time for processing the material (Lorenz et al. 2004; Vlek et al. 2006; Buss et al. 2015). However, for biology conservation studies, where a comprehensive taxonomic list is required, reduced sampling effort may increase the risk of mistakes in interpreting results (Petersen and Meier 2003; Hering et al. 2010).

Concerning taxonomic resolution, we observed greater stabilization in family taxonomic richness than in EPT genera richness at the basin scale (Fig. 2). As expected, the coarser the taxonomic resolution, the faster its stabilization relative to a finer taxonomic resolution (Gotelli and Colwell 2001). Despite this difference, the decision for a taxonomic level is usually determined by the study objectives (Jones 2008; Buss et al. 2015). Generally, the information provided by higher taxonomic levels, like family or even orders, are enough to identify environmental gradients in bioassesment and biomonitoring programs (Mao and Colwell 2005; Whittier and Van Sickle 2010; Buss and Vitorino 2010; Monk et al. 2012), and it is suggested when there are limitations in financial support, time for processing and knowledge for identification to lower levels (Buss and Vitorino 2010). However, refined taxonomic levels are preferred to distinguish subtle differences in water quality, when a high degree of confidence in conclusions is required, for conservation studies where a taxonomic list is benefited by precise identification (Lenat and Resh 2001) and when species function offers ecological insights that would be missing otherwise (Leitão et al. 2016).

Models that best explained the slopes of site-scale curves included metrics that describe habitat heterogeneity, habitat structure, anthropogenic disturbance, and water quality. Dissolved oxygen is well known as an important factor influencing the composition of freshwater communities (Dodds 2002; Connolly et al. 2004), but to our knowledge, no association with sampling effort had been done prior to our study. In our study, dissolved oxygen positively influenced the slopes of both family and EPT genera sampling curves. The greater the dissolved oxygen concentration, the greater the sampling effort needed for estimating assemblage richness. Mean bankfull width and bankfull width to depth ratio can influence the sampling effort in different ways. Those metrics describe stream size and previously have been associated with assemblage structure (Cole 2004). However, their effect on observed taxonomic richness is highly variable among studies (Clarke et al. 2008). Brush and small debris are important streambed components, providing habitat, food, and flow refuges for macroinvertebrates (Johnson and Kennedy 2003; Kaller and Kelso 2007). Those microhabitats can exhibit higher diversity and abundance of macroinvertebrates than occur in structurally simpler microhabitats (Drury and Kelso 2000). The abundance of brush and small debris influences taxonomic richness and consequently can influence sampling effort. Logging and natural

cover are commonly used metrics to explain observed taxa richness (Haggerty et al. 2004; Mereta et al. 2012) and probably influence sampling effort in the same way. Ferreira et al. (2014), analyzing the same dataset, found that a different set of habitat metrics explained site EPT richness compared with those we found explaining sampling effort. Their research and ours suggest that the factors driving the sampling effort needed are relatively independent from those explaining the total observed richness of the sites. Thus, our study provides an innovative method for evaluating the relationship between sampling effort and the environmental factors influencing it.

Sampling effort evaluations are of paramount importance for obtaining high-quality data that are capable of reliably describing diversity patterns and providing useful information for theoretical advancements and ecosystem management. We showed that it is very important to consider greater basin- and site-scale sampling effort than usually expended in stream ecological surveys if one is concerned with assessing taxonomic richness. In addition, taxa accumulation curves and richness estimators are useful for assessing sampling effort adequacy. Lastly, physical habitat characteristics affect the amount of sampling effort required for a site in a different manner than they do for explaining observed taxonomic richness. Therefore, we suggest that future studies continue to explore the possible implications of habitat characteristics before establishing sampling designs, in this way maximizing the representation of species diversity in different environmental contexts.

Acknowledgments We thank Companhia Energética de Minas Gerais (CEMIG), Projeto de Pesquisa e Desenvolvimento da Agência Nacional de Energia Elétrica (P&D ANEEL) (GT-487), and Programa Peixe-Vivo for research funding. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do estado de Minas Gerais (FAPEMIG), and Fulbright-Brasil provided financial support. MC was awarded research productivity CNPq (no. 303380/2015-2) and research project CNPq (no. 446155/2014-4), and Minas Gerais research grant FAPEMIG PPM-IX - 00525-15. Special thanks to colleagues from the Laboratório de Ecologia de Bentos-UFMG for sample collection and processing assistance.

References

Allan, J. D. (2004). Landscapes and riverscapes: the influence of land use on stream ecosystems. *Annual Review of Ecology, Evolution, and Systematics, 35*, 257–284.

Allan, J. D., & Flecker, A. S. (1993). Biodiversity conservation in running waters. *BioScience, 43*(1), 32–43.

Arita, H. T., & Vázquez-Domínguez, E. (2008). The tropics: cradle, museum or casino? A dynamic null model for latitudinal gradients of diversity. *Ecology Letters, 11*(7), 653–663.

Bady, P., Doledec, S., Fesl, C., Gayraud, S., Bacchi, M., & Scholl, F. (2005). Use of invertebrate traits for biomonitoring of European large rivers: the effects of sampling effort on genus richness and functional diversity. *Freshwater Biology, 50*(1), 159–173.

Bartsch, L. A., Richardson, W. B., & Naimo, T. J. (1998). Sampling benthic macroinvertebrates in a large flood-plain river: considerations of study design, sample size, and cost. *Environmental Monitoring and Assessment, 52*(3), 425–439.

Basualdo, C. V. (2011). Choosing the best non-parametric richness estimator for benthic macroinvertebrates databases. *Revista de la Sociedad Entomológica Argentina, 70*(1–2), 27–38.

Bonada, N., Prat, N., Resh, V. H., & Statzner, B. (2006). Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annual Review of Entomology, 51*, 495–523.

Brasil. (1992). Normais Climatológicas (1960–1990). Ministério da Agricultura e Reforma Agrária, Secretaria Nacional de Irrigação, Departamento Nacional de Meteorologia, Brasília, 84 p

Brown, R. L., Jacobs, L. A., & Peet, R. K. (2007). Species richness: small scale. In *Encyclopedia of Life Sciences* (pp. 1–8). John Wiley & Sons Ltd.

Burnham, K. P., & Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika, 65*(3), 625–633.

Buss, D. F., & Vitorino, A. S. (2010). Rapid bioassessment protocols using benthic macroinvertebrates in Brazil: evaluation of taxonomic sufficiency. *Journal of the North American Benthological Society, 29*(2), 562–571.

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. *Environmental Monitoring and Assessment, 187*(1), 4132.

Cao, Y., Larsen, D. P., & Hughes, R. M. (2001). Evaluating sampling sufficiency in fish assemblage surveys: a similarity-based approach. *Canadian Journal of Fisheries and Aquatic Sciences, 58*(9), 1782–1793.

Cao, Y., Larsen, D. P., Hughes, R. M., Angermeier, P. L., & Patton, T. M. (2002). Sampling effort affects multivariate comparisons of stream communities. *Journal of the North American Benthological Society, 21*(4), 701–714.

Cao, Y., Williams, D. D., & Williams, N. E. (1998). How important are rare species in aquatic community ecology bioassessment? *Limnology and Oceanography, 43*(7), 1403–1409.

Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics, 11*(4), 265–270.

Chao, A. (2005). Species estimation and applications. In N. Balakrishnan, C. B. Read, & B. Vidakovic (Eds.), *Encyclopedia of statistical sciences* (pp. 7909–7916). New York: Wiley.

Chen, K., Hughes, R. M., Xu, S., Zhang, J., Cai, D., & Wang, B. (2014). Evaluating performance of macroinvertebrate-based

adjusted and unadjusted multi-metric indices (MMI) using multi-season and multi-year samples. *Ecological Indicators, 36*, 142–151.

Clarke, A., Macnally, R., Bond, N. R., & Lake, P. S. (2008). Macroinvertebrate diversity in headwater streams: a review. *Freshwater Biology, 53*(9), 1707–1721.

Cole, M. B. (2004). *Assessment of macroinvertebrate communities in and adjacent to the city of Wilsonville, Oregon*. Oregon: Unpublished report prepared for the City of Wilsonville.

Colwell, R. K. (2006). EstimateS: statistical estimation of species richness and shared species from samples. Version 8. Persistent URL < purl.oclc.org/estimates>

Connolly, N. M., Crossland, M. R., & Pearson, R. G. (2004). Effect of low dissolved oxygen on survival, emergence, and drift of tropical stream macroinvertebrates. *Journal of the North American Benthological Society, 23*(2), 251–270.

Costa, C., Ide, S., & Simonka, C. E. (2006). *Insetos imaturos - metamorfose e identificação*. Ribeirão Preto: Holos Editora.

Dodds, W. K. (2002). *Freshwater ecology: concepts and environmental applications*. San Diego: Academic Press.

dos Anjos, M. B., & Zuanon, J. (2007). Sampling effort and fish species richness in small *terra firme* forest streams of central Amazonia, Brazil. *Neotropical Ichthyology, 5*(1), 45–52.

Drury, D. M., & Kelso, W. E. (2000). Invertebrate colonization of woody debris in coastal plain streams. *Hydrobiologia, 434*(1), 63–72.

Feio, M. J., Ferreira, W. R., Macedo, D. R., Eller, A. P., Alves, C. B. M., França, J. S., & Callisto, M. (2013). Defining and testing targets for the recovery of tropical streams based on macroinvertebrates communities and abiotic conditions. *River Research and Applications, 31*(1), 70–85.

Fernández, H. R., & Domínguez, E. (2001). *Guia para la determinación de los artrópodos bentónicos sudamericanos*. Tucumán: Universidad Nacional de Tucumán.

Ferreira, W. R., Ligeiro, R., Macedo, D. R., Hughes, R. M., Kaufmann, P. R., Oliveira, L. G., & Callisto, M. (2014). Importance of environmental factors for the richness and distribution of benthic macroinvertebrates in tropical headwater streams. *Freshwater Science, 33*(3), 860–871.

Ferreira, W. R., Paiva, L. T., & Callisto, M. (2011). Development of a benthic multimetric index for biomonitoring of a neotropical watershed. *Brazilian Journal of Biology, 71*(1), 15–25.

Freemark, K. E., Meyers, M., White, D., Warman, L. D., Kiester, A. R., & Lumban-Tobing, P. (2006). Species richness and biodiversity conservation priorities in British Columbia, Canada. *Canadian Journal of Zoology, 84*(1), 20–31.

Gaston, K. J. (2000). Global patterns in biodiversity. *Nature, 405*, 220–227.

Gerth, W. J., & Herlihy, A. T. (2006). Effect of sampling different habitat types in regional macroinvertebrate bioassessment surveys. *Journal of the North American Benthological Society, 25*(2), 501–512.

Gotelli, N. J., & Colwell, R. K. (2001). Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters, 4*(4), 379–391.

Gotelli, N. J., & Colwell, R. K. (2010). Estimating species richness. In A. E. Magurran & B. J. McGill (Eds.), *Biological diversity: frontiers in measurement and assessment* (pp. 39–54). Oxford: Oxford University Press.

Haggerty, S. M., Batzer, D. P., & Jackson, C. R. (2004). Macroinvertebrate response to logging in coastal headwater streams of Washington, USA. *Canadian Journal of Fisheries and Aquatic Sciences, 61*(4), 529–537.

Hawkins, C. P., Norris, R. H., Hogue, J. N., & Feminella, J. W. (2000). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications, 10*(5), 1456–1477.

Heino, J. (2002). Concordance of species richness patterns among multiple freshwater taxa: a regional perspective. *Biodiversity and Conservation, 11*(1), 137–147.

Heino, J. (2011). A macroecological perspective of diversity patterns in the freshwater realm. *Freshwater Biology, 56*(9), 1703–1722.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., Heiskanen, A.-S., Johnson, R. K., Moe, J., Pont, D., Solheim, A. L., & van de Bund, W. (2010). The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. *Science of the Total Environment, 408*(19), 4007–4019.

Hughes, R. M., & Peck, D. V. (2008). Acquiring data for large aquatic resource surveys: the art of compromise among science, logistics, and reality. *Journal of the North American Benthological Society, 27*(4), 837–859.

Hughes, R. M., Herlihy, A. T., Gerth, W. J., & Pan, Y. (2012). Estimating vertebrate, benthic macroinvertebrate and diatom taxa richness in raftable Pacific Northwest rivers for bioassessment purposes. *Environmental Monitoring and Assessment, 184*(5), 3185–3198.

Hughes, R. M., Kaufmann, P. R., Herlihy, A. T., Intelmann, S. S., Corbett, S. C., Arbogast, M. C., & Hjort, R. C. (2002). Electrofishing distance needed to estimate fish species richness in raftable Oregon rivers. *North American Journal of Fisheries Management, 22*(4), 1229–1240.

Hulbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology, 52*(4), 577–585.

Johnson, Z. B., & Kennedy, J. H. (2003). Macroinvertebrate assemblages of submerged woody debris in the Elm Fork of the Trinity River, Texas. *Journal of Freshwater Ecology, 18*(2), 187–197.

Jones, F. C. (2008). Taxonomic sufficiency: the influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environmental Reviews, 16*, 45–69.

Kaller, M. D., & Kelso, W. E. (2007). Association of macroinvertebrate assemblages with dissolved oxygen concentration and wood surface area in selected subtropical streams of the southeastern USA. *Aquatic Ecology, 41*(1), 95–110.

Kaufmann, P. R., Levine, P., Robison, E. G., Seeliger, C., & Peck, D. V. (1999). *Quantifying physical habitat in wadeable streams EPA/620/R-99/003*. Washington, DC: US Environmental Protection Agency.

Klemm, D. J., Blocksom, K. A., Fulk, F. A., Herlihy, A. T., Hughes, R. M., Kaufmann, P. R., Peck, D. V., Stoddard, J. L., Thoeny, W. T., Griffith, M. B., & Davis, W. S. (2003). Development and evaluation of a macroinvertebrate Biotic Integrity Index (MBII) for regionally assessing mid-atlantic highlands streams. *Environmental Management, 31*(5), 656–669.

Leitão, R. P., Zuanon, J., Villéger, S., Williams, S. E., Baraloto, C., Fortunel, C., Mendoça, F. P., & Mouillot, D. (2016). Rare species contribute disproportionately to the functional

structure of species assemblages. *Proceedings of the Royal Society B*. doi:10.1098/rspb.2016.0084.

Lenat, D. R., & Resh, V. H. (2001). Taxonomy and stream ecology—the benefits of genus- and species-level identifications. *Journal of the North American Benthological Society, 20*(2), 287–298.

Li, J., Herlihy, A., Gerth, W., Kaufmann, P., Gregory, S., Urquhart, S., & Larsen, D. P. (2001). Variability in stream macroinvertebrates at multiple spatial scales. *Freshwater Biology, 46*(1), 87–97.

Li, L., Liu, L., Hughes, R. M., Cao, Y., & Wang, X. (2014). Towards a protocol for stream macroinvertebrate sampling in China. *Environmental Monitoring and Assessment, 186*(1), 469–479.

Ligeiro, R., Ferreira, W., Hughes, R. M., & Callisto, M. (2013a). The problem of using fixed-area subsampling methods to estimate macroinvertebrate richness: a case study with Neotropical stream data. *Environmental Monitoring and Assessment, 185*(5), 4077–4085.

Ligeiro, R., Hughes, R. M., Kaufmann, P. R., Macedo, D. R., Firmiano, K. R., Ferreira, W. R., Oliveira, D., Melo, A. S., & Callisto, M. (2013b). Defining quantitative stream disturbance gradients and the additive role of habitat variation to explain macroinvertebrate taxa richness. *Ecological Indicators, 25*, 45–57.

Ligeiro, R., Melo, A. S., & Callisto, M. (2010). Spatial scale and the diversity of macroinvertebrates in a Neotropical catchment. *Freshwater Biology, 55*(2), 424–435.

Lorenz, A., Kirchner, L., & Hering, D. (2004). 'Electronic subsampling' of macrobenthic samples: how many individuals are needed for a valid assessment result? *Hydrobiologia, 516*(1), 299–312.

Macedo, D. R., Hughes, R. M., Ferreira, W. R., Firmiano, K. R., Silva, D. R., Ligeiro, R., Kaufmann, P. R., & Callisto, M. (2016). Development of a benthic macroinvertebrate multimetric index (MMI) for Neotropical Savanna headwater streams. *Ecological Indicators, 64*, 132–141.

Macedo, D. R., Hughes, R. M., Ligeiro, R., Ferreira, W. R., Castro, M. A., Junqueira, N. T., Oliveira, D. R., Firmiano, K. R., Kaufmann, P. R., Pompeu, P. S., & Callisto, M. (2014). The relative influence of catchment and site variables on fish and macroinvertebrate richness in cerrado biome streams. *Landscape Ecology, 29*(6), 1001–1016.

Magurran, A. E. (2004). *Measuring biological diversity*. Oxford: Blackwell Science Ltd.

Mao, C. X., & Colwell, R. K. (2005). Estimation of species richness: mixture models, the role of rare species, and inferential challenges. *Ecology, 86*(5), 1143–1153.

Martínez-Sanz, C., García-Criado, F., Fernández-Aláez, C., & Fernández-Aláez, M. (2010). Assessment of richness estimation methods on macroinvertebrate communities of mountain ponds in Castilla y León (Spain). *Annales de Limnologie - International Journal of Limnology, 46*(2), 101–110.

McGarvey, D. J., & Terra, B. F. (2015). Using river discharge to model and deconstruct the latitudinal diversity gradient for fishes of the Western Hemisphere. *Journal of Biogeography*. doi:10.1111/jbi.12618.

Meir, E., Andelman, S., & Possingham, H. P. (2004). Does conservation-planning matter in a dynamic and uncertain world? *Ecology Letters, 7*(8), 615–622.

Melo, A. S. (2004). A critique of the use of jackknife and related non-parametric techniques to estimate species richness. *Community Ecology, 5*(2), 149–157.

Melo, A. S., & Froehlich, C. G. (2001). Evaluation of methods for estimating macroinvertebrate species richness using individual stones in tropical streams. *Freshwater Biology, 46*(6), 711–721.

Mereta, S. T., Boets, P., Bayih, A. A., Malu, A., Ephrem, Z., Sisay, A., Endale, H., Yitbarek, M., Jemal, A., De Meester, L., & Goethals, P. L. M. (2012). Analysis of environmental factors determining the abundance and diversity of macroinvertebrate taxa in natural wetlands of Southwest Ethiopia. *Ecological Informatics, 7*(1), 52–61.

Merritt, R. W., & Cummins, K. W. (1996). *An introduction to the aquatic insects of North America*. Dubuque: Kendall/Hunt.

Monk, W. A., Wood, P. J., Hannah, D. M., Extence, C. A., Chadd, R. P., & Dunbar, M. J. (2012). How does macroinvertebrate taxonomic resolution influence ecohydrological relationships in riverine ecosystems. *Ecohydrology, 5*(1), 36–45.

Moreno, P., França, J. S., Ferreira, W. R., Paz, A. D., Monteiro, I. M., & Callisto, M. (2010). Factors determining the structure and distribution of benthic invertebrate assemblages in a tropical basin. *Neotropical Biology and Conservation, 5*(3), 135–145.

Moya, N., Hughes, R. M., Domínguez, E., Gibon, F. M., Goitia, E., & Oberdorff, T. (2011). Macroinvertebrate-based multimetric predictive models for measuring the biotic condition of Bolivian streams. *Ecological Indicators, 11*(3), 840–847.

Mugnai, R., Nessimian, J. L., & Baptista, D. F. (2010). *Manual de identificação de macroinvertebrados aquáticos do Estado do Rio de Janeiro*. Rio de Janeiro: Technical Books.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*, 853–858.

Oberdorff, T., Guegan, J. F., & Hugueny, B. (1995). Global scale patterns of fish species richness in rivers. *Ecography, 18*(4), 345–352.

Oliveira, R. B. S., Baptista, D. F., Mugnai, R., Castro, C. M., & Hughes, R. M. (2011). Towards rapid bioassessment of wadeable streams in Brazil: development of the Guapiaçu-Macau Multimetric Index (GMMI) based on benthic macroinvertebrates. *Ecological Indicators, 11*(6), 1584–1593.

Oliveira, R. B. S., Mugnai, R., Castro, C. M., & Baptista, D. F. (2010). Determining subsampling effort for the development of a rapid bioassessment using benthic macroinvertebrates in streams of Southeastern Brazil. *Environmental Monitoring and Assessment, 175*(1), 75–85.

Olsen, A. R., & Peck, D. V. (2008). Survey design and extent estimates for the Wadeable Streams Assessment. *Journal of the North American Benthological Society, 27*(4), 822–836.

Peck, D. V., Herlihy, A. T., Hill, B. H., Hughes, R. M., Kaufmann, P. R., Klemm, D. J., Lazorchak, J. M., McCormick, F. H., Peterson, S. A., Ringold, P. L., Magee, T., & Cappaert, M. R. (2006). *Environmental Monitoring and Assessment Program-Surface Waters: Western Pilot Study field operations manual for wadeable streams* (EPA/620/R-06/003). DC: USEPA.Washington.

Pérez, G. R. (1988). *Guía para el estudio de los macroinvertebrados acuáticos del Departamento de Antioquia*. Bogotá: Editorial Presencia Ltda.

Petersen, F. T., & Meier, R. (2003). Testing species-richness estimation methods on single-sample collection data using the Danish Diptera. *Biodiversity and Conservation, 12*(4), 667–686.

Schneck, F., & Melo, A. S. (2010). Reliable sample sizes for estimating similarity among macroinvertebrate assemblages in tropical streams. *Annales de Limnologie - International Journal of Limnology, 46*(2), 93–100.

Silva, D. R., Ligeiro, R., Hughes, R. M., & Callisto, M. (2014). Visually determined stream mesohabitats influence benthic macroinvertebrate assessments in headwater streams. *Environmental Monitoring and Assessment, 186*(9), 5479–5488.

Stoddard, J. L., Herlihy, A. T., Peck, D. V., Hughes, R. M., Whittier, T. R., & Tarquinio, E. (2008). A process for creating multimetric indices for large-scale aquatic surveys. *Journal of the North American Benthological Society, 27*(4), 878–891.

Stout, J., & Vandermeer, J. (1975). Comparison of species richness for stream-inhabiting insects in tropical and midlatitude streams. *The American Naturalist, 109*(967), 263–280.

Systat Software Inc. (2009). *SYSTAT statistical package, version 13.0*. San Jose, CA: Systat Software, Inc.

Vlek, H. E., Sporka, F., & Krno, I. (2006). Influence of macroinvertebrate sample size on bioassessment of streams. *Hydrobiologia, 566*(1), 523–542.

Walther, B. A., & Morand, S. (1998). Comparative performance of species richness estimation methods. *Parasitology, 116*(4), 395–405.

Wang, L., Lyons, J., Kanehl, P., & Gatti, R. (1997). Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. *Fisheries, 22*(6), 6–12.

Wantzen, K. M. (2003). Cerrado streams - characteristics of a threatened freshwater ecosystem type on the tertiary shields of South America. *Amazoniana, 17*(3–4), 485–502.

Wantzen, K. M., Siqueira, A., Nunes Da Cunha, C., & Sá, M. F. P. (2006). Stream-valley systems of the Brazilian cerrado: impact assessment and conservation scheme. *Aquatic Conservation: Marine and Freshwater Ecosystems, 16*(7), 713–732.

Whittier, T. R., & Van Sickle, J. (2010). Macroinvertebrate tolerance values and an assemblage tolerance index (ATI) for western USA streams and rivers. *Journal of the North American Benthological Society, 29*(3), 852–866.