

Should phylogenetic and functional diversity metrics compose macroinvertebrate multimetric indices for stream biomonitoring?

Victor S. Saito · Tadeu Siqueira ·
Alaíde Ap. Fonseca-Gessner

Received: 21 July 2014 / Revised: 29 October 2014 / Accepted: 30 October 2014 / Published online: 7 November 2014
© Springer International Publishing Switzerland 2014

Abstract The idea behind multimetric indices is to integrate information from several metrics to provide a general classification of water bodies without losing the particularities of each individual metric. Historically, multimetric indices use information on richness, taxon sensitivity, and taxonomic diversity. Recently, functional and phylogenetic diversity proved to capture different dimensions of biodiversity. Here we asked if these new metrics provide complementary information to classical metrics and should be included in multimetric indices. We used an index construction protocol based on statistical filters to test candidate metrics for range, sensitivity, and redundancy. We used macroinvertebrate data from streams

located in a Savanna region of Brazil, encompassing a gradient of impact, to test our ideas. Of 41 candidate metrics, functional dispersion of functional diversity, mean nearest neighbor distance of phylogenetic diversity, and four classical metrics passed the filter selection composing the final multimetric index. Our results indicated that functional and phylogenetic diversity metrics indeed responded to environmental impact and complemented the information provided by classical metrics. We suggest that future indices should consider including new metrics of functional and phylogenetic diversity to properly monitor multiple dimensions of biodiversity.

Keywords Taxonomic diversity · Environmental assessment · Trait diversity · Functional traits · Aquatic insects · Biotic index

Handling editor: Nuria Bonada

Electronic supplementary material The online version of this article (doi:[10.1007/s10750-014-2102-3](https://doi.org/10.1007/s10750-014-2102-3)) contains supplementary material, which is available to authorized users.

V. S. Saito
Programa de Pós-Graduação em Ecologia e Recursos Naturais, Centro de Ciências Biológicas e da Saúde, UFSCar, Rod. Washington Luís, km 235, São Carlos, SP 13.565-905, Brazil

V. S. Saito (✉)
Laboratório de Ecologia de Insetos Aquáticos,
Departamento de Hidrobiologia, UFSCar, Rod.
Washington Luís, km 235, São Carlos, SP 13.565-905,
Brazil
e-mail: victor.saito@gmail.com

T. Siqueira
Departamento de Ecologia, Instituto de Biociências,
UNESP – Universidade Estadual Paulista, Rio Claro, SP,
Brazil

A. Ap. Fonseca-Gessner
Departamento de Hidrobiologia Centro de Ciências
Biológicas e da Saúde, UFSCar, São Carlos, SP, Brazil

Introduction

The implementation of modern biomonitoring programs in Australia, Europe, and North America has been due to, in part, the increasing amount of supporting evidence from studies that used ecological and statistical theory to discuss biomonitoring practices in the past three decades (Bonada et al., 2006; Dolédec & Statzner, 2010). Among the most used approaches in biomonitoring, the multimetric index approach (Kerans & Karr, 1994; Barbour et al., 1996) is based on the idea of using metrics that represent major aspects of composition and diversity of biological communities that change with human impact. The benthic community of streams, rivers, and lakes—more specifically macroinvertebrates—are the focus of many well-established multimetric indices (Bonada et al., 2006). Among tropical countries, for example, although Brazil still lacks a formal nationwide biomonitoring program, there has been scientific advance with the proposition of multimetric indices (e.g., Ferreira et al., 2011; Oliveira et al., 2011; Suriano et al., 2011; Couceiro et al., 2012; Melo et al., 2014). The strength of this approach is in integrating information from a biological community to provide a general classification without losing the particularities of each individual metric (Barbour et al., 1996; Bonada et al., 2006; Hering et al., 2006). Metrics like the number of species (or genera), taxonomic, and trophic composition, and diversity indices are rated to what would be expected at a reference preserved site and combined into an index of biological integrity.

Recently, functional and phylogenetic diversity metrics have gained attention from ecologists because they can be used to identify aspects of biodiversity that are not captured by traditional indices based on taxonomic diversity (e.g., Petchey & Gaston, 2002; Cianciaruso et al., 2009; Vellend et al., 2011; Weiher, 2011). Traditional taxonomic diversity indices treat species as equivalent units—i.e., they all have the same importance for ecosystem services, the same ecological characteristics, and evolutionary history, and therefore, all have the same responses to environmental impacts (Magurran, 2004). In this sense, functional and phylogenetic approaches provide a means to consider ecological or evolutionary differences among species. Functional diversity indices measure the diversity of functional traits, i.e., characteristics of organism phenotype that influence

ecosystem services (Tilman, 2001), and organism survival (Weiher, 2011). Phylogenetic diversity indices, instead, measures the evolutionary relationship among species. The reasoning is that a community with species from different genera is more ecologically diverse than a community that has all species from the same genus. This is because close related species are likely to have similar phenotype and traits due to niche conservatism (Harvey & Pagel, 1991).

Within the multimetric index approach, functional-based metrics have been used mainly as the abundance of feeding groups (Cummins & Klug, 1979; Cummins et al., 2005), but not as functional diversity metrics per se (including recent studies, after the development of functional diversity as Couceiro et al., 2012). Regarding phylogenetic diversity, the study by Warwick & Clarke (1995) on taxonomic distinctiveness in marine environments was pioneering in comparing the phylogenetic structure of communities in impacted and preserved sites. This metric, however, was not used in a multimetric approach context. Moreover, Warwick & Clarke (1995) used taxonomic classification (e.g., order, family, genus, species) rather than branch lengths (which measures the evolutionary distance of species by molecular distance or divergence time) to calculate the distance between pairs of species. With recent advances in aquatic insect phylogenies, it is possible to construct phylogenetic supertrees (Sander-son et al., 1998) of metacommunities with branch lengths based on divergence time, creating, thus, a good opportunity to test them in multimetric protocols.

Functional and phylogenetic diversity are potentially useful for biomonitoring programs in a number of ways. Functional diversity metrics capture variation in trait states within a community. Since traits are phenotypic characteristics linked to organism survival (Weiher, 2011), we can expect that impacted sites restrict the life forms available for colonization, resulting in lower functional diversity when compared to pristine sites. Also, because functional diversity considers the richness of trait states of multiple traits, it is reasonable to consider it as more informative than simple metrics of functional groups.

Phylogenetic diversity metrics, on the other hand, indicate the amount of shared evolutionary history by the community (Faith, 1992; Vellend et al., 2011; Swenson, 2014). We could expect it to be higher in pristine environments because disturbed sites created

by anthropogenic impacts are recent in evolutionary scale. Thus, there would be few *taxa* that evolved traits to maintain populations there. Moreover, phylogenetic diversity indicates the adaptive potential of a community because more genetic diversity means greater capacity to respond to environmental changes (Vellend et al., 2011).

Based on these advantages, functional and phylogenetic diversity metrics have potential to be useful for aquatic biomonitoring programs (Vandewalle et al., 2010). In this sense, we asked if multimetric indices should include these new metrics to properly monitor several facets of biodiversity. Although there is information about many functional traits of aquatic macroinvertebrates from the Northern hemisphere (Dolédec et al., 1999; Poff et al., 2006; Dolédec & Statzner, 2010) and some from tropical environments (Tomanova et al., 2006; Colzani et al., 2013), we selected only four traits—voltinism, refuge building, respiration, and trophic groups (see Appendix 1 in Supplementary Material)—for two reasons. First, because these traits can be considered response traits—traits filtered by the environment that represent organism fitness in that environment (Violle et al., 2007). Second, because simplicity is one of the most desirable characteristic of biomonitoring tools (Bonada et al., 2006). We chose voltinism because we expected that impacted streams would have mainly multivoltine species—fast life cycle and less dependency on stable environments, whereas pristine streams would have both, multivoltine, and univoltine species—slower life cycle and higher dependency on stable environments to maintain populations (Weiher, 2011). We expected that impacted streams would have mainly species with aerial respiration because of low levels of dissolved oxygen in the water, whereas pristine streams would have species with cutaneous and gill respiration (Dolédec et al., 2006). Because impacted streams usually do not have riparian vegetation, they have lower levels of dissolved oxygen due to higher light entrance, higher daily variation in dissolved oxygen concentrations, and higher levels of organic matter (in case of presence of cattle) (Suriano et al., 2011). For the refuge building trait—i.e., trait states representing if species construct mobile shelters of sand, twigs, or leaves (mainly trichopterans and chironomids larvae)—and the trophic group trait, we reasoned that impacted streams would have species with no refuge building and mostly detritivorous

species, whereas pristine streams with preserved riparian vegetation would have all states of both traits (including *taxa* that construct shelters of leaves and twigs and leaf shredder species) (Cummins et al., 2005). Thus, because the results of functional diversity metrics should be a response of traits to the environment, we expected that anthropic impacts would diminish the values of functional diversity metrics.

We used a well-established protocol for multimetric index construction (Hering et al., 2006) to compare the responses of classical monitoring metrics with recent metrics of functional and phylogenetic diversity in face of environmental impacts of land use changes. Specifically, we compared the ability of traditional metrics and recent metrics in distinguishing between impacted areas and reference conditions of headwater streams, and also compared the redundancy among these different metrics. As we expected that functional and phylogenetic diversity metrics would access different dimensions of the communities, we predicted that they would be selected by the multimetric index protocol. This prediction is supported by the idea that functional diversity and phylogenetic diversity account for species differences (Vellend et al., 2011; Weiher, 2011), and therefore, they could indicate if impacts reduce the amount of life forms capable to inhabit these streams.

Materials and methods

Study area

We used a dataset collected in a region highly relevant for conservation, in São Paulo State (Brazil). The region is inserted in a Cerrado area (Brazilian Savanna), which is considered the most threatened tropical savanna in the world (Silva & Bates, 2002) and the only savanna biodiversity “hotspot” (Myers et al., 2000).

The Cerrado vegetation is composed of trees and C4 grasses that grow in acid soils rich in aluminum (Durigan et al., 2007). The climate is Köppen’s Cwa, with dry winters and hot wet summers (>22°C). The study region has less than 6% of its original area protected by law (Silva & Bates, 2002), and it is located on a plateau that favored the expansion of sugarcane monocultures and pastures.

We chose reference areas a priori by selecting 1st and 2nd order streams within minimally disturbed

Table 1 Measured variables in Cerrado streams of São Paulo State (Brazil)

Variables	Reference		Impacted		Statistic	
	Mean	SD	Mean	SD	<i>F</i>	<i>P</i>
Width (cm)	153.42	68.76	113.85	54.43	1.06	0.324
Depth (cm)	32.08	20.26	53.5	36.37	0	0.994
Velocity (m/s)	0.05	0.06	0.05	0.02	1.31	0.266
Temperature (°C)	17.29	2.2	20.03	3.31	12.42	0.003*
Conductivity (μS/cm)	5.8	3.57	45.42	51.22	21.77	<0.001*
pH	6.09	0.6	6.37	1.9	0.264	0.614
DO (mg/L)	6.37	0.7	7.4	1.33	4.011	0.06
RCE	257	9.7	30	23.93	125.1	<0.000*
Boulder (%)	16.71	30.47	7.07	22.71	1.39	0.253
Gravel (%)	6.71	11.33	2.42	3.05	0.65	0.428
Sand (%)	67.42	32.75	27.07	37.88	6.98	0.016*
Silt (%)	13.14	8.09	67.42	40.87	4.69	0.043*

SD standard deviation

* Significant difference

areas (Stoddard et al., 2006). The impacted streams were located in areas of sugarcane monoculture and pasture, which are the most common land uses in the region (Durigan et al., 2007). We based ours a priori selection on the degree of impact of streams according to an adaptation of the *Riparian, Channel, and Environmental protocol*—RCE (Petersen, 1992). This protocol evaluates land use, composition of the riparian vegetation and physical structure of the channel, and gives a final score of the environmental quality of the stream. We classified reference streams as “excellent” or “very good”, and impacted streams as “poor”. In total, we selected 21 streams that included seven for each type of land use (Cerrado vegetation, pasture, and sugarcane monoculture).

Sampling environmental and biological variables

In the dry season (May to August) of 2011, we measured the following physical and chemical variables in each stream: water temperature (°C), pH, conductivity (μS/cm), and dissolved oxygen (mg/L), superficial velocity (m/s) and average width (cm) and depth (cm) of the channel. We visually described the composition of the substrate by estimating the percentage of boulders, gravel, sand, and silt following Ward (1992).

The reference streams had significantly lower temperature [$F(1, 14) = 12.42, P = 0.003$] and electrical conductivity [$F(1, 14) = 21.77, P = 0.001$]. The substrate was composed mostly of sand and partly

of rocks and silt, which differs from impacted stream substrate that is mostly composed of silt (Table 1).

Following the protocol established by Suriano et al., (2011), we selected 100 meter-long stretches of stream and took six samples within each stretch by using a Surber sampler with an area of 30 × 30 cm (total of 0.54 m² per stream) and a mesh opening of 250 μm to sample macroinvertebrates. Additionally, for 30 s we scanned with a D-net to sample microhabitats where a Surber sampler was difficult to be used. We sorted the specimens live on transilluminated trays and identified them, at the laboratory, to the lowest feasible taxonomic level (mostly on genus level).

Metric selection

Classical metrics

We chose the classical metrics based on Barbour et al. (1996), Tetratich (2000), Bonada et al. (2006), Hering et al. (2006), Oliveira et al. (2008), Suriano et al. (2011), and Couceiro et al. (2012). The full list and the abbreviations for metric names can be found in Table 2.

We separated the 38 classical tested metrics into the following four groups: taxonomic diversity metrics; taxonomic composition metrics; tolerance metrics; and feeding group metrics (Table 2). We based the identification of feeding groups on the compilation by Cummins et al. (2005) for Brazilian groups, and the biotic indices we used were adaptations of the BMWP

Table 2 Tested metrics and the result of filters to range and sensitivity to anthropic impact

Metric		Range	Validation	Mann–Whitney <i>U</i>	<i>P</i>	Validation	Expected response
<i>Diversity measures</i>							
Family richness	S_f	8	Yes	4.26	0.030	Yes	Decrease
Genus richness	S_g	17	Yes		Ns	No	Decrease
Richness to genus (Chironomidae)	S_g-Ch	9	Yes	4.85	0.020	Yes	Decrease
EPT family richness	S_fEPT	5	Yes	11.62	<0.001	Yes	Decrease
EPT genus richness	S_gEPT	12	Yes	10.9	0.001	Yes	Decrease
EPTC family richness	S_fEPTC	6	Yes	9.70	0.001	Yes	Decrease
EPTC genus richness	S_gEPTC	10	Yes	9.44	0.002	Yes	Decrease
Ephemeroptera family richness	S_fE	1	No				Decrease
Plecoptera family richness	S_fP	2	No				Decrease
Trichoptera family richness	S_fT	3	No				Decrease
Coleoptera family richness	S_fC	1	No				Decrease
Ephemeroptera genus richness	S_gE	2	No				Decrease
Plecoptera genus richness	S_gP	2	No				Decrease
Trichoptera genus richness	S_gT	4	No				Decrease
Coleoptera genus richness	S_gC	3	No				Decrease
Chironomidae genus richness	S_Chiro	8	Yes		Ns	No	Decrease
Shannon diversity index (families)	H'_f	*		9.35	0.002	Yes	Decrease
Shannon diversity index (genus)	H'_g	*		5.69	0.010	Yes	Decrease
<i>Composition measures</i>							
Abundance	Abund	*			Ns	No	Increase
Percentage of EPT	%_EPT	44.05	Yes	6.82	0.009	Yes	Decrease
Percentage of EPTC	%_EPTC	48.91	Yes	6.43	0.011	Yes	Decrease
Percentage of Ephemeroptera	%_E	11.32	Yes		Ns	No	Decrease
Percentage of Plecoptera	%_P	2.3	No			No	Decrease
Percentage of Trichoptera	%_T	15.05	Yes	6.90	0.008	Yes	Decrease
Percentage of Coleoptera	%_C	4.28	No			No	Decrease
Percentage of Chironomidae	%_Chiro	41.73	Yes		Ns	No	Increase
<i>Tolerance measures</i>							
EPT/Chironomidae	EPT/Chiro	*		4.68	0.035	Yes	Decrease
EPTC/Chironomidae	EPTC/Chiro	*		4.36	0.034	Yes	Decrease
Baetidae/Ephemeroptera	Bae/Eph	*			Ns	No	Increase
BMWP_1	BMWP_1	*		6.63	0.013	Yes	Decrease
BMWP_2	BMWP_2	*		8.25	0.004	Yes	Decrease
<i>Feeding group measures</i>							
Percentage of collector	%_col	39.5	Yes		Ns	No	Increase
Percentage of filter	%_fil	17.52	Yes		Ns	No	Increase
Percentage of scraper	%_scra	11.75	Yes		Ns	No	Increase
Percentage of predator	%_pre	18.4	Yes		Ns	No	Decrease
Percentage of shredder	%_shre	1.6	No			No	Decrease
[Scra + col – filt]/[shre + col – gat]	SF/SC	*			Ns	No	Increase
Scra/(shre + col)	S/SC	*			Ns	No	Increase
<i>Phylogenetic diversity measures</i>							
Phylogenetic diversity	PD	*			Ns	No	Decrease

Table 2 continued

Metric		Range	Validation	Mann–Whitney <i>U</i>	<i>P</i>	Validation	Expected response
Mean pairwise distance	MPD	*			Ns	No	Decrease
Mean nearest neighbour distance	MNND	*		8.30	0.009	Yes	Decrease
<i>Functional diversity measures</i>							
Functional diversity	FD	*			Ns	No	Decrease
Functional dispersion	Fdisp	*		8.80	0.009	Yes	Decrease

The range values are the highest variation of the metric among all sites. Metrics of richness with range below five and of percentage below 10 were considered not valid. The Mann–Whitney test assessed the potential of metrics in distinguishing impact from preserved streams. The expectation is related to how we expected the metric to respond in face of impacts

* Not tested

(Biological Monitoring Working Party) by Junqueira & Campos (1998) (BMWP_1) and Loyola (2000) (BMWP_2). Expectations regarding the response of metrics can be found in Table 2.

Functional diversity metrics

The matrix of functional traits constructed to calculate functional diversity metrics was based on the traits compiled by Colzani et al. (2013) and on Poff et al. (2006) (see Appendix 1). The matrix included information only on insects because most non-insect taxa (e.g., Gastropoda and Oligochaeta) were not identified to genus level—which means that we did not have reliable trait information for these taxa. This unlikely biased our comparison with other metrics of diversity, because non-insect taxa were very few and generally were present in almost all communities (e.g., Oligochaeta).

To measure the functional diversity of each community (the group of six surber samples within each stream), we used two indices Functional Diversity (FD, Faith, 1992; Petchey & Gaston, 2002) and Functional Dispersion (FDisp, Laliberté & Legendre, 2010). FD is related to functional richness since it is calculated based on a trait-based dendrogram where the result is the sum of all branches connecting the species of a community, not accounting for the abundance of species (Petchey & Gaston, 2002). We used the Gower distance to construct the dendrogram, through cluster analysis, following recommendations by Podani & Schmera (2006). FDisp is a metric of functional structure, because it measures the average distance of species (abundance weighted) to the centroid of the communities in a trait-space, where the centroid is the average of the traits of species present in the

community (Anderson et al., 2006; Laliberté & Legendre, 2010). Both metrics were chosen because they are widely used in studies of functional ecology and can be calculated using only categorical traits (Cianciaruso et al., 2009; Laliberté & Legendre, 2010).

Phylogenetic diversity metrics

We used only insects to estimate phylogenetic diversity metrics because non-insect taxa were very distant phylogenetically from insects. This could bias our analysis by overestimating the phylogenetic diversity of sites that had non-insect taxa and by decreasing the importance of the phylogenetic structure of insects (Swenson, 2014).

We selected phylogenetic diversity (PD, Faith, 1992), mean pairwise distance (MPD, Webb, 2000), and mean nearest neighbor distance (MNND, Webb, 2000) as phylogenetic diversity metrics to access the evolutionary shared history of communities. PD is a metric of phylogenetic richness, where the value of PD is the sum of the branch lengths a phylogenetic supertree constructed using the species pool of the metacommunity. MPD is the average phylogenetic distance (based on the branch length separating species) of all combinations of pairs of species of a community. MNND is the average phylogenetic distance of species to the closest species of a community (Webb, 2000; Cianciaruso et al., 2009). These last two metrics were balanced by the abundance of the species, so we considered them as metrics of phylogenetic structure; MPD indicating the overall value of this structure and MNND indicating the terminal structure of the tree (Webb, 2000; Cianciaruso et al., 2009).

Following well described procedures that have been successfully used in several previous studies (Webb,

2000; Vamosi & Vamosi, 2007), we constructed a phylogenetic supertree (see Appendix 2 in Supplementary Material) with the insects present in our species pool. We used the software Mesquite (Maddison & Maddison, 2011), with reference to several manuscripts that used molecular data in phylogenies construction (see Appendix 2 to the complete list of references used in the supertree construction). When these data were not available, the topology was solved using taxonomy because it is a reasonable surrogate to phylogeny (Vamosi & Vamosi, 2007). Ages of 32 nodes were compiled (see Appendix 2) and branch lengths were assigned using the BLADJ algorithm in Phylocom (see Appendix 2). BLADJ spaces undated nodes evenly between dated nodes using an adjuster algorithm (Webb et al., 2008).

Multimetric index protocol

Statistical filters

We created our index following the protocol created by Hering et al. (2006), but also following the suggestions by Barbour et al. (1996), Tetratich (2000), and Klemm et al. (2003). The protocol consisted of statistical filters testing the metrics for range, sensitivity to anthropic impacts, and redundancy among them.

The first filter tested the range of metrics calculated from values of richness and percentages (e.g., percentage of EPT; see Table 2). The range is represented by values between the 25 and 75% quartiles of a boxplot. We excluded the richness metrics with a range lower than five (e.g., if a metric of richness vary between two and four it will be excluded because the metric's range is only two) and a percentage metric that does not vary more than 10%, as proposed by Klemm et al. (2003).

The metrics that passed the first filter were submitted to the sensitivity test, which analyzes the ability to distinguish reference conditions from impacted sites (Barbour et al., 1996). We performed a non-parametric Mann–Whitney *U* test to test the difference between reference and impacted sites (Hering et al., 2006). Because our metrics do not necessarily have a normal distribution, this test is appropriate for comparisons between the two groups (Ellison & Gotelli, 2004).

To test redundancy between the metrics of the same group, we used Spearman's correlation matrices. We

considered metrics with *r* values higher than 0.75 and *P* < 0.05 redundant (Hering et al., 2006). Among the redundant metrics, we excluded those with the lowest *U* values (Mann–Whitney *U* test) from the previous filter, which indicates lower sensitivity. Since we have a particular interest in the performance of phylogenetic and functional metrics, we also tested the redundancy between these metrics and metrics from other groups of metrics.

Standardization of metrics

The scoring method was continuous and was based on the index for streams of the Environmental Protection Agency of the United States (EPA) (Tetratich, 2000). Each metric assumed a value from 0 to 100 with an upper limit of 100 and is represented by the highest value of the 95% quartile using all sampling sites and a lower limit of 0.

To calculate the value of each metric we used the following formula:

$$\text{Metric value} = \text{Observed value} \times 100 \div 95\text{th upper value}$$

The standardization between metrics occurs in the equation itself, therefore generating values from 0 to 100. The highest score of the index is the number of final metrics multiplied by 100 (e.g. for an index of five metrics the highest score is 500). The value is then divided into five classes of equal size to represent the different classes of quality: very good, good, regular, poor, and bad.

Results

We collected 8638 aquatic macroinvertebrates that were identified as belonging to 199 *taxa*, mostly identified to genus level, but some groups also to higher levels (e.g., Gastropoda and Hirudinea). On an average, in the reference conditions, *taxa* richness was 41 (SD = 11.33) and abundance was 312 (SD = 223.70) individuals, while in the impacted sites richness was 34 *taxa* (SD = 13.22) with 460 (SD = 283.19) individuals.

After applying the range filter, we excluded metrics that did not vary significantly between streams, as individual richness for families and genera of Ephemeroptera, Plecoptera, Trichoptera, and Coleoptera. We also excluded the composition metrics percentage of

Coleoptera and percentage of Plecoptera, and the feeding group metric percentage of shredders (see Table 2).

The sensitivity filter retained metrics that did not respond to environmental impacts as richness to genus, richness of Chironomidae genus, total abundance, percentage of Chironomidae, percentage of Ephemeroptera, and ratio between Baetidae abundance and others Ephemeroptera. The filter also retained all feeding group metrics and the functional diversity metric Functional Diversity and the phylogenetic diversity metrics Phylogenetic Diversity and Mean Pairwise Distance (Table 2).

All taxonomic diversity metrics presented high correlation values, indicating redundancy among them. Therefore, we kept the metric richness of EPT families (Table 3). The composition metrics percentage of EPT, percentage of EPTC, and percentage of Trichoptera were also redundant; thus, we selected only the last one. Within the tolerance metrics, we selected BMWP_2 instead of BMWP_1. After comparing the phylogenetic and functional metrics functional dispersion and mean nearest neighbor distance to all other metrics, we did not find redundancy among them.

Finally, we selected the following metrics to integrate our multimetric index: richness of EPT families, percentage of Trichoptera, ratio between EPT abundance and Chironomidae abundance, BMWP_2, functional dispersion, and mean nearest neighbor distance. The multimetric index we developed (Table 4) had a score ranging from 0 to 600 and is divided into the following five classes: very good (score > 480), good (480 > score > 361), regular (360 > score > 241), poor (240 > score > 121), and bad (120 > score). The final index was composed of metrics of taxonomic diversity, composition, tolerance, and metrics of phylogenetic and functional diversity.

Discussion

Using a well-established protocol of statistic filters (Hering et al., 2006), we found that all the selected metrics, including one functional and one phylogenetic, discriminated impacted streams from reference ones, while not been redundant. Thus, functional and phylogenetic diversity metrics indeed responded to environmental impact and complemented the information

provided by classical metrics. Our main message is that future indices should consider including new metrics of functional and phylogenetic diversity to properly monitor multiple dimensions of biodiversity. This, however, has to be done carefully as some metrics might not have the potential to discriminate human impact.

The sensitivity filter excluded Functional Diversity and Phylogenetic Diversity—metrics that do not account for abundance, but selected functional dispersion, mean pairwise distance, and mean nearest neighbor distance—metrics that account for abundance (Cianciaruso et al., 2009). These results indicate that information about the functional structure (how many individuals and with which trait state) and phylogenetic structure (how many individuals of which clades) are important when one aims to separate impacted sites from reference ones. The sensitivity of Functional Dispersion also indicates that voltinism, respiration, refuge building, and trophic group indeed responded to environmental impacts, as we expected. Voltinism should determine the importance of environmental stability to species colonization. This is because univoltine species need stable environments for long periods to reach maturity and reproduce, while multivoltine species can reach maturity within a short period of time, and therefore can maintain populations in constantly disturbed sites. In other words, an environment that fits the fundamental niche of a multivoltine species in short intervals will be sufficient to complete its life cycle, but will unlikely be sufficient for univoltine ones (Weiher, 2011).

Regarding respiration, a transition from *taxa* with gill respiration to *taxa* with cutaneous respiration and ultimately, air respiration with spiracles, plastrons, or tracheae is expected with increasing environmental impacts (Dolédec et al., 2006). This prediction is based on species with aerial respiration be independent of water dissolved oxygen levels and species with cutaneous respiration be more resistant to lower dissolved oxygen levels (higher ratio of respiratory surface/body volume) than species with gill respiration (lower ratio of respiratory surface/body volume; Buck, 1962). However, against our expectations, we did not find lower levels of dissolved oxygen in impacted streams. A likely explanation is that, because our measures of oxygen were punctual, we were not able to detect variation in dissolved oxygen levels within a day, and probably, the respiration trait is

Table 3 Spearman’s correlation matrix testing the redundancy between metrics of the same group (classical diversity, composition, tolerance, phylogenetic diversity and functional diversity)

Measure group	S_f	S_g-Ch	S_fEPT	S_gEPT	S_fEPTC	S_gEPTC	H'_f	H'_g	%_EPT
Classical diversity	S_f	#####	0.80	0.72	0.79	0.73	0.45	0.70	
Classical diversity	S_g-Ch	#####	0.79	0.79	0.78	0.78	0.52	0.72	
Classical diversity	S_fEPT	#####	#####	0.94	0.92	0.90	0.70	0.70	
Classical diversity	S_gEPT	#####	#####	#####	0.93	0.97	0.76	0.67	
Classical diversity	S_fEPTC	#####	#####	#####	#####	0.95	0.77	0.71	
Classical diversity	S_gEPTC	#####	#####	#####	#####	#####	0.75	0.69	
Classical diversity	H'_f	#####	#####	#####	#####	#####	#####	0.76	
Classical diversity	H'_g	#####	#####	#####	#####	#####	#####	#####	#####
Composition	%_EPT	#####	#####	#####	#####	#####	#####	#####	#####
Composition	%_EPTC	#####	#####	#####	#####	#####	#####	#####	#####
Composition	%_T	#####	#####	#####	#####	#####	#####	#####	#####
Tolerance	EPT/Chiro	#####	#####	#####	#####	#####	#####	#####	#####
Tolerance	EPTC/Chiro	#####	#####	#####	#####	#####	#####	#####	#####
Tolerance	BMWP_1	#####	#####	#####	#####	#####	#####	#####	#####
Tolerance	BMWP_2	#####	#####	#####	#####	#####	#####	#####	#####
Phylogenetic diversity	MNND	#####	#####	#####	#####	#####	#####	#####	#####
Functional diversity	FDisp	#####	#####	#####	#####	#####	#####	#####	#####

Measure group	%_EPTC	%_T	EPT/Chiro	EPTC/Chiro	BMWP_1	BMWP_2	MNND	FDisp
Classical diversity	S_f						0.53	0.40
Classical diversity	S_g-Ch						0.61	0.44
Classical diversity	S_fEPT						0.65	0.54
Classical diversity	S_gEPT						0.68	0.58
Classical diversity	S_fEPTC						0.72	0.60
Classical diversity	S_gEPTC						0.65	0.60
Classical diversity	H'_f						0.60	0.51
Classical diversity	H'_g						0.52	0.53
Composition	%_EPT	0.88					0.63	0.51
Composition	%_EPTC	#####					0.71	0.69
Composition	%_T	#####	#####				0.68	0.54
Tolerance	EPT/Chiro	#####	#####	0.95			0.70	0.54
Tolerance	EPTC/Chiro	#####	#####	#####	0.55		0.71	0.65

Table 3 continued

Measure group	BMWP_1	BMWP_2	MNND	FDisp
Tolerance	0.94	0.60	0.58	
Tolerance	#####	0.55	0.46	
Phylogenetic diversity	#####	#####	0.63	
Functional diversity	#####	#####	#####	#####

See Table 2 for the acronyms. Bold values are correlations higher than 0.75, and considered redundant among them

Table 4 Macroinvertebrate Multimetric index and classification method of Cerrado streams (São Paulo State, Brazil)

Metric	Metric value
S_fEPT	Observed × 100/10
%_T	Observed × 100/31
EPT/Chiro	Observed × 100/13
BMWP_2	Observed × 100/142
MNND	Observed × 100/419
FDisp	Observed × 100/331
<i>Final classification</i>	
>480	Very good
<480 > 361	Good
<360 > 241	Regular
<240 > 121	Poor
<120	Bad

The final classification is the sum of all metric values

responding to that. This response to impacts in time is one of the main desirable characteristic of a biomonitoring tool (Bonada et al., 2006), and our functional traits are probably satisfying this goal.

Regarding refuge building, *taxa* that shelter on leaves and twigs are dependent on riparian vegetation and on constant input of allochthonous material to streams. The input of leaves is lower in impacted streams than in pristine sites, and likely restricted the colonization of these impacted sites to free-living species. The same reasoning applies to trophic groups. Despite the abundance of feeding groups that did not detect differences between the streams, in a multi-trait approach the lower abundance of leaf shredders (a trait state dependent of allochthonous material and riparian vegetation) in impacted streams should have contributed to the functional differences between impacted and reference streams. Moreover, the classical feeding groups approach decreases the amount of community responses as it clusters species in few categories, whereas in the functional diversity approach the possibility of responses is increased because it classifies each species according to multiple traits (Petchev & Gaston, 2002).

Functional approaches in stream ecology have been used for a long time (e.g., Townsend & Hildrew, 1994) and have been applied in biomonitoring with success (see Dolédec & Statzner, 2010 and Statzner & Beche 2010 for reviews about the use of functional approach in biomonitoring). Besides this tradition, little

attention has been given to integrate biomonitoring protocols and functional approaches. Most studies have focused in comparing taxonomic approaches against functional approaches, and in finding the relationship between trait states and environmental filters (Dolédec & Statzner, 2010; but see Feio & Dolédec, 2012 in attempting to integrate functional diversity and predictive models). Here, we gave an example on how traditional tools, as multimetric indices, can be improved and complemented by inserting functional diversity metrics.

Among phylogenetic structure metrics, only mean nearest neighbor distance was sensitive, indicating that the information needed to observe higher phylogenetic diversity in reference streams is the terminal structure of the phylogenetic tree (Swenson, 2014). Initially, mean nearest neighbor distance was used to infer about the role of competitive exclusion and environmental filtering assembling communities (Webb, 2000). This approach should be viewed with caution because coexistence mechanisms can lead to patterns contrary to the expected according to the community phylogenetics logic (Mayfield & Levine, 2010). So, we only suggest that environmental filters are more restrictive in impacted streams and drive communities to lower terminal phylogenetic diversity than reference streams, due to closely related tolerant clades with similar niche being filtered by these impacts. For example, considering Ephemeroptera, one could find three Baetidae genera (tolerant family) in impacted streams (lower MNND value), whereas one could find one Baetidae, one Leptophlebiidae, and one Leptohiphidae genera in reference streams (higher MNND value).

The lack of redundancy between functional and phylogenetic diversity metrics and all other metrics indicates that phylogenetic and functional metrics accessed different dimensions of the communities. This information proved to be different from classical taxonomic diversity and reliable for multimetric indices. In this sense, using functional diversity indeed proved that preserved streams enable a higher variety of life forms to inhabit them, including life strategies dependent on the presence of riparian vegetation—e.g., shredder species that build shelters of leaves. The higher phylogenetic diversity in preserved streams also indicates that environmental impacts diminish not only the taxonomic diversity of streams, but also the evolutionary history shared by

the component species. It means that by preserving streams we would also be preserving high levels of evolutionary diversity.

Despite the potential advantages of including functional and phylogenetic metrics in multimetric indices, there are also some potential drawbacks that we cannot ignore. First, using phylogenetic metrics needs a phylogenetic supertree for each ecoregion that is intended to monitor, because the results of the metrics are dependent on the regional species pool. That is, the construction of a phylogenetic supertree is unavoidable. Although the construction of supertrees can be a difficult task, recently, several molecular phylogenies of aquatic insects have been published, making the construction of these trees more accessible. Second, the autoecological knowledge for aquatic insects from temperate environments is much more complete than for tropical environments. This makes tropical studies more speculative, because we sometimes need to use information from studies from temperate environments or higher *taxa* generalizations (but see recent advances by Tomanova et al., 2006 and Colzani et al., 2013). Third, the calculation of functional and phylogenetic diversity is mathematically more complex than classical taxonomic diversity, and since one desirable characteristic of biomonitoring tools is simplicity (Bonada et al., 2006), we think that one should decide between ease calculation and information of multiple dimensions of biodiversity accessed. Besides, all functional and phylogenetic diversity metrics used in our study can be calculated in R using free available functions and packages (FD package, Laliberté et al., 2010, Picante package, Kembel et al., 2010), turning its use less costly. So, considering advantages and disadvantages, we suggest that new multimetric indices can benefit from the inclusion of functional and phylogenetic metrics to properly identify what should be monitored, restored, and protected.

Acknowledgments We thank Fabio T. T. Hanashiro for comments on an earlier version of this manuscript, Francisco Valente-Neto, Marcia C. P. Bueno and Luiz A. Joaquim for helping in the field, and Susana Trivinho-Strixino and Melissa O. Segura for helping in the identification of Chironomidae and Coleoptera, respectively. This study was partially funded by grant #2013/50424-1, #2011/15077-3 and #2013/20540-0, São Paulo Research Foundation (FAPESP) and by grant #480933/2012-0, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Conflict of interest All authors declare no conflict of interest in the preparation of this manuscript.

References

- Anderson, M. J., K. E. Ellingsen & B. H. Mccardle, 2006. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 9: 683–693.
- Barbour, M. T., J. Gerritsen, G. E. Griffith, R. Frydenbourg, E. Mccarron, J. S. White & M. L. Bastian, 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* 15: 185–211.
- Bonada, N., N. Prat, V. H. Resh & B. Statzner, 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annual Review of Entomology* 51: 495–523.
- Buck, J., 1962. Some physical aspects of insect respiration. *Annual Review of Entomology* 7: 27–56.
- Cianciaruso, M. V., I. A. Silva & M. A. Batalha, 2009. Diversidades filogenética e funcional: novas abordagens para a Ecologia de comunidades. *Biota Neotropica* 9: 93–103.
- Colzani, E., T. Siqueira, M. T. Suriano & F. O. Roque, 2013. Responses of aquatic insect functional diversity to landscape changes in atlantic forest. *Biotropica* 45: 343–350.
- Couceiro, S. R. M., N. Hamada, B. R. Forsberg, T. P. Pimentel & S. L. B. Luz, 2012. A macroinvertebrate multimetric index to evaluate the biological condition of streams in the Central Amazon region of Brazil. *Ecological Indicators* 18: 118–125.
- Cummins, K. W. & M. J. Klug, 1979. Feeding ecology of stream invertebrates. *Annual Review of Ecology and Systematics* 10: 147–172.
- Cummins, K. W., R. W. Merritt & P. Andrade, 2005. The use of invertebrate functional groups to characterize ecosystem attributes in selected streams and rivers in southeast Brazil. *Studies on Neotropical Fauna and Environment* 40: 69–89.
- Dolédéc, S. & B. Statzner, 2010. Responses of freshwater biota to human disturbances: contribution of J-NABS to developments in ecological integrity assessments. *Journal of the North American Benthological Society* 29: 286–311.
- Dolédéc, S., B. Statzner & M. Bournaud, 1999. Species traits for future biomonitoring across ecoregions: patterns along a human impacted river. *Freshwater Biology* 42: 737–758.
- Dolédéc, S., N. Phillips, M. Scarsbrook, R. H. Riley & C. R. Townsend, 2006. Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. *Journal of the North American Benthological Society* 25: 44–60.
- Durigan, G., M. F. D. Siqueira, G. Antonio & D. Correa, 2007. Threats to the cerrado remnants of the State of São Paulo, Brazil. *Scientia Agricola* 64: 355–363.
- Ellison, G. N. & N. J. Gotelli, 2004. A Primer of Ecological Statistics. Sinauer, Massachusetts.
- Faith, D. P., 1992. Systematics and conservation: on predicting the feature diversity of subsets of taxa. *Cladistics* 8: 361–373.
- Feio, M. J. & S. Dolédéc, 2012. Integration of invertebrate traits into predictive models for indirect assessment of stream functional integrity: a case study in Portugal. *Ecological Indicators* 15: 236–247.
- Ferreira, W. R., L. T. Paiva & M. Callisto, 2011. Development of a benthic multimetric index for biomonitoring of a neotropical watershed. *Brazilian Journal of Biology* 71: 15–25.
- Harvey, P. H. & M. D. Pagel, 1991. *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- Hering, D., C. K. Feld, O. Moog & T. Ofenböck, 2006. Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. *Hydrobiologia* 566: 311–324.
- Junqueira, V. M. & S. C. M. Campos, 1998. Adaptation of the BMWP method for water quality evaluation to Rio Das Velhas watershed (Minas Gerais Brazil). *Acta Limnologica Brasiliensia* 10: 125–135.
- Kembel, S. W., P. D. Cowan, M. R. Helmus, W. K. Cornwell, H. Morlon, D. D. Ackerly, S. P. Blomberg & C. O. Webb, 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.
- Kerans, B. L. & J. R. Karr, 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications* 4: 768–785.
- Klemm, D. J., K. Blocksom, F. Fulk, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann, D. V. Peck, J. L. Stoddard, W. T. Thoeny, M. B. Griffith & W. S. Davis, 2003. Development and evaluation of a macroinvertebrate biotic integrity index (MBII) for regionally assessing Mid-Atlantic Highlands streams. *Environmental Management* 31: 656–669.
- Laliberté, E. & P. Legendre, 2010. A distance-based framework for measuring functional diversity from multiple traits. *Ecology* 91: 299–305.
- Laliberté, E., B. Shipley, & M. E. Laliberté, 2010. Package ‘FD’. Measuring functional diversity (FD) from multiple traits, and other tools for functional ecology.
- Loyola, R. G. N., 2000. Atual estágio do IAP no uso de índices biológicos de qualidade. ACIESP, São Paulo.
- Maddison, W. P. & D. R. Maddison, 2011. *Mesquite: a modular system for evolutionary analysis*. Version 2: 75.
- Magurran, A. E., 2004. *Measuring Biological Diversity*. Blackwell Publishing, Oxford.
- Mayfield, M. M. & J. M. Levine, 2010. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters* 13: 1085–1093.
- Melo, S., C. Stenert, M. S. Dalzochio & L. Maltchik, 2014. Development of a multimetric index based on aquatic macroinvertebrate communities to assess water quality of rice fields in southern Brazil. *Hydrobiologia* 7: 1–14.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. Fonseca & J. Kent, 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- Oliveira, R. B. S., C. M. Castro & D. F. Baptista, 2008. Desenvolvimento de índices multimétricos para utilização em programas de monitoramento biológico da integridade de ecossistemas aquáticos. *Oecologia Brasiliensis* 12: 487–505.
- Oliveira, R., D. F. Baptista, R. Mugnai, C. M. Castro & R. M. Hughes, 2011. Towards rapid bioassessment of wadeable streams in Brazil: development of the Guapiacu-

- Macau Multimetric Index (GMMI) based on benthic macroinvertebrates. *Ecological Indicators* 11: 1584–1593.
- Petchev, O. L. & K. J. Gaston, 2002. Functional diversity (FD), species richness and community composition. *Ecology Letters* 5: 402–411.
- Petersen, R. C., 1992. The RCE: a riparian, channel, and environmental inventory for small streams in the agricultural landscape. *Freshwater Biology* 27: 295–306.
- Podani, J. & D. Schmera, 2006. On dendrogram based measures of functional diversity. *Oikos* 115: 179–185.
- Poff, N. L., J. D. Olden, N. K. Vieira, D. S. Finn, M. P. Simmons & B. C. Kondratieff, 2006. Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society* 25: 730–755.
- Sanderson, M. J., A. Purvis & C. Henze, 1998. Phylogenetic supertrees: assembling the trees of life. *Trends in Ecology and Evolution* 13: 105–109.
- Silva, J. M. C. & J. M. Bates, 2002. Biogeographic patterns and conservation in the South American Cerrado: a Tropical Savanna Hotspot. *BioScience* 52: 225–233.
- Statzner, B. & L. A. Beche, 2010. Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? *Freshwater Biology* 55: 80–119.
- Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson & R. H. Norris, 2006. Setting expectations for the ecological condition of streams: the concept for reference condition. *Ecological Applications* 16: 1267–1276.
- Suriano, M. T., A. A. Fonseca-Gessner, F. O. Roque & C. G. Froehlich, 2011. Choice of macroinvertebrate metrics to evaluate stream conditions in Atlantic Forest, Brazil. *Environmental Monitoring and Assessment* 175: 87–101.
- Swenson, N. G., 2014. *Functional and Phylogenetic Ecology* in R. Springer, New York.
- Tetratich, 2000. *A Stream Condition Index for West Virginia Wadeable Streams*. U.S. EPA Region 3 Environmental Services Division and Office of Water. Wheeling, Washington.
- Tilman, D., 2001. Functional diversity. *Encyclopedia of Biodiversity* 3: 109–120.
- Tomanova, S., E. Goitia & J. Helešić, 2006. Trophic levels and functional feeding groups of macroinvertebrates in neotropical streams. *Hydrobiologia* 556: 251–264.
- Townsend, C. R. & A. G. Hildrew, 1994. Species traits in relation to a habitat template for river systems. *Freshwater Biology* 31: 265–275.
- Vamosi, J. C. & S. M. Vamosi, 2007. Body size, rarity, and phylogenetic community structure: insights from diving beetle assemblages of Alberta. *Diversity and Distributions* 13: 1–10.
- Vandewalle, M., F. de Bello, M. P. Berg, T. Bolger, S. Dolédec, F. Dubs, C. K. Feld, R. Harrington, P. A. Harrison, S. Lavorel, M. Da Silva, M. Moretti, J. Niemela, P. Santos, T. Sattler, J. P. Sousa, M. T. Sykes, A. J. Vanbergen & B. A. Woodcock, 2010. Functional traits as indicators of biodiversity response to land use changes across ecosystems and organisms. *Biodiversity and Conservation* 19: 2921–2947.
- Vellend, M., W. K. Cornwell, K. Magnuson-Ford & A. Ø. Mooers, 2011. Measuring phylogenetic biodiversity. In Magurran, A. E. & B. J. McGill (eds), *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford: 194–207.
- Violle, C., M. L. Navas, D. Vile, E. Kazakou, C. Fortunel, I. Hummel & E. Garnier, 2007. Let the concept of trait be functional! *Oikos* 116: 882–892.
- Ward, J. V., 1992. *Aquatic Insect Ecology*. Wiley, New York.
- Warwick, R. M. & K. R. Clarke, 1995. New ‘biodiversity’ measures reveal a decrease in taxonomic distinctness with increasing stress. *Marine Ecology* 129: 301–305.
- Webb, C. O., 2000. Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *American Naturalist* 156: 145–155.
- Webb, C. O., D. D. Ackerly & S. W. Kembel, 2008. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24: 2098–2100.
- Weiherr, E., 2011. A primer of trait and functional diversity. In Magurran, A. E. & B. J. McGill (eds), *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford: 175–193.