Reliable sample sizes for estimating similarity among macroinvertebrate assemblages in tropical streams

Fabiana Schneck[1]* and Adriano S. Melo[2,3]

[1] Programa de Pós-Graduação em Ecologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, CP 15007, CEP 91501–970, Porto Alegre, RS, Brazil
[2] Departamento de Ecologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, CP 15007, CEP 91501–970, Porto Alegre, RS, Brazil
[3] Present address: Departamento de Ecologia, ICB, Universidade Federal de Goiás, CP 131, CEP 74001–970, Goiânia, GO, Brazil

Abstract – Studies in tropical streams are relatively few, and one of the still-unresolved methodological issues is sample size. Adequate sample size for temperate streams cannot be extrapolated for tropical sites, because of the differences in species richness and the proportions of rare species. We evaluated reliable sample size for estimation of resemblance among samples of macroinvertebrate assemblages inhabiting riffles of tropical streams, using the autosimilarity approach. Sample sizes were much larger than those currently employed in tropical studies. Sampling units consisted of individuals associated with single stones (15–20 cm). Evaluations employed the Bray-Curtis index for abundance data and the equivalent Sorensen index for presence-absence data. Autosimilarity curves were constructed using both sampling units and individuals. The estimation of resemblance among samples was strongly dependent upon sample size at reduced sampling effort, particularly for the Bray-Curtis index. For the same sampling effort, fixed counts of individuals obtained randomly from sampling units gave better estimations of resemblance, and their similarity curves tended to stabilize earlier than those using sampling units. A minimum of 9–15 sampling units (stones) or 150–850 individuals is necessary for adequate estimations of resemblance using presence-absence data, and 13–18 sampling units or 750–1550 individuals are required for relative abundance data in tropical streams.

Key words: Autosimilarity / sample representativeness / sampling effort / species richness / species relative abundances

Introduction

Most ecological surveys aim toward the recognition of spatio-temporal patterns of community or assemblage structure. The ability to detect such patterns (*e.g.*, community or assemblage-environment relationships), attributes (*e.g.*, species richness, diversity), and human impacts often varies with sample size (Lorenz *et al.*, 2004; Kennard *et al.*, 2006). Different results may be obtained by changing the sample size, and consequently erroneous conclusions may be reached. Adequate sample size depends on the effect size of the study. For instance, large effects caused by human impact can be detected using small sample sizes, whereas larger sample sizes are needed to detect slight differences among non-impacted nearby sites (Doberstein *et al.*, 2000). Adequate sample size is also dependent on the metric used to compare assemblages.

In species-rich assemblages, such as macroinvertebrates in tropical streams, most species are rare in the sample, and thus a large sampling effort (area sampled or number of individuals counted) is needed to obtain a reliable estimate of species richness. On the other hand, smaller sample sizes are sufficient to estimate diversity using diversity or biotic indices (Lloyd *et al.*, 1968; Magurran, 2004). For instance, Hughes and Herlihy (2007) and Maret *et al.* (2007) found that the estimation of adequate scores for indices of biotic integrity (IBI) required less sampling effort than estimates of fish species richness.

Sample representativeness can be characterized by its accuracy and precision. Accuracy measures how close the estimated value is to the real value, and in most cases it cannot be measured because the true composition and relative abundances of the members of an assemblage are rarely known (Cao *et al.*, 2003). Precision refers to how similar are repeated measurements. It can be estimated by randomly taking two replicate samples of the assemblage

---

*Corresponding author: fabiana.schneck@gmail.com

under the same conditions. Many studies with assemblages are based on resemblance among sites, measured in terms of species composition and relative abundances. Accordingly, a good sample size should result in high similarity between two samples obtained from the same assemblage and in the same conditions.

A straightforward approach to determine adequate sample size for studies relying on resemblance is the construction of curves of similarity-area/individuals, using large data sets (Weinberg, 1978; Kronberg, 1987; Schleier and van Bernem, 1998; Cao et al., 2001; Schmera and Eros, 2006). Such curves are constructed by calculating resemblance values between random draws of two replicate samples of a given size, of a single data set, and then taking the average. The procedure is repeated for different sample sizes up to the maximum possible sample size (half of the full data set). Such autosimilarity curves (Cao et al., 2002) usually attain an asymptote and thus an adequate or representative sample size is obtained when increases in sample sizes do not result in higher similarity values.

Repeated samples, even those of large sample sizes, are rarely identical in terms of species composition and relative abundances, reflecting small-scale heterogeneities in species distributions (Kronberg, 1987). This is particularly evident for samples of species-rich assemblages such as tropical streams, where many rare species are present in one or two sampling units (uniques, duplicates) or with one or two individuals (singleton, doubletons). These differences among repeated samples are similar to those observed among subsamples from a large sample and have two practical consequences. First, a resampling procedure should be employed to estimate statistics from subsamples. Second, the maximum similarity of two random subsamples of the same sample can be very different from the theoretical maximum value of the index (Wolda, 1981). Nevertheless, autosimilarity curves tend to attain an asymptote at a level that reflects the within-assemblage natural variation in species composition and relative abundances (Schleier and van Bernem, 1998).

An important aspect of the sampling design is whether sample size should be expressed using sampling units (area, volume, traps; hereafter sample-based) or number of individuals (hereafter individual-based) (Gotelli and Colwell, 2001). The tradition in many subfields of assemblage ecology, including stream macroinvertebrates, is the standardization of sampling effort by number of sampling units (e.g., Surber, artificial substrates, stones, quadrats, or traps; Gotelli and Colwell, 2001). However, Gotelli and Colwell (2001) argued that species are accumulated according to the number of individuals counted, not the area. The long-known positive relationship between species richness and area should thus result mostly from the need to sample large areas to obtain a large number of individuals. This distinction is of particular relevance in the comparison of regions that differ in the density of individuals. In this case, equal-area samples of two regions containing the same number of species can result in a very different number of species. A second important consequence of whether sample- or

individual-based sampling is employed refers to the patchy distribution of individuals. If individuals of many species are aggregated in space, species accumulation curves of individual-based collections will be steeper than those curves produced by sample-based collections (Gotelli and Colwell, 2001). In fact, the difference between the two curves can be used as a measure of patchiness (Chazdon et al., 1998).

Stream macroinvertebrates are widely used in basic research and monitoring programs in temperate regions. Studies on tropical streams are relatively few, and more recent (Melo et al., 2006; Wantzen et al., 2006). Although many studies have evaluated adequate sample size for macroinvertebrates in temperate streams (Doberstein et al., 2000; Lorenz et al., 2004; Schmera and Eros, 2006), only a few studies have carried out such evaluations for their tropical counterparts (Stout and Vandermeer, 1975). Adequate sample size for temperate streams cannot be easily extrapolated for tropical sites. Previous evidence suggests that tropical streams harbor more species than temperate streams (Stout and Vandermeer, 1975). Most importantly, tropical faunas are composed of many species that are only rarely detected in samples (Stout and Vandermeer, 1975; Melo, 2004). Accordingly, Stout and Vandermeer (1975) showed that early accounts of low species richness in tropical streams were artifacts resulting from low sample size. They used large data sets and a method to estimate species richness in extrapolated sample sizes to show that species accumulation curves of tropical streams increase slowly, and that for small sample sizes these curves remain below the levels of temperate curves. However, for increased sample sizes, curves of temperate streams tend to stabilize and attain an asymptote much earlier than those of tropical streams.

We evaluated sample representativeness of macroinvertebrate assemblages inhabiting riffles in tropical streams, using the autosimilarity approach. Three large sample-based data sets were obtained in three rocky streams at least 150 km apart in the Atlantic Rain Forest, southeast Brazil. The evaluations employed the widely used Bray-Curtis index and its qualitative or presence-absence version, the Sorensen index. Curves were obtained for sampling units and individuals. Because the Bray-Curtis index is based on more detailed information (relative abundances), we expected that adequate sample sizes would be larger than those for the Sorensen index. Also, we expected that adequate sample sizes for sample-based curves would be larger than those using individuals, because of the more-rapid accumulation of species in the latter.

Materials and methods

Study sites

We used three data sets for macroinvertebrates collected in streams located in protected Atlantic Rain Forest areas in southern Brazil. All sites were shaded

by primary or old- growth vegetation, their streambeds were free of deposited terrestrial sediments and were not subjected to local human impacts. Streambeds were similar among sites and composed of gravel, stones and boulders. The first data set was collected in the Carmo River (24°18'S, 48°25'W), a fourth-order stream, 10 m wide, baseflow during austral winter of 0.66 $m^3 \cdot s^{-1}$ and at an elevation of 520 m. The precipitation in the area is 1700–2000 mm and the vegetation is Tropical Ombrophilous Submontane-Montane Forest. The survey was carried out in July 1997 during the dry season, when discharge is constant, no spates are observed and streambed remains stable (Melo and Froehlich, 2004). The second data set is from the Ermida Stream (23°14'S, 46°56'W), a third-order stream located in the Serra do Japi at an elevation of 860 m. The mean annual precipitation in the region is 1400 mm and the vegetation is Tropical Semi-Deciduous Montane Forest. Sampling was carried out from September to mid-November 1996, during the end of the dry and the beginning of the wet seasons. The third data set was collected in the Cedro Stream (22°45'S, 45°28'W), a third-order stream, at an elevation of 950 m in the Serra da Mantiqueira, Pindamonhangaba. The vegetation is Tropical Evergreen Seasonal Submontane Forest. The collection was carried out in December 1998 and January 1999, in the middle of the rainy season, although no spates occurred during the sampling period. The three data sets are hereafter called Carmo, Japi, and Pinda, respectively. Additional information of the studied sites can be found in Morellato (1992) and Melo and Froehlich (2001a, 2001b, 2004).

Sampling and data processing

The sampling and sorting procedures were the same for all samples and were done by the same person. Sampling units were individual stones of 15–20 cm maximum diameter sampled in riffles. At each site, stones were obtained from many riffles in reaches 300–500 m long. We used a U-net with a 250-μm mesh to avoid the loss of active swimmers. Attached organisms were removed from the stones, and together with all visible invertebrates collected by the U-net, were fixed in 80% ethanol. Individuals associated to each stone were stored in separate plastic vials containing an appropriate label (Melo and Froehlich, 2001a). Seventy-five stones were collected at the Japi and Carmo sites, and 150 stones at Pinda. Individuals were identified to the lowest possible taxonomic level (usually genus for Ephemeroptera, Plecoptera and Trichoptera and family for the remaining groups) and then sorted as morphospecies. Sorting was aided by a reference collection, draws and photos. When separation of organisms into one or two morphospecies was doubtful, we used a conservative approach and pooled them in a single morphospecies. A previous evaluation has shown the effectiveness of using morphospecies, in comparison to genus or family data, in the recovery of small differences among stream assemblages (Melo, 2005).

Mites and chironomid larvae were not included in the analysis. These sample sizes are 3–6 times the size of samples used in previous studies in the region (Melo and Froehlich, 2001a).

The Carmo data set comprised 71 morphospecies and 2673 individuals, while the Japi data set had 66 morphospecies and 3759 individuals. The Pinda data set was richer than the Carmo and Japi, and included on average 101 morphospecies on 75 stones. A total of 117 morphospecies and 10 339 individuals were included in the full Pinda data set. Detailed descriptions of the data are available elsewhere (Melo and Froehlich, 2001a, 2001b).

Estimation of autosimilarity

For each data set, we drew randomly and without replacement an even number of sampling units ($n = 2$, 4, 6, 8, …) or a fixed number of individuals ($n = 100$, 200, 300, 400, …) from the total sampling units. We pooled the first $n/2$ sampling units or individuals to create a new sample, and the other $n/2$ sampling units or individuals to create another sample. We then calculated the Bray-Curtis and Sorensen similarity indices for the pair of samples. The process was repeated 10 000 times, and the average for each similarity index was plotted against sampling effort (number of sampling units or number of individuals pooled). We examined the behavior of the autosimilarity curves and assessed whether they reached an asymptote.

We constructed curves based on the difference between similarity values of two adjacent sample sizes, and used differences of 0.01 in similarity as a criterion to define the attainment of an asymptote and determine adequate sample sizes. We also calculated the 95% confidence interval (95% CI) using the 10 000 similarity values for 37 sampling units. Here, the criterion used to determine an adequate sample size was the smallest sample size included in the 95% CI of 37 stones. This was the maximum sample size available (half of the full data set) for the Carmo and Japi data sets. For the Pinda site, the entire data set (150 stones) was used to construct curves, but, similarly to the other two data sets, the 95% CI was 37 sampling units. The same procedure was employed for individuals, using the 95% CI for the pooled mean number of individuals observed in 37 sampling units. We opted to use the Bray-Curtis similarity index and its presence-absence Sorensen version because they are widely used in the ecological literature and are usually among the best-scored indices in previous evaluations (Faith *et al.*, 1987; Legendre and Legendre, 1998). The resampling procedure was automated using an algorithm written in the R environment (The R Development Core Team, 2008).

Results

Autosimilarity curves of the three stream sites were similar in general form (Fig. 1). They produced increased
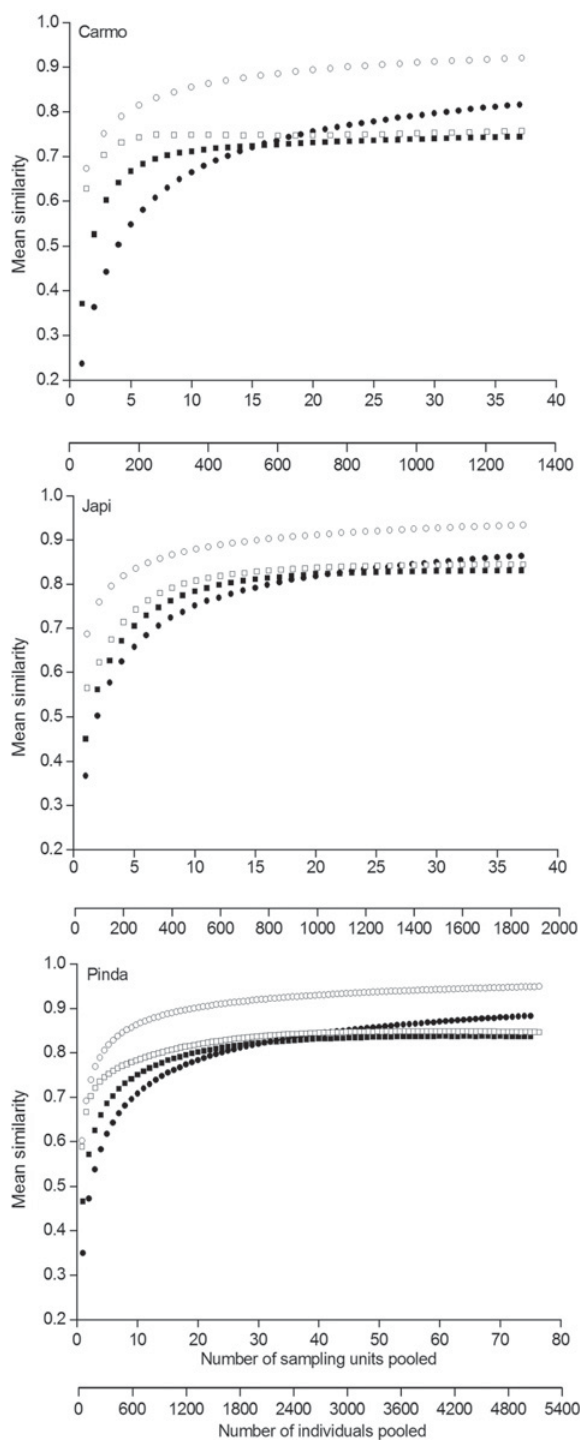
Fig. 1. Autosimilarity curves for three large data sets of stream macroinvertebrates in southeast Brazil. Individual-based curves (open symbols) and sample-based curves (filled symbols) are shown for the Sorensen index (squares) and for the Bray-Curtis index (circles).

similarity values as sample sizes increased, and then tended to stabilize and attain an asymptote. However, they differed in the sample size needed to attain the asymptote. Curves for the Sorensen index reached an asymptote in

all three data sets and for both forms of sampling effort (sampling units or individuals). On the other hand, curves using the Bray-Curtis index did not attain an asymptote, although they tended toward it. Individual-based curves tended to show a higher autosimilarity value than sample-based curves, and this was particularly evident for the Bray-Curtis index. Additionally, individual-based curves tended to stabilize earlier than those using sampling units.

Larger sample sizes were necessary for reliable estimates of similarity based on species relative abundance data (Bray-Curtis index) than those based on species presence-absence (Sorensen index) for both criteria used (differences of 0.01 and 95% CI; Tables 1 and 2). Using differences of 0.01 in similarity between samples as the criterion to define the attainment of an asymptote (see Fig. 2 for an example), a minimum of 6–9 sampling units (15–20-cm stones) or 250–400 individuals was necessary for the adequate estimation of similarity using presence-absence data, while 10–12 sampling units or 400–450 individuals were needed for abundance data (Table 1). However, the sample sizes determined by this criterion did not appear to produce reliable estimates, since most of them are located on the steep part of the curve, and not on the asymptote (Fig. 1).

The smallest sample size necessary to estimate a similarity value included within the 95% confidence interval for the large 37-stone sample size was larger than the sample sizes obtained using differences of 0.01. Sample sizes obtained with the 95% CI criterion were close to the asymptote (see Fig. 3 for an example; Fig. 1; Table 2). At least 9–15 sampling units or 150–850 individuals were necessary for reliable estimates of similarity using presence-absence data (Table 2). For abundance data, a minimum of 13–18 sampling units or 750–1550 individuals (Table 2) was required.

Discussion

Our results indicate that, at a relatively low sampling effort, the estimation of resemblance among samples of macroinvertebrate assemblages is strongly dependent upon sample size. However, at increased sample size, this dependence tends to disappear and attain a constant similarity value. Similar results were found by Wolda (1981) and Cao et al. (1997, 2002). This result shows the importance of determining adequate sampling efforts, because small sample sizes may underestimate similarities among samples of macroinvertebrate assemblages. Chao et al. (2005) suggested new indices for estimation of similarity that are less dependent on sample sizes. However, as pointed out by Cao et al. (1997), an index may be independent of sample size but have a low ability to distinguish different communities or assemblages. Accordingly, care should be taken in the interpretation of results based on sampling efforts determined by similarity indices that are not sensitive to sample size. Because species composition and relative abundance change with

Table 1. Number of sampling units and individuals needed to attain an asymptote, using as the criterion, differences of 0.01 in similarity between adjacent sample sizes in autosimilarity curves. Carmo and Japi datasets included 75 sampling units (individual stones). The Pinda dataset was composed of 150 sampling units and rarefied species richness expected for 75 stones is shown in parenthesis.

Sample	Observed richness	Total individuals	Sampling units (stones)		Individuals	
			Sorensen	Bray-Curtis	Sorensen	Bray-Curtis
Carmo	71	2673	6	12	250	400
Japi	66	3759	9	10	400	400
Pinda	117 (101)	10 339	8	11	300	450

Table 2. Number of sampling units and individuals needed to obtain autosimilarity values within the 95% confidence interval of sample sizes of 37 sampling units for each data set. Observed species richness and total individuals are shown in Table 1.

Sample	Sampling units (stones)		Individuals	
	Sorensen	Bray-Curtis	Sorensen	Bray-Curtis
Carmo	9	15	150	750
Japi	12	13	550	1000
Pinda	15	18	850	1550

increasing sample size, a similarity index that is not sensitive to these changes likely will not detect differences between natural communities or assemblages (Cao et al., 1997), and small sample sizes would erroneously be considered sufficient.

The sampling effort necessary for the estimation of resemblance was dependent on the use of presence-absence or relative abundance data, with the latter requiring additional sampling effort. Schmera and Eros (2006) found similar results, estimating sample representativeness of sample-based stream caddisfly fauna in Hungary. They observed asymptotes for presence-absence data (Jaccard index) but not for relative abundance data (Bray-Curtis index), with the latter being strongly dependent on sample size. Likewise, Cao et al. (2002) using macroinvertebrates, and Kennard et al. (2006) studying stream fish assemblages in Australia, found that smaller sample sizes are required to estimate resemblance using species composition than relative abundances. As in our study, Cao et al. (2002) and Schmera and Eros (2006) observed lower autosimilarity values for presence-absence data than for relative abundance data.

Our results showed that, for the same sampling effort, individual-based samples gave better estimates of resemblance (higher autosimilarity values) and tended to stabilize earlier than sample-based ones. Most river invertebrate assemblages have patchy spatial distributions, and this spatial aggregation causes species to occur nonrandomly among samples (Gotelli and Colwell, 2001). Therefore, sample-based efforts aggregate individuals and, consequently, accumulation of species is slow, resulting in lower similarity values when pairs of samples are compared. However, when individuals are collected randomly within a site, species accumulate faster, since spatial aggregation is eliminated and thus probabilities of collecting unseen species are increased in low sample

sizes. This results in high autosimilarity values and early curve stabilization. Similarly, Cao et al. (2002) found that individual-based autosimilarity curves of stream macroinvertebrates stabilized earlier than sample-based ones, for both presence-absence and relative abundance data. However, in contrast to our results, they observed similar autosimilarity values for individual- and sample-based curves.

Some studies have already outlined that high sampling effort may be needed in species-rich systems (Resh and Jackson, 1993; Li et al., 2001) or those with a high proportion of rare species (singletons or doubletons; Cao et al., 2001; Kanno et al., 2009), such as tropical streams (Melo and Froehlich, 2001a, 2001b; Melo, 2004). Using the criterion of inclusion in the 95% CI of the largest sample size, our data indicate that a minimum of 9–15 sampling units or 150–850 individuals should be obtained for the estimation of resemblance using assemblage composition in tropical streams. However, at least 13–18 stones or 750–1550 individuals are required to obtain reliable autosimilarity estimations for relative abundance data. Melo and Froehlich (2001b) sampled 25 stones to assess macroinvertebrate richness in Brazilian tropical streams, and Lake et al. (1994) collected 28 stones with the purpose of comparing species richness in Australian temperate and tropical streams.

In temperate streams, when individual-based methods are applied, usually from 100 to 300 organisms are collected and identified in biomonitoring programs (Carter and Resh, 2001). However, some biomonitoring programs are now counting a large number of individuals. For instance, the Environmental Monitoring and Assessment Program (EMAP) funded by the US Environmental Protection Agency (EPA) mandates fixed-count of 500 individuals (Hughes and Peck, 2008) and the European Union AQEM project suggests the use of all individuals found in 20 sampling units (Hering et al., 2004). Lorenz et al. (2004) evaluated the effect of sample sizes ranging from 100 to 700 individuals of macroinvertebrates on 45 metrics, and found a significant increase of reliability if at least 300 individuals were sampled. They also showed that many metrics, especially those based on abundance, attained good reliability only with 700 individuals. Another study using macroinvertebrates evaluated sample sizes ranging from 100 to 1000 individuals, and compared them with the results obtained when the whole samples are counted (Doberstein et al., 2000). The authors found that
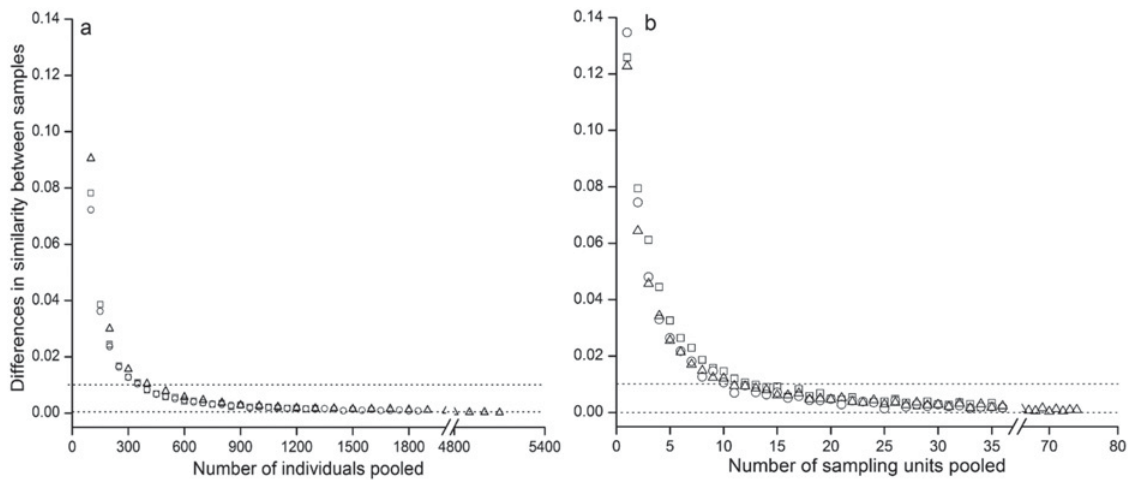
Fig. 2. Determination of adequate sample sizes using as criteria differences of 0.01 between similarity values of consecutive sample sizes, expressed as (a) number of individuals and (b) number of sampling units (15–20 cm stones) pooled using the Bray-Curtis index. □ = Carmo (adequate sample sizes: 400 individuals; 12 sampling units). ○ = Japi (400 individuals; 10 sampling units). △ = Pinda (450 individuals; 11 sampling units).
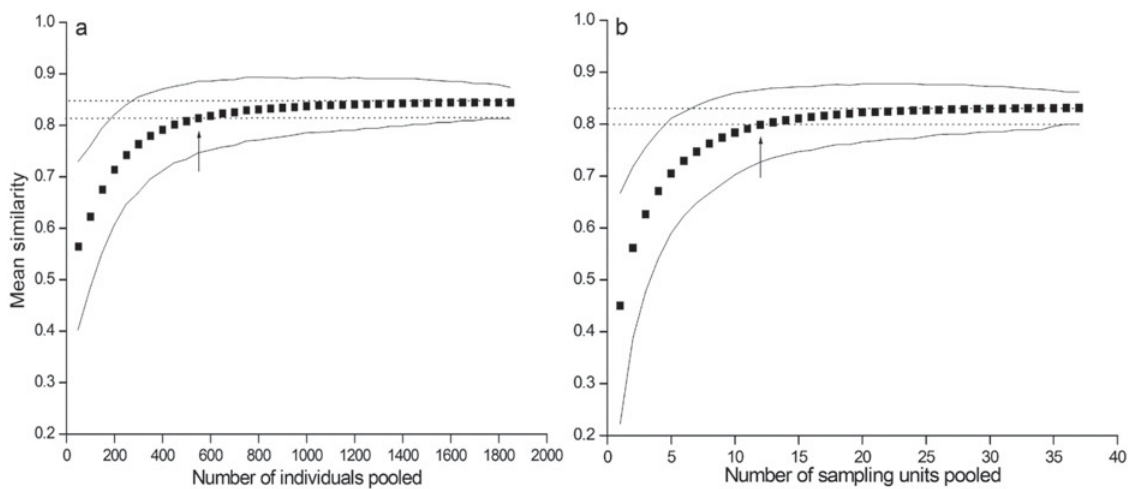
Fig. 3. Determination of adequate sample sizes using as criteria the smallest sample size necessary to estimate a similarity value included within the 95% confidence interval for a large 37-stone sample size. Data expressed as (a) number of individuals and (b) number of sampling units (15–20 cm stones) pooled using the Sorensen index for the Japi data set.

counting 100–300 individuals introduces high variability among same-site replicates, resulting in a low discriminatory power, and concluded that the entire sample had to be counted to obtain reliable results (*i.e.*, from 810 to over 3000 individuals). Adequate sample sizes, however, should likely vary according to the metric under study. For instance, Li *et al.* (2001) found that macroinvertebrate richness increased rapidly with the first 500–1000 individuals counted, and that curves did not attain an asymptote until more than 2000 individuals had been accumulated.

We suggest that, using presence-absence data, at least 9–15 sampling units (15–20-cm stones) or 150–850 individuals are necessary for estimation of resemblance among samples of macroinvertebrate assemblages in tropical streams, and 13–18 sampling units or 750–1550 individuals are necessary for estimations using relative abundance data. These numbers should be interpreted as a starting point, because adequate sample size is always dependent on the effect size of the study.

References

Cao Y., Williams W.P. and Bark A.W., 1997. Effects of sample size (replicate number) on similarity measures in river benthic Aufwuchs community analysis. *Water Environ. Res.*, *69*, 107–114.

Cao Y., Larsen D.P. and Hughes R.M., 2001. Evaluating sampling sufficiency in fish assemblage surveys: a similarity-based approach. *Can. J. Fish. Aquat. Sci.*, *58*, 1782–1793.

Cao Y., Larsen D.P., Hughes R.M., Angermeier P.L. and Patton T.M., 2002. Sampling effort affects multivariate comparisons of stream assemblages. *J. N. Am. Benthol. Soc.*, *21*, 701–714.

Cao Y., Hawkins C.P. and Vinson M.R., 2003. Measuring and controlling data quality in biological assemblage surveys with special reference to stream benthic macroinvertebrates. *Freshwater Biol.*, *48*, 1898–1911.

Carter J.L. and Resh V.H., 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. *J. N. Am. Benthol. Soc.*, *20*, 658–682.

Chao A., Chazdon R.L., Colwell R.K. and Shen T., 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.*, *8*, 148–159.

Chazdon R.L., Colwell R.K., Denslow J.S. and Guariguata M.R., 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of NE Costa Rica. *In*: Dallmeier F. and Comiskey A. (eds.), Forest biodiversity research, monitoring and modeling: Conceptual background and Old World case studies, Parthenon Publishing, Paris, 285–309.

Doberstein C.P., Karr J.R. and Conquest L.L., 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biol.*, *44*, 355–371.

Faith D.P., Minchin P.R. and Belbin L., 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, *69*, 57–68.

Gotelli N.J. and Colwell R.K., 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.*, *4*, 379–391.

Hering D., Moog O., Sandin L. and Verdonschot P.F.M., 2004. Overview and application of the AQEM assessment system. *Hydrobiologia*, *516*, 1–20.

Hughes R.M. and Herlihy A.T., 2007. Electrofishing distance needed to estimate consistent index of biotic integrity (IBI) scores in raftable Oregon Rivers. *T. Am. Fish. Soc.*, *136*, 135–141.

Hughes R.M. and Peck D.V., 2008. Acquiring data for large aquatic resource surveys: the art of compromise among science, logistics, and reality. *J. N. Am. Benthol. Soc.*, *27*, 837–859.

Kanno Y., Vokoun J.C., Dauwalter D.C., Hughes R.M., Herlihy A.T., Maret T.R. and Patton T.M., 2009. Influence of rare species on electrofishing distance when estimating species richness of stream and river reaches. *T. Am. Fish. Soc.*, *138*, 1240–1251.

Kennard M.J., Pusey B.J., Harch B.D., Dore E. and Arthington A.H., 2006. Estimating local stream fish assemblage attributes: sampling effort and efficiency at two spatial scales. *Mar. Freshwater Res.*, *57*, 635–653.

Kronberg I., 1987. Accuracy of species and abundance minimal areas determined by similarity area curves. *Mar. Biol.*, *96*, 555–561.

Lake P.S., Schreiber E.S.G., Milne B.J. and Pearson R.G., 1994. Species richness in streams: patterns over time, with stream size and with latitude. *Verh. Internat. Verein. Theor. Angew. Limnol.*, *25*, 1822–1826.

Legendre P. and Legendre L., 1998. Numerical Ecology, Second Edition, Elsevier Scientific, Amsterdam, 853 p.

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

Lloyd M., Inger R.F. and King F.W., 1968. On the diversity of reptile and amphibian species in a Bornean rain forest. *Am. Nat.*, *102*, 497–515.

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

Magurran A.E., 2004. Measuring Biological Diversity, Second Edition, Blackwell Science Ltd, Oxford, 260 p.

Maret T.R., Ott D.S. and Herlihy A.T., 2007. Electrofishing effort required to estimate biotic condition in southern Idaho rivers. *N. Am. J. Fish. Manage.*, *27*, 1041–1052.

Melo A.S., 2004. A critic of the use of jackknife and related nonparametric techniques to estimate species richness in assemblages. *Comm. Ecol.*, *5*, 149–157.

Melo A.S., 2005. Effects of taxonomic and numeric resolution on the ability to detect ecological patterns at a local scale using stream macroinvertebrates. *Arch. Hydrobiol.*, *164*, 309–323.

Melo A.S. and Froehlich C.G., 2001a. Macroinvertebrates in Neotropical streams: richness patterns along a catchment and assemblage structure between 2 seasons. *J. N. Am. Benthol. Soc.*, *20*, 1–16.

Melo A.S. and Froehlich C.G., 2001b. Evaluation of methods for estimating macroinvertebrate species richness using individual stones in tropical streams. *Freshwater Biol.*, *46*, 711–721.

Melo A.S. and Froehlich C.G., 2004. Substrate stability in streams: effects of stream size, particle size, and rainfall on frequency of movement and burial of particles. *Acta Limnol. Bras.*, *16*, 381–390.

Melo A.S., Bini L.M. and Carvalho P., 2006. Brazilian articles in international journals on Limnology. *Scientometrics*, *67*, 187–199.

Morellato L.P.C., 1992. História Natural da Serra do Japi, Editora da Unicamp/Fapesp, Campinas, Brazil.

Resh V.H. and Jackson J.K., 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. *In*: Rosenberg D.M. and Resh V.H. (eds.), Freshwater Biomonitoring and Benthic Macroinvertebrates, Chapman and Hall, New York, 195–233.

Schleier U. and van Bernem K.H., 1998. Statistical methods to determine sample size for the estimation of species richness and species' abundances in benthic communities. *Arch. Fish. Mar. Res.*, *46*, 205–223.

Schmera D. and Eros T., 2006. Estimating sample representativeness in a survey of stream caddisfly fauna. *Ann. Limnol. - Int. J. Lim.*, *42*, 181–187.

Stout J. and Vandermeer J., 1975. Comparison of species richness for stream-inhabiting insects in tropical and mid-latitude streams. *Am. Nat.*, *109*, 263–280.

The R Development Core Team, 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, http://www.R-project.org.

Wantzen K.M., Ramírez A. and Winemiller K.O., 2006. New vistas in Neotropical stream ecology: Introduction to the special volume. *J. N. Am. Benthol. Soc.*, *25*, 61–65.

Weinberg S., 1978. Minimal area problem in invertebrate communities of Mediterranean rocky substrata. *Mar. Biol.*, *49*, 33–40.

Wolda H., 1981. Similarity indexes, sample-size and diversity. *Oecologia*, *50*, 296–302.