APPLIED ISSUES

Untangling associations between chironomid taxa in Neotropical streams using local and landscape filters

FABIO O. ROQUE*, TADEU SIQUEIRA[†,‡], LUIS M. BINI[‡], MILTON C. RIBEIRO[§,¶], LEANDRO R. TAMBOSI[§], GIORDANO CIOCHETI[§] AND SUSANA TRIVINHO-STRIXINO**

Faculdade de Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados, Dourados, Brazil
[†]*Programa de Pós-graduação em Ecologia e Recursos Naturais, Universidade Federal de São Carlos, São Carlos, Brazil*
[‡]*Laboratório de Ecologia Teórica e Síntese, Universidade Federal de Goiás, Goiânia, Brazil*
[§]*Laboratório de Ecologia da Paisagem e Conservação, Universidade de São Paulo, São Paulo, Brazil*
[¶]*Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, ON, Canada*
**Laboratório de Entomologia Aquática, Universidade Federal de São Carlos, São Carlos, Brazil*

SUMMARY

1. Analyses of species association have major implications for selecting indicators for freshwater biomonitoring and conservation, because they allow for the elimination of redundant information and focus on taxa that can be easily handled and identified. These analyses are particularly relevant in the debate about using speciose groups (such as the Chironomidae) as indicators in the tropics, because they require difficult and time-consuming analysis, and their responses to environmental gradients, including anthropogenic stressors, are poorly known.

2. Our objective was to show whether chironomid assemblages in Neotropical streams include clear associations of taxa and, if so, how well these associations could be explained by a set of models containing information from different spatial scales. For this, we formulated *a priori* models that allowed for the influence of local, landscape and spatial factors on chironomid taxon associations (CTA). These models represented biological hypotheses capable of explaining associations between chironomid taxa. For instance, CTA could be best explained by local variables (e.g. pH, conductivity and water temperature) or by processes acting at wider landscape scales (e.g. percentage of forest cover).

3. Biological data were taken from 61 streams in Southeastern Brazil, 47 of which were in well-preserved regions, and 14 of which drained areas severely affected by anthropogenic activities. We adopted a model selection procedure using Akaike's information criterion to determine the most parsimonious models for explaining CTA.

4. Applying Kendall's coefficient of concordance, seven genera (*Tanytarsus/Caladomyia*, *Ablabesmyia*, *Parametriocnemus*, *Pentaneura*, *Nanocladius*, *Polypedilum* and *Rheotanytarsus*) were identified as associated taxa. The best-supported model explained 42.6% of the total variance in the abundance of associated taxa. This model combined local and landscape environmental filters and spatial variables (which were derived from eigenfunction analysis). However, the model with local filters and spatial variables also had a good chance of being selected as the best model.

5. Standardised partial regression coefficients of local and landscape filters, including spatial variables, derived from model averaging allowed an estimation of which variables

Correspondence: Fabio O. Roque, Faculdade de Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados, Dourados CEP 79825-070, Brazil. E-mail: roque.eco@gmail.com

were best correlated with the abundance of associated taxa. In general, the abundance of the associated genera tended to be lower in streams characterised by a high percentage of forest cover (landscape scale), lower proportion of muddy substrata and high values of pH and conductivity (local scale).

6. Overall, our main result adds to the increasing number of studies that have indicated the importance of local and landscape variables, as well as the spatial relationships among sampling sites, for explaining aquatic insect community patterns in streams. Furthermore, our findings open new possibilities for the elimination of redundant data in the assessment of anthropogenic impacts on tropical streams.

Keywords: aquatic insects, Atlantic Forest, landscape ecology, model selection, spatial eigenvectors

Introduction

A long-standing goal in stream ecology has been to relate species distributions to key environmental factors (Thienemann, 1954; Hynes, 1970; Townsend & Hildrew, 1994; Statzner, Hildrew & Resh, 2001; Poff et al., 2006). The ability to track how patterns of biodiversity change in response to environmental factors is important in the development of strategies for monitoring biodiversity. This goal has proved challenging, not only because of taxonomic impediments, but also because lotic communities are controlled by a number of factors that are in turn, influenced by the characteristics of the catchment (Allan, 1995; Wiens, 2002; Townsend et al., 2003). Nevertheless, the way the local environment interacts with the wider and dynamic landscape remains poorly understood (Bengtsson et al., 2002).

The 'species association' concept stems from the classic debate about biotic communities as discrete entities (after Clements) or as coincident groups of individualistic species (after Gleason) (Odum, 1971). Although in recent years, this dichotomy has been substituted by a combination of both views (Heino et al., 2003), applied ecologists are interested in taxon associations, i.e. groups of taxa recognised as clusters following the application of a clearly stated set of rules (Legendre, 2005). When an association has been identified, one can then concentrate on finding the ecological requirements common to most or all members of the group rather than describing the biology and habitat of each species. Conversely, species associations may be used to predict environmental characteristics (Legendre, 2005). Clearly, this approach has major implications for the selection of indicators for freshwater biomonitoring and conser-

vation, because it allows for a reduction in redundant information and focuses on taxa that can be easily handled in the field, and are suitable for rapid identification. Furthermore, in times of great concern about the effects of climate change, nations must follow the requirements of international conventions (UNFCCC, 1992) and initiate research programmes aimed at identifying species indicating large-scale environmental change. In this context, it would be highly desirable to find a relationship between groups of associated taxa and more easily measured predictor variables such as, for instance, landscape metrics (Williams, Margules & Hilbert, 2002; Metzger, 2008).

The dipteran family Chironomidae is a potentially useful group for exploring species associations because of its ecological importance and applied significance. In the tropics, it represents one of the most speciose groups in any aquatic environment. Their larvae feed in a variety of ways and occupy different habitats (Ferrington, 2008). Furthermore, Chironomidae species have wide-ranging life-histories, modes of locomotion and tolerance to oxygen deficit (Pinder, 1986). They are also considered to be a group requiring difficult and time-consuming analysis for identification to the genus level or lower, and there is much debate on the importance of including chironomid data in biomonitoring and conservation programmes (Rosenberg, 1992; Cranston, 2000; Hawkins & Norris, 2000; Rabeni & Wang, 2001; Nijboer, Verdonschot & Van Der Werf, 2005). Thus, the species association approach is relevant in the debate about using chironomids in biomonitoring programmes in the tropics. This is because, if a group of taxa exhibit similar responses to major environmental factors, particularly anthropogenic stressors, then one could select members of that group that are easily collected

and identified as a surrogate for the others. Such an approach could be a way of monitoring biodiversity in freshwater ecosystems, in situations where reliable species-abundance lists are lacking.

The problem of explaining chironomid taxon associations (CTA) based on environmental variables is a typical problem of model selection (see Stephens, Buskirk & del Rio, 2007). For this reason, the null hypothesis significance testing approach may be inappropriate, and an alternative is preferable. Here, we use a model selection approach based on Akaike's Information Criteria (AIC) (see Burnham & Anderson, 2002) to analyse CTA patterns. As the ecological processes interact dynamically at multiple spatial scales (Levin, 1992; Wagner & Fortin, 2005), we used local, landscape and spatial variables as predictors in our analyses. Specifically, our objectives here were twofold. First, we wanted to know whether chironomid assemblages in Neotropical streams have clear taxon associations (are there groups of concordant taxa that respond similarly to environmental gradients?). Second, if so, how well can these associations be explained by a set of models containing information from multiple spatial scales (local, landscape and spatial variables)? We used the idea of multiple filters, in which to establish a viable population in a particular site each species has to pass through several environmental filters operating at different spatial scales (Poff, 1997). We expected that the filtering processes would act on taxon characteristics, and only groups of concordant species that possess suitable traits would be able to pass the filter at each scale. We then considered the implications of our findings for the biomonitoring and conservation of Neotropical streams.

Methods

Study area and selection of streams

São Paulo state has an area of approximately 248 800 km^2 and a human population of about 40 million (22% of the Brazilian population). The region is composed of four major vegetation types, the coastal forest or Atlantic Rain Forest, the tropical seasonal forest or Atlantic semi-deciduous forest, the Mixed Atlantic Rain Forest and Cerrado (Brazilian savanna). The Atlantic Rain Forest grows at low to medium altitude (<1000 m a.s.l.) on the eastern slopes

of mountain chains running close to the coastline from southern to Northeastern Brazil. The Atlantic semi-deciduous forest and Cerrado extend across the plateau (usually >600 m a.s.l.) in the central and Southeastern parts of the country. The climate of São Paulo varies from tropical to subtropical. The Atlantic Rain Forest and the Mixed Atlantic Rain Forest have a warm and wet climate without a dry season, while the Atlantic semi-deciduous forest and Cerrado have a more seasonal climate with a comparatively severe dry season, generally from April to September. The Atlantic Forest is among the most threatened tropical forests in the world, having been reduced to 7% of their original cover in the interior region of São Paulo state, although the coastal region retains more than 30% of its original cover (Ribeiro *et al.*, 2009). Most remnants are found on sheltered, steep mountain slopes and are <50 ha in size (Ribeiro *et al.*, 2009).

The present study was conducted in 61 first- and second-order streams across the São Paulo state (Fig. 1). Of these, 47 were located in protected conservation areas, and 14 drained an impacted landscape. The streams from well-preserved regions are typical of Brazilian forested headwater streams, with water depth <50 cm, tree canopy coverage exceeding 70% of the channel, an absence of macrophytes, high dissolved oxygen concentration, low conductivity and slightly alkaline to acid water. Water temperature typically ranges from 15 to 23 °C. The riparian vegetation along all streams is well preserved. The 14 impacted streams lay in three of the most common land uses in the São Paulo state: *Eucalyptus* plantation ($n = 4$), sugar cane plantation ($n = 5$) and pasture ($n = 5$).

Chironomid sampling and identification

We used data on chironomid larvae distribution extracted from the 'Macroinvertebrates database' compiled by the research group of the 'Laboratório de Entomologia Aquática da Universidade Federal de São Carlos (SP)'. These data were collected during the dry seasons of 2001 (three streams), 2005 (24 streams) and 2006 (34 streams) using a Surber sampler (0.1 m^2 area and 250- μm mesh size). Our group has been continually visiting all sampling areas since 2001 and has not noticed any drastic change in land use during this period, particularly with regard to forested streams, which are in protected conservation areas.
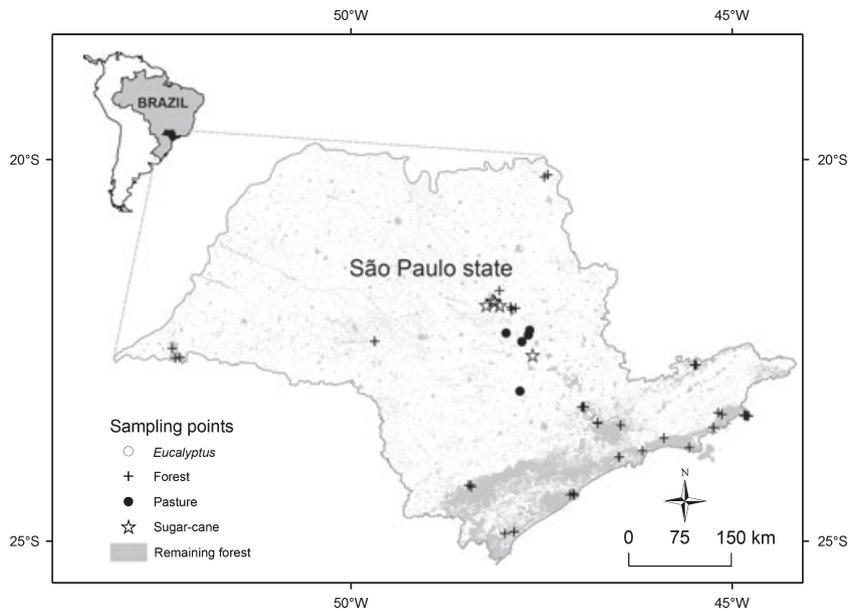
Fig. 1 The location of the study sites in São Paulo state, Brazil. The dominant land use for each sampling point is indicated by symbols.

Samples were taken at each stream on one occasion; six samples (three samples in pools and three in riffle sites) were taken randomly along a 100 m stretch of each stream. All specimens were counted and identified to the genus level, bearing in mind the limited knowledge available for Neotropical fauna. The specimens have been deposited in the collection of the 'Laboratório de Entomologia Aquática da Universidade Federal de São Carlos (SP)'. Illustrations of most taxa can be obtained from the senior author.

Chironomid taxon association response variables

We applied Kendall's coefficient of concordance (*W*), following Legendre (2005), to identify significantly associated groups of genera. This coefficient is a measure of the concurrence among several judges who are evaluating a given set of objects (Legendre, 2005). In our case, the 'judges' are chironomid genera. First, we conducted an overall global test of independence using the Hellinger-transformed abundance of all taxa collected (for details about Hellinger transformation, see Legendre & Gallagher, 2001). In this first test, the rank values in all genera are permuted at random, independently from genus to genus. Secondly, if the null hypothesis (independent rankings of the genera) was rejected, we performed an *a posteriori* analysis of concordance using the modified form of the permutation test to identify the genera that were significantly associated. Here, the null hypothesis is

the independence of the genus subjected to the test, with respect to all the other genera. The modified test differs from the overall test in the sense that only the genus under test is permuted. If a genus is concordant with one or several other genera, shuffling its values at random should break the concordance and have a conspicuous influence on the *W* statistics (Legendre, 2005). Subsequently, we explored through a principal component analysis (PCA) the possibility of more than one group of concordant taxa, and we eliminated taxa that occurred at low abundance and frequency and showed distinct patterns in the PCA plot (isolated groups). Finally, we used the sum of the abundances of the significantly associated genera [log (*x* + 1)] as the response variable representing 'CTA'.

Local variables

Conductivity and pH were measured at three different sections of each stream using a Horiba U-10 or a Yellow Springs-556 water checker equipped with multiple probes. We used the mean of these measures in the statistical analyses. Predominant substrata were estimated visually as the proportion of the stream bottom covered by boulder and cobble (>256 mm), gravel (2–255 mm), sand (0.125–2 mm) and mud (<0.125 mm). To assess physical and biological conditions in the riparian zone and stream channel morphology at the local scale, seven metrics and their respective scores from RCE – the 'Riparian Channel

and Environmental Inventory' for small streams (Petersen, 1992) were applied to each sampling site (consisting of 100 m of stream). These metrics were (i) land-use pattern beyond the immediate riparian zone; (ii) width of riparian zone from stream edge to field; (iii) completeness of riparian zone; (iv) vegetation of riparian zone within 10 m of channel; (v) retention devices; (vi) channel sediments and (vii) stream–bank structure. Thus, our RCE final score for each stream refers to the sum of the scores of these seven metrics.

Landscape metrics and macro-regional variables

Landscape ecology has been recognised as an important approach for understanding biodiversity patterns and the underlying processes and mechanisms in aquatic systems (Wiens, 2002). Part of its importance comes from the potential to provide data about landscape modification that can directly impact management efforts (Johnson, 2007). In this study, we incorporated compositional and configuration landscape metrics as potential predictors of CTA.

To compute the landscape metrics, we delineated a circular buffer area (500 m radius, *c.* 78.5 ha) around a point located in the centre of the stream channel of each of the 61 sampling sites (see Umetsu, Metzger & Pardini, 2008; for application of landscape metrics based on buffers of varying width). Buffers of different width were also delineated, but added no new significant information relative to the metrics derived from 500 m buffer. Some of our landscape metrics were derived from a land cover map at a scale of 1 : 50 000 from the Forestry Institute of São Paulo (Kronka *et al.*, 2005; Metzger *et al.*, 2008). Macro-regional climatic variables, derived from coarse-scale maps were also included in this set of variables. Landscape metrics derived from the land cover map were the land use cover, the total edge contrast index (TECI) and the edge density (see later for explanation). The macro-regional variables were the enhanced vegetation index (EVI), the rainfall and the irradiance, all of which are indicators of primary production and biomass accumulation. We also included a measure of altitude in our set of landscape metrics. A description of the metrics presented earlier is given below:

1. Land use cover. Area of forest, area of *Eucalyptus* plantation, area of agriculture, area of sugar cane plantation and area of urban areas. These metrics represent the total area cover by a land use within the buffer. This information was provided by the Forestry Institute of São Paulo for the year 2005. The cover classes were obtained by visual inspection of Landsat images, on a scale of 1 : 50 000 and with a spatial resolution of 30 m.

2. Total edge contrast index (TECI). This metric represents a comparison of all the edges of a particular patch across the entire buffer according to an 'edge contrast weight' (e.g. forest to forest = 0, forest to *Eucalyptus* plantation = 0.5, forest to agriculture = 0.7, forest to sugar cane plantation = 0.8 and forest to urban areas = 1).

3. Edge density (ED). This metric represents the total edge length divided by the total landscape area. ED = (E/A) 10000, where E is the total length (m) of edge in the landscape area and A is the total landscape area (m^2).

4. Enhanced vegetation index (EVI). This is a general measure of foliar biomass. The index is derived from the moderate resolution imaging spectroradiometer sensor (Bagan *et al.*, 2005) and was developed to optimise the vegetation signal with improved sensitivity in high biomass regions and to improve vegetation monitoring through decoupling of the canopy background signal and a reduction in atmospheric influences. We included four variables based on EVI measures: the EVI for the autumn (EVI AR) and for the winter (EVI AM) and both the range (EVI WR) and mean (EVI WM) between the two seasonal variables.

5. Rainfall was obtained from the São Paulo state Rainfall database.

6. Irradiance was obtained from the Brazilian Irradiation Atlas. We used mean values of the irradiation bands of São Paulo state.

7. Altitude was calculated using the mean value of the altitude in a circle of 500 m radius around each sampling site. We used the Digital Elevation Model provided by the Shuttle Radar Topography Mission (SRTM/NASA; http://www2.jpl.nasa.gov/srtm).

The cover area, TECI and ED metrics were calculated using FRAGSTATS 3.3 (McGarigal *et al.*, 2002). The EVI and altitude values were calculated using the Zonal statistics tool in ArcGis 9.

Spatial variables

Geographic space can be considered either as a factor responsible for ecological structures, or as a confounding variable that introduces biases when it is

ignored in a spatial analysis (see Legendre, 1993). In the second case, a correlation between pairs of observations (spatial autocorrelation) is commonly found in ecological data. Thus, more complex strategies for data analyses are required (Diniz-Filho, Bini & Hawkins, 2003).

We created spatial variables following the approach proposed by Borcard & Legendre (2002), where eigenvectors with positive eigenvalues were extracted from a spatial matrix consisting of the distances between all pairs of streams (see also Thioulouse, Chessel & Champely, 1995 for an early use of this approach). These spatial eigenvectors can be considered as independent schemes of how streams are geographically related to each other, expressed as a set of new variables indicating the spatial relationships (i.e. spatial structure) among them (Borcard *et al.*, 2004; Diniz-Filho & Bini, 2005; Griffith & Peres-Neto, 2006).

Recently, Diniz-Filho, Rangel & Bini (2008) stressed that shifting from a classical significance testing of a null hypothesis to an information theoretical approach does not guarantee a solution to the problem of spatial autocorrelation. The inclusion of spatial eigenvectors is a further safeguard against the problems caused by the presence of spatial structure in the model residuals in information theoretic approaches, just as much as in classical null hypothesis testing. Furthermore, spatial eigenvectors can account for the inability to measure an environmental variable which is spatially structured. We selected eigenvectors with significant patterns of spatial autocorrelation. We evaluated these patterns using Moran's I based correlograms with 10 distance class intervals, established in such a way that the number of pairs of sites compared at each distance class tends to be the same. A correlogram allows one to evaluate the level of autocorrelation as a function of spatial distance, thus providing a description of the level of spatial dependence in the data and, at the same time, is useful for examining the nature of the spatial structure (e.g. gradients or patches) displayed by each variable (eigenvectors in our case). We assessed the significance level of each Moran's I coefficient with 999 permutations (Sokal & Oden, 1978a,b), and a correlogram as a whole was considered to be significant (at a significance level of $P \leq 0.05$) if at least one of the Moran's I coefficients was significant at P/k, where k is the number of distance classes used (following the Bonferroni criterion; see Oden, 1984).

It is important to note that a spatial eigenvector is not a variable with ecological meaning, it is a mathematical construction. The amount of variation attributed to spatial variables is probably related to an 'unexplored' variable (e.g. dispersal) that could generate the observed pattern. Although there is an increasing number of studies on chironomids in the Neotropical region (Fittkau, 2001) and some information on dispersal patterns of adults and larvae in other regions (e.g. Delettre & Morvan, 2000; Silver, Wooster & Palmer, 2004), to date no study has examined specifically the ability of chironomids to disperse in heterogeneous landscapes as those sampled in our studied. Thus, given the absence of specific information, we consider our exploratory approach, including all distances among the sites, as the best option.

Data analyses

For each group of variables (i.e. local and landscape-scale variables), we used a PCA to reduce data dimensionality. Subsequently, we applied a randomisation procedure based on eigenvalues to determine the number of non-trivial axes to be used as local and landscape-scale predictors of CTA abundance. This method is based on the randomisation of the values within variables and the subsequent recalculation of the correlation matrix and associated eigenvalues. After repeating this process many times, a null distribution is generated for comparison with the observed eigenvalues. The P-value associated with the i-th eigenvalue is calculated as the number of random eigenvalues (which were obtained after reshuffling the data) that were equal to or larger than the observed (plus one) divided by the number of randomisations (plus one; see Peres-Neto, Jackson & Somers, 2005). We used 999 randomisations to assess the significance level of the eigenvalues. For each PCA, a method based on the broken-stick model (see Peres-Neto, Jackson & Somers, 2003) was used to determine which original variables were significantly correlated with the selected axes. As principal component axes scores are new orthogonal variables, their use in multiple regression analysis, as explanatory variables summarising each multidimensional dataset, can minimise the problems caused by multicollinearity (King & Jackson, 1999). Because of the reduced number of variables (non-trivial axes), over-parametrisation problems are also minimised, avoid-

ing the Freedman's paradox, i.e. too many variables with a small sample size (see Burnham & Anderson, 2002).

Following the recommendations of Burnham & Anderson (2002), we formulated *a priori* models containing different explanatory variables (local, landscape and spatial variables) representing biological hypotheses that could explain CTA. Model 1: CTA can be best explained by principal component axes representing the main patterns of local variables, such as pH, conductivity and substratum characteristics; Model 2: processes acting on wider landscape scales (e.g. the landscape structure) are responsible for shaping CTA, so it can be best explained by principal component axes representing the main patterns of landscape variables, such as land cover, edge density and rainfall; Model 3: there may be a spatially structured missing variable that influences the CTA (e.g. dispersal), so it can be best explained by spatial eigenvectors indicating the spatial relationships among streams. Nevertheless, it could be that CTA can be better explained by more than one model. So, based on the combination of these previous three models and using these parameters (local and landscape PCAs and spatial eigenvectors), we constructed four additional hypothetical models (Table 1), totalling seven models.

The minimum adequate model was selected using the AIC (Akaike, 1974; Burnham & Anderson, 2002). For each model, we calculated the AIC using the

Table 1 A set of *a priori* candidate models reflecting our hypotheses relating local variables, landscape metrics and spatial filters to chironomid taxon association; $n = 61$ in all cases. The number of parameters to be estimated in each model is also given. Notice that the maximum number of parameters (12) was estimated for the seventh model, because we extracted two nontrivial principal components (PCs) from local variables, three PC from landscape metrics and seven eigenvectors with significant spatial structures (see Results)

Model code	Model	Number of parameters
1	Local variables	2 (PCs)
2	Landscape variables	3 (PCs)
3	Spatial variables	7 (Spatial eigenvectors)
4	Local + landscape variables	5
5	Local + spatial variables	9
6	Landscape + spatial variables	10
7	Local + landscape + spatial variables	12

modified index for small sample sizes (AIC_c). The best models are those with the lowest AIC_c values, which are computed by:

$$AIC_c = n\log\left(\sum e_i^2 / n\right) + 2K + [2K(K+1)/(n-K-1)] \quad (1)$$

where e_i is the residual of the *i*-th observation for a candidate model, K is the number of parameters in the model and n is the sample size.

Each model was represented by the entire set of explanatory variables, i.e. all selected PCA axes representing the local filters, all selected PCA axes representing the landscape filters, and spatial variables representing geographical relationships between sampling sites. For example, model 1 (local filters alone) was represented by:

$$E(y) = \beta_0 + \beta_1 PCA1 + \beta_2 PCA2 \quad (2)$$

where β_0 is the intercept, and β_i is the partial regression coefficient, estimated by ordinary least squares.

We first calculated AIC_c values for each model using the spatial analysis in Macroecology software v.3 (Rangel, Diniz-Filho & Bini, 2006). After calculating AIC_c values for all models, we used the model with the lowest AIC_c value (the best model) and derived other metrics from it (Burnham & Anderson, 2002). We used the AIC_c of each model *i* to calculate the Δ_i value, which is the difference between AIC_{ci} and the minimum AIC_c found for the set of models analysed. Values of Δ_i higher than seven indicate models that have poor fit relative to the best model, whereas values <2 indicate models that are equivalent to the minimum AIC_c model (Burnham & Anderson, 2002). These cut-off values are arbitrary, but, more importantly, Δ_i can also be used to compute the Akaike weight of each model (w_i), which is a measure of the chance that a model is the best one. These weights can be calculated by:

$$w_i = \exp(-1/2\Delta_i) / \sum_i [\exp(-1/2\Delta_i)] \quad (3)$$

where $\exp(-1/2\Delta_i)$ approximates the likelihood of the model. w_i values were normalised across the set of candidate models to sum to one, and interpreted as probabilities (Johnson & Omland, 2004).

When no single model is clearly superior to the others in the set (i.e. $w_{best} < 0.9$), then model averaging is recommended (Johnson & Omland, 2004). This is usually performed to obtain robust estimates of linear

regression coefficients and to reduce the extent of bias that would arise if one uses only one model to make conclusions. In model averaging, it is possible to average the estimated coefficients of the explanatory variables across all models in which they are present, but weighting this average according to the w_i value of the model. Burnham & Anderson (2002) advise that, when possible, one should use inference based on more than one model, via model averaging. Thus, when necessary, we estimated new linear regression coefficients by averaging each coefficient across all models in which it was present (see details in Burnham & Anderson, 2002). For example, $\beta_1$localPC1 would be the average of the coefficients found in models 1, 4, 5 and 7, whereas $\beta_2$landPC2 would be the average of the coefficients found in models 2, 4, 6 and 7 (see Table 1).

Results

Local variables

The first two principal component axes (PC1 and PC2) were significant ($P < 0.001$) according to the randomisation procedure based on the eigenvalues (Table 2). The first principal component (explaining 37.1% of the total variability) tended to segregate 'forested streams', mostly characterised by high values for RCE, cobble and gravel, from 'impacted streams' or

Table 2 Loadings of the local environmental variables on the first three principal component axes

	Axes	
Variables	1	2
pH	0.40	**0.76**
Cond	−0.21	**0.75**
DO	0.20	**−0.69**
RCE	**0.87**	0.03
% cobble	**0.54**	**0.36**
% gravel	**0.72**	−0.12
% sand	**0.47**	−0.27
% mud	**−0.97**	0.02
Eigenvalues	2.97	1.83
% of variance	37.08	22.9
P	<0.001	<0.001

Significant loadings (according to the broken-stick criterion) are indicated in bold. The eigenvalues associated with the non-significant axes (third up to eighth) were 1.13, 0.81, 0.62, 0.34, 0.30 and 0.01.
Cond, Conductivity; DO, dissolved oxygen; RCE, Riparian Channel Environmental Protocol.

lowland streams, mostly covered by *Eucalyptus* plantation and dominated by mud (Table 2; Fig. 2). Most sites with low and positive scores on PC1 were in the Atlantic forest mountain range (eastern coast), while those with higher and negative PC1 values were in the non-mountainous areas of the state (western part; Fig. 2). The second principal component (22.9%) was positively correlated with pH and conductivity and negatively correlated with dissolved oxygen concentration, clearly separating streams located in karstic regions and some anthropogenically impacted streams (with high pH and conductivity values and low dissolved oxygen concentrations) from the others (Table 2; Fig. 2).

Landscape metrics and macro-regional variables

The first three axes selected based on landscape metrics accounted for 52.9% of the total variability (Table 3). The first axis expressed a gradient of land use; most sites in areas dominated by agriculture and sugar cane had high negative scores. Streams in areas dominated by *Eucalyptus* plantations had intermediate scores, whereas streams in forested areas had high positive scores (eastern coastal region). The second axis was related to plant biomass around the streams (EVI measures) and urban areas; it also roughly differentiated streams in forested areas and *Eucalyptus* plantations from those in other land uses. The third axis was most closely related to agriculture, *Eucalyptus*, rainfall and altitude, which approximately corresponds to the dominant landscape gradient at the state scale. The western part of São Paulo state is located on a plateau with a continental climate and has areas historically dominated by agriculture, whereas the eastern coastal part is mountainous with higher precipitation (Fig. 3).

Spatial variables

Of the eight eigenvalues extracted from the geographic distance matrix among sampling sites, seven presented significant ($P \leq 0.005$) patterns of spatial autocorrelation (i.e. at least one Moran's *I* coefficient of the correlogram was significant at the Bonferroni corrected level of probability). Specifically, the seventh eigenvector was discarded. The first eigenvector accounted for the highest level of spatial autocorrelation, representing a pattern with two major groups of similar values: one group with relatively high values
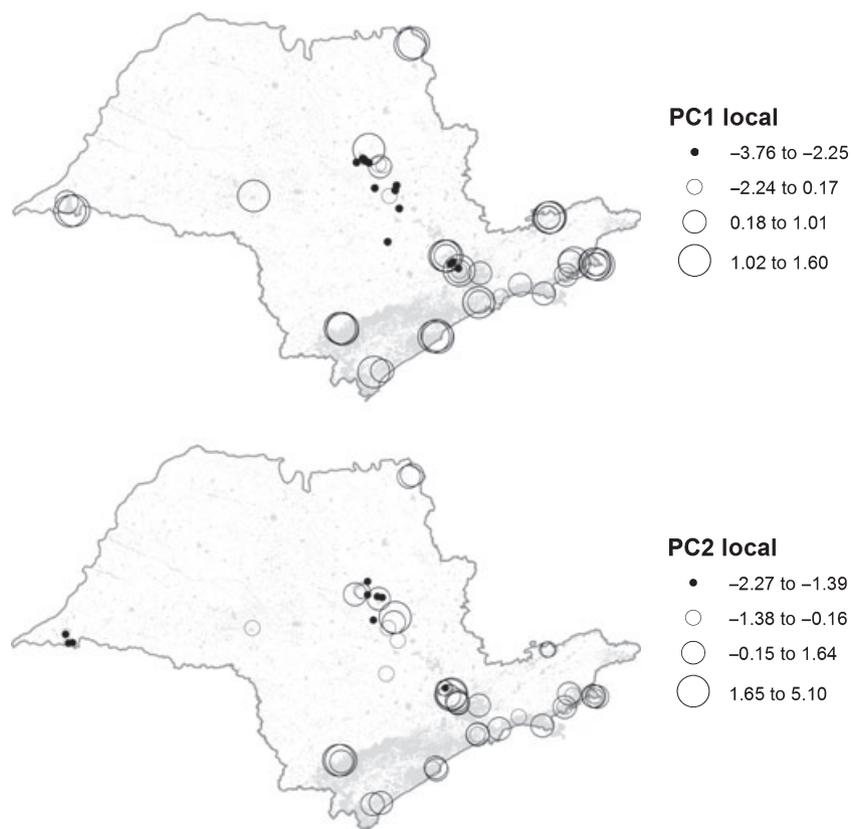
Fig. 2 Geographical patterns of principal component scores generated from local environmental variables.

and the other with relatively low values (Fig. 4). The other eigenvectors represented spatial patterns at more local scales. In these cases, significant autocorrelations coefficients were, in general, detected for only a few distance classes (detailed results not shown).

Chironomid assemblage descriptors: associated taxa

We detected a significant level of concordance in the first step of the Kendall's concordance analysis involving the 50 chironomid taxa collected in our study ($W = 0.09991$; $P = 0.001$ after 9999 permutations). *Tanytarsus/Caladomyia*, *Ablabesmyia*, *Tanypus*, *Parametriocnemus*, *Pentaneura*, Orthocladiinae gen. 2, *Nanocladius*, *Beardius*, *Cricotopus*, *Polypedilum* and *Rheotanytarsus* were identified as concordant taxa (Table 4). After applying a PCA to this group, we found that *Beardius*, *Cricotopus*, Orthocladiinae gen. 2 and *Tanypus* seemed to be unrelated to the others, and as they were also scarce, we excluded them from the concordant group. It is highly probable that *Beardius*, *Cricotopus*, Orthocladiinae gen. 2 and *Tanypus* were

grouped as concordant because of their absence from most sampling areas. The summed abundances of the seven remaining taxa (representing 63.95% of total abundance) were the response variable, i.e. the CTA, used in the models described later (Fig. 5). In general, associated genera were less abundant in streams located in the Atlantic Forest region in the eastern part of São Paulo state (Fig. 6).

Relationships between chironomid taxon association and filters: the minimum adequate model

The best-supported model according to the AIC_c criterion was the global model (Model 7: Table 5), which represented local and landscape environmental filters and spatial eigenvectors. This model explained 42.6% (adjusted R^2) of the variance in the CTA (Table 5) and had about a 49.4% probability of being the best model according to the Akaike weights (w_i). However, model 5 (local environmental filters and spatial eigenvectors) had a Δ_i value <2, with a 26.4% probability of being the best (Table 5) and also seemed to provide support for explaining CTA.

Table 3 Loadings of the landscape variables on the first three principal component axes

Variable	Axes		
	1	2	3
Area of agriculture	**−0.46**	−0.06	**0.59**
Area of sugar cane	**−0.51**	0.31	0.36
Area of *Eucalyptus*	−0.27	**−0.45**	**−0.58**
Area of forest	**0.80**	0.03	−0.16
Urban area	−0.06	**0.47**	**−0.41**
Landscape ED	**−0.45**	−0.26	0.30
Landscape TECI	**0.47**	**−0.44**	0.08
EVI WR	**0.71**	0.14	−0.07
EVI WM	**0.53**	**−0.67**	0.07
EVI AR	0.00	0.24	−0.14
EVI AM	0.16	**−0.62**	**0.47**
Altitude	**−0.54**	**−0.49**	**−0.45**
Irradiance	**−0.72**	0.06	0.01
Rainfall	**0.65**	0.38	0.33
Eigenvalues	3.66	2.09	1.66
% of variance	26.13	14.92	11.88
P	<0.001	<0.001	0.014

Significant loadings (according to the broken-stick criterion) are indicated in bold. The eigenvalues associated with the non-significant axes (fourth up to 14th) were 1.44, 1.33, 1.09, 0.76, 0.73, 0.48, 0.34, 0.20, 0.13, 0.11 and 0.00.
EVI, enhanced vegetation index; WR, winter range; WM, winter mean; AR, autumn range; AM, autumn mean; ED, edge density; TECI, total edge contrast index.

Because of this uncertainty, a model averaging procedure based on all models was necessary (see Methods). The third ($\Delta_i = 2.77$) and fourth ($\Delta_i = 4.62$) best models had w_i values of 0.124 and 0.049, respectively, whereas the summed weights for the models ranked 5–7 were 0.069.

Standardised partial regression coefficients of local and landscape filters, including seven spatial eigenvectors derived from model averaging, allowed an estimation of which axes were most strongly correlated with taxon associations, independently of spatial autocorrelation at distinct scales. CTA was negatively related to local PC1 (proportion of mud in the stream bed; standardised regression coefficient, $b' = -0.46$) and PC2 (pH and conductivity; $b' = -0.20$) as well as to landscape PC1 (area covered by forest; $b' = -0.43$) and PC2 (area covered by urban areas; $b' = -0.11$), and positively related to landscape PC3 (area covered by in *Eucalyptus* plantations; $b' = 0.11$). The comparison among the standardised regression coefficients (b') indicates that variation in CTA was better explained by local PC1 and landscape PC1.

Discussion

Ecological filters and chironomid associations

Even though it has been highly recognised that stream communities are structured by processes operating at several spatial scales (Vinson & Hawkins, 1998; Clarke *et al.*, 2008), disagreements prevail about which scales and factors are the most influential. The best model in our study was the global model, which included local and landscape variables and spatial information (i.e. the model with the highest number of parameters). In other words, models including variables from any single scale performed worse in fitting the data. Thus, our results add more evidence for the importance of different scales and spatial structure on community organisation (Richards, Johnson & Host, 1996; Johnson *et al.*, 2007).

The fact that the two models selected by AIC were a combination of local, landscape and spatial components indicates that variables at different scales can interact directly and indirectly in structuring CTA. In general, the abundance of the associated genera tended to be lower in streams characterised by a low proportion of mud in the stream bed (local PC1) and on streams located in forested areas (landscape PC1), particularly those located in the well-preserved areas of the Atlantic coastal region. Geomorphology influences stream ecosystems and their communities through differences in water velocity, input of sediments, nutrient loading, material retention, decomposition rates and food availability (Frissel *et al.*, 1986). The low abundance of associated chironomids in the Atlantic Forest region seems to be related to the geomorphology of the area, which is characterised by high level of declivity and by a rock matrix. These may result in streams with poor retention of detritus and low nutrient concentrations (Richards *et al.*, 1996), features that are directly related to the predictor variables local PC1 and landscape PC1. The low abundance of chironomids in the coastal streams may also be related to the high abundance of predators like shrimps and some plecopterans and trichopterans, when compared to streams inland (see De Souza & Moulton, 2005; Roque & Trivinho-Strixino, 2007).

We found that some physicochemical factors also contributed to explaining the abundance of associated
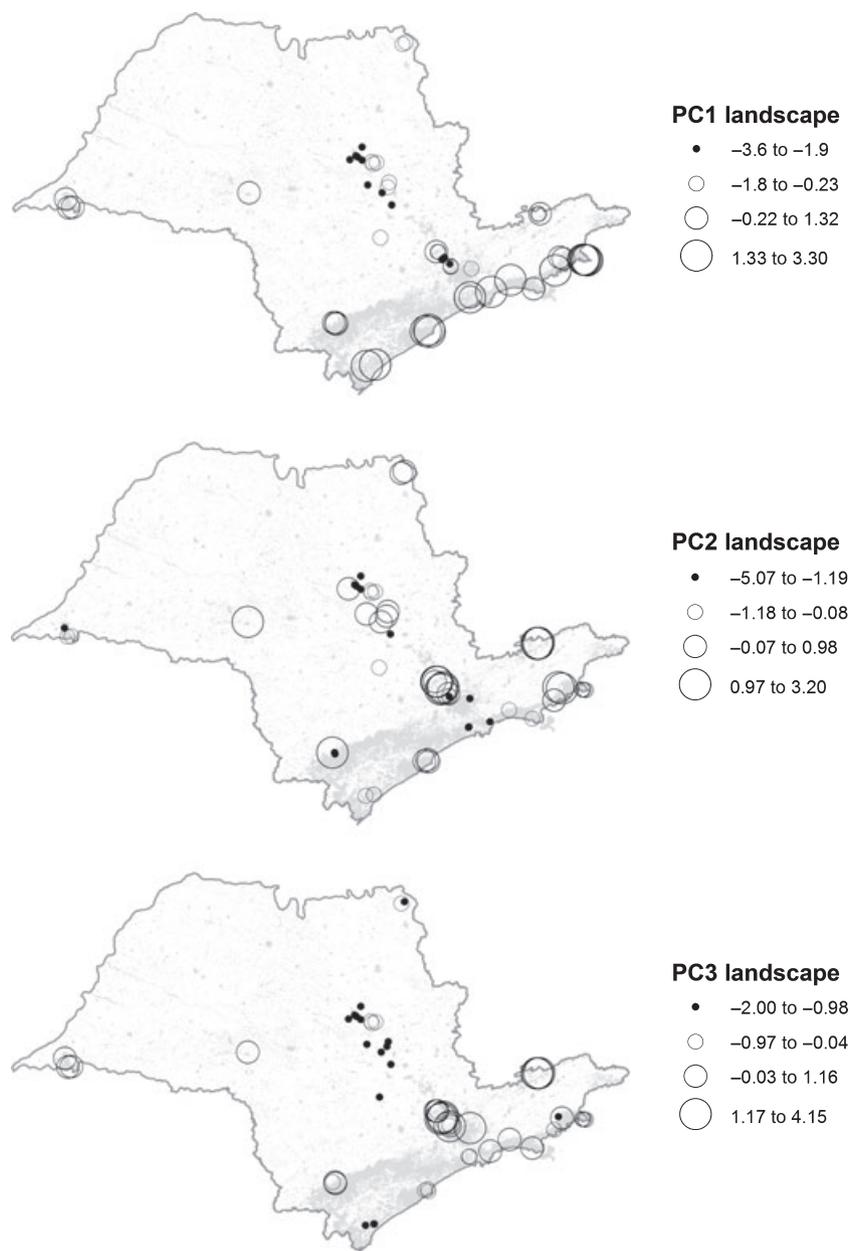
Fig. 3 Geographical patterns of principal component scores generated from local landscape variables.

taxa. A low abundance of chironomids was found in streams with high pH and conductivity values (local PC2), located in the karstic region of the São Paulo state. The responses of chironomids to pH and conductivity are variable with different patterns being observed in temperate and tropical regions (Cranston *et al.*, 1997). In our case, the low abundance of chironomids could be attributable to the high conductivity, because of high $CaCO_3$ concentration, high pH and nutrient-poor waters, which may negatively filter populations of some associated genera with

acidophilic tropical species (e.g. *Polypedilum* and *Tanytarsus*). Larvae of these genera are usually numerous in naturally acidic and nutrient-enriched and/or polluted streams (see Cranston *et al.*, 1997). Model average coefficients also indicate that landscape PC2 and PC3 were the predictors that contributed the least (smallest partial coefficients) to explaining CTA. The abundance of associated genera tended to be higher in streams surrounded by forested catchments (as indicated by EVI) and in streams not subjected to the direct influence of
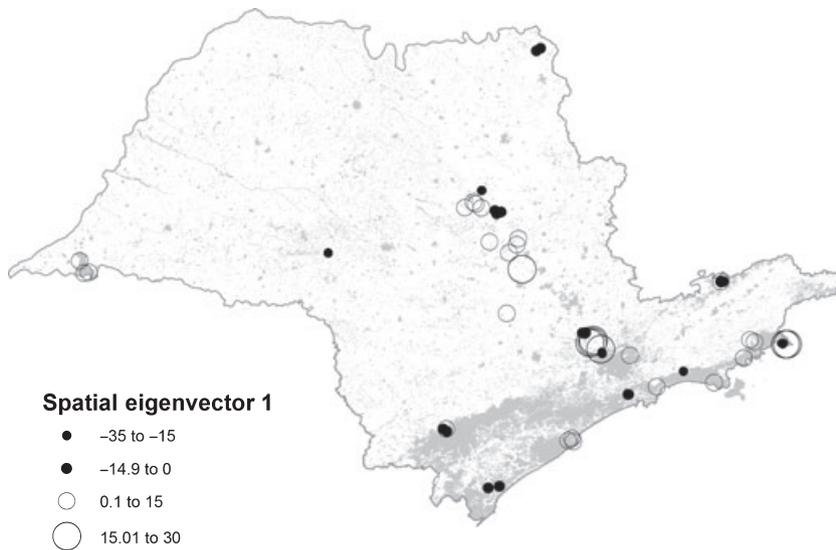
Fig. 4 Patterns depicted by spatial eigen-vector 1, showing the relationship among streams. Open symbols refer to positive values and closed to negative.

urban areas (landscape PC2). This last result was unexpected, particularly with regard to the genus *Polypedilum*, because this taxon is usually abundant in urban streams with high concentrations of nutrients and low dissolved oxygen (Cranston *et al.*, 1997; Helson, Williams & Turner, 2006). However, it should be noted that, although some of our streams were located near urban areas, none receive organic effluents. Finally, streams located in *Eucalyptus* plantations (landscape PC3) also supported a lower number of chironomid larvae. This agrees with previous studies that have shown low abundance of macroinvertebrates in *Eucalyptus* streams, which may be a consequence of the changed hydrology of the system and/or of the low quality of detritus as food (Abelho & Graça, 1996). Note that streams surrounded by *Eucalyptus* forests were also characterised by a high proportion of muddy substratum and, according to our analysis, such streams should have supported a high abundance of chironomid larvae. Thus, the interplay between local and landscape filters can be visualised in this example. Specifically, although the abundance of chironomids is predicted to increase in muddy streams, the occurrence of *Eucalyptus* may alter this pattern.

The inclusion of spatial information among the best-supported model indicates that other factors affected assemblage structures, such as colonisation processes, unmeasured spatially autocorrelated local environmental variables and dispersal pathways among habitat patches (Caley & Schluter, 1997). Furthermore, anthropogenic impacts are not randomly distributed in Southeastern Brazil. Most degraded areas are in the sedimentary basin, where intensive land conversion to agriculture (e.g. sugar cane and coffee) and pasture has occurred. These factors probably reflected some environmental variables that were spatially structured and that were captured by the spatial variables given by the eigenfunction analysis. Last but not least, models that included solely local or landscape variables were among the worst in explaining CTA, but when the spatial information was accounted for, and combined with environmental variables at both scales, they explained CTA in an appreciable manner. So, we urge future studies on stream insects' distribution to incorporate spatial information into data analysis.

Overall, our main result adds to the increasing number of studies that have pointed out the importance of considering both local and regional variables for explaining aquatic insect community patterns in streams. Moreover, CTA seems to respond in a predictable way to land use-gradients. This opens a promising new avenue for the exploration of patterns of chironomid distribution in tropical streams.

Implication for biomonitoring and environmental assessment in Neotropical streams

Most studies on surrogacy of aquatic biodiversity have focused on examining the effects of taxonomic resolution on assemblage patterns and the utility of

Table 4 Results of the *a posteriori* test of concordance involving all chironomid larvae collected

Taxa	Total abundance	Mean abundance (SD)	r_j	W_j	P	PH
Tanytarsus/Caladomyia	2307	37.81 (74.45)	0.176	0.192	0.000	0.005
Ablabesmyia	311	5.09 (9.99)	0.163	0.180	0.000	0.005
Tanypus	2	0.03 (0.17)	0.162	0.178	0.001	0.033
Parametriocnemus	1035	16.96 (46.71)	0.157	0.174	0.000	0.005
Pentaneura	178	2.91 (6.49)	0.154	0.171	0.000	0.005
Orthocladiinae gen. 2	3	0.04 (0.38)	0.149	0.166	0.001	0.025
Nanocladius	54	0.88 (2.10)	0.149	0.166	0.000	0.005
Orthocladiinae gen. 3	183	3 (19.5)	0.141	0.158	0.016	0.470
Beardius	18	0.29 (0.92)	0.139	0.156	0.000	0.013
Cricotopus	195	13.19 (13.30)	0.128	0.146	0.000	0.017
Parachironomus	33	0.54 (3.33)	0.126	0.143	0.008	0.269
Polypedilum	905	14.83 (20.43)	0.125	0.143	0.000	0.005
Pelomus	31	0.50 (1.44)	0.122	0.139	0.002	0.062
Clinotanypus	30	0.49 (1.51)	0.121	0.139	0.005	0.153
Rheotanytarsus	952	15.60 (32.87)	0.121	0.138	0.001	0.033
Onconeura	80	1.31 (3.76)	0.106	0.124	0.002	0.089
Coelotanypus	32	0.52 (3.19)	0.105	0.123	0.005	0.165
Labrundinia	65	1.06 (2.03)	0.105	0.123	0.004	0.137
Cryptochironomus	66	1.08 (2.82)	0.103	0.121	0.004	0.133
Stenochironomus	135	2.21 (4.68)	0.101	0.119	0.002	0.080
Pseudochironomini gen.3	223	3.65 (18.75)	0.094	0.112	0.010	0.319
Larsia	509	8.34 (29.56)	0.093	0.111	0.013	0.375
Pseudochironomini gen. 2	4	0.06 (0.30)	0.090	0.108	0.043	1.022
Zavreliella	14	0.22 (1.01)	0.086	0.105	0.017	0.476
Fittkauimyia	20	0.32 (1.05)	0.078	0.096	0.046	1.060
Dicrotendipes	5	0.08 (0.27)	0.073	0.092	0.080	1.670
Stempellinella	88	1.44 (5.01)	0.066	0.085	0.036	0.578
Thienenanimyia	11	0.18 (0.73)	0.065	0.084	0.036	0.936
Cladotanytarsus	4	0.06 (0.39)	0.065	0.084	0.036	0.990
Harnischia gen. 2	8	0.13 (0.66)	0.064	0.083	0.086	1.670
Djalmabatista	199	3.26 (7.95)	0.063	0.082	0.092	1.670
Thienemanniella	149	2.44 (8.65)	0.061	0.079	0.064	1.404
Stempellina	8	0.13 (0.68)	0.052	0.071	0.082	1.670
Paratendipes	114	1.86 (5.57)	0.051	0.070	0.082	1.670
Rheocricotopus	16	0.26 (1.19)	0.049	0.068	0.169	2.535
Corynoneura	114	1.86 (4.56)	0.043	0.063	0.106	1.702
Procladius	22	0.36 (1.66)	0.024	0.043	0.221	3.088
Lauterborniella	4	0.06 (0.50)	0.019	0.039	0.247	3.208
Chironomus	212	3.47 (11.38)	0.014	0.034	0.385	4.411
Paraphaenocladius	3	0.04 (0.38)	0.005	0.025	0.421	4.411
Harnischia gen. 1	4	0.06 (0.30)	0.003	0.023	0.368	4.411
Lopescladius	49	0.80 (2.14)	0.001	0.021	0.407	4.411
Macropelopini	2	0.03 (0.17)	−0.005	0.015	0.425	4.411
Aff. *Stackelbergina*	3	0.04 (0.28)	−0.006	0.014	0.487	4.411
Gymnometriocnemus	2	0.03 (0.17)	−0.007	0.013	0.452	4.411
Endotribelos	550	9.01 (16.45)	−0.014	0.006	0.430	4.411
Fissimentum	7	0.11 (0.40)	−0.014	0.006	0.605	4.411
Zavrelimyia	5	0.08 (0.41)	−0.018	0.002	0.593	4.411
Psedouchironomini gen.1	3	0.04 (0.21)	−0.029	−0.008	0.687	4.411
Orthocladiinae gen. 1	11	0.18 (0.91)	−0.049	−0.028	0.947	4.411

The 50 taxa were ordered by the values of the partial concordance statistics, r_j and W_j, to facilitate interpretation.
P = permutational probability based upon 9999 random permutations. PH = probability after Holm adjustment. H_0: This species is not concordant with most of the others – Reject H_0 at $\alpha = 0.05$.

Abundance of CTA
log scale

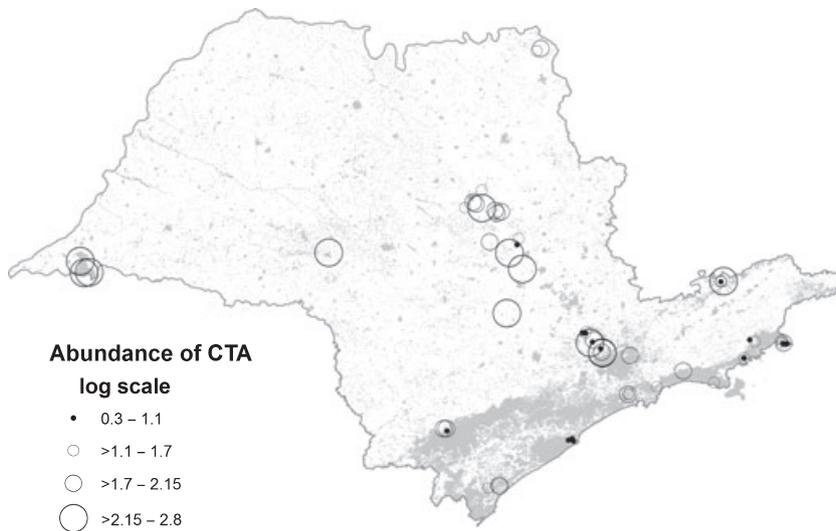- 　0.3 – 1.1
- ○　>1.1 – 1.7
- ○　>1.7 – 2.15
- ◯　>2.15 – 2.8

Fig. 5 Geographical distribution of chironomid taxon association across São Paulo state.

higher taxa in bioassessment (Heino & Soininen, 2007). We have found a low number of associated genera (14% of the total number of taxa), which means that most taxa responded individualistically to the environmental gradients. In this context, although few data are available about the evolution of traits among chironomids, it should be remembered that some genera seem to have high phylogenetic inertia in some traits (e.g. all species of *Rheotanytarsus* are rheophilous), while others do not (e.g. different species of *Polypedilum* live in different habitats). In other words, the use of genera as bioindicators may be adequate for detecting general patterns only when within-genus variability, in terms of responses to environmental gradients, is low (see Poff *et al.*, 2006). In relation to the utility of higher taxa as surrogate of groups of taxa, our findings indicate that the responses of chironomid associations to environmental gradients do not follow any high-level taxonomic classification (i.e. associated taxa belong to different tribes and subfamilies). This indicates that previous suggestions that subfamilies or tribes might be useful in biomonitoring of Neotropical streams are not well founded (Roque, Corbi & Trivinho-Strixino, 2000).

Although associated taxa represented a small proportion of the total generic richness, they accounted for 63.95% of total numerical abundance. This result indicates that we were dealing with the most common genera and that groups of these taxa respond similarly to environmental gradients. However, despite the fact that CTA responded predictably to environmental

and anthropogenic factors at distinct scales, the important question remains of whether monitoring just one of the associated genera as a surrogate for the whole group of concordant taxa would be sufficient for assessing impacts in streams? Unfortunately, the situation is not so simple. First, our study included a marked environmental gradient, ranging from extensive monocultures of sugar cane to well-preserved areas, so the pattern detected may be valid only for strong gradients arising from human activities. In these cases, other well-established bioindicators are also expected to perform well (see Bonada *et al.*, 2006). Second, the concordant taxa *Tanytarsus/Caladomyia* complex, *Polypedilum* and *Pentaneura* are among the most diverse, frequent and abundant genera in Neotropical streams (Coffman *et al.*, 1992; Roque & Trivinho-Strixino, 2007; Roque *et al.*, 2007). Together with the *Corynoneura* group, they account for more than 40% of chironomid diversity in Southeastern Brazil's Atlantic Forest streams (Roque *et al.*, 2007). These taxa are characterised by different species that inhabit distinct types of environments. In this context, taxonomic resolution and taxon associations have practical consequences for freshwater biomonitoring. Identification to genus or the analysis of associations of genera should be used with caution for some monitoring purposes, because a number of genera contain species with different tolerances to a particular impact (see Cranston, 2000; Raunio, Paavola & Muotka, 2007). For example, some species of *Chironomus* can colonise extremely acid or hot habitats (Armitage, Cranston & Pinder, 1995; Cranston *et al.*,
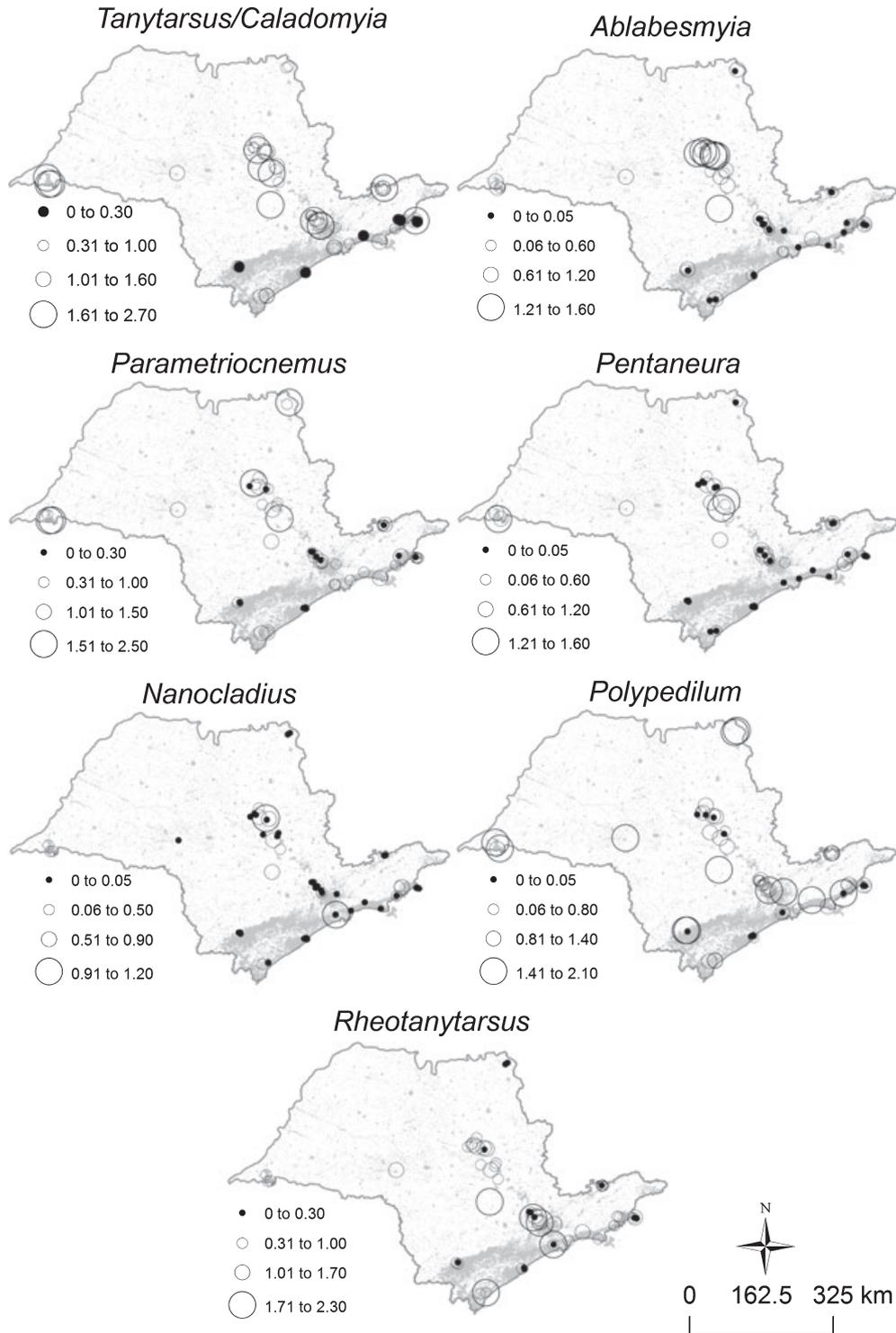
Tanytarsus/Caladomyia

Ablabesmyia

Parametriocnemus

Pentaneura

Nanocladius

Polypedilum

Rheotanytarsus

Fig. 6 Spatial patterns of chironomid-associated genera across São Paulo state. Symbols refer to classes of abundance (log scale).

1997), whereas others are found only in pristine streams (Correia & Trivinho-Strixino, 2007). Thus, using generic level identification could in some cases result in misinterpretations of stream conditions.

Another fundamental issue is whether GIS metrics can explain patterns in community structure and make predictions in biomonitoring programmes. The use of landscape metrics to explain patterns in

Table 5 Models of chironomid taxon association in the studied streams, built with different sets of explanatory variables, with their relative weights (w_i) derived from the Akaike information criterion (AIC$_c$) and adjusted R^2

Model code	Model	R^2-adjus.	AIC$_c$	Δ_i	w_i
7	Local + landscape + spatial	0.426	108.18	0	0.494
5	Local + spatial	0.351	109.43	1.251	0.264
6	Landscape + spatial	0.356	110.95	2.770	0.124
2	Landscape	0.195	112.80	4.623	0.049
3	Spatial	0.266	113.30	5.125	0.038
4	Local + landscape	0.214	114.26	6.086	0.024
1	Local	0.122	116.74	8.563	0.007

Δ_i = difference between AIC$_c$-i and the minimum AIC$_c$ found for the set of models compared.

the biodiversity of stream macroinvertebrates is controversial. Some authors have reported strong relationships between catchment-derived variables and macroinvertebrate community organisation (e.g. Townsend *et al.*, 1997; Feld & Hering, 2007), whereas other studies have revealed a stronger explanatory power of habitat characteristics measured at the reach scale (e.g. Death & Joy, 2004; Heino, Mykra & Kotanen, 2008). In general, our results suggest that single-scale variables should not be taken into account alone, which adds more evidence that GIS metrics are valuable for predicting biological patterns in streams only when linked with local environmental features (Pyne, Rader & Christensen, 2007).

In conclusion, chironomid associations in Southeastern Brazilian streams appear to be influenced by local and landscape filters. Moreover, the influence of some unmeasured and spatially structured environmental factors, as suggested by the importance of the spatial variables, cannot be discarded. This finding, although requiring caution, opens new possibilities for the use of a reduced number of taxa in the assessment of anthropogenic impacts on tropical streams.

Acknowledgments

References

Abelho M. & Graca M.A.S. (1996) Effects of eucalyptus afforestation on leaf litter dynamics and macroinvertebrate community structure of streams in Central Portugal. *Hydrobiologia*, **324**, 195–204.

Akaike H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Allan J.D. (1995) *Stream Ecology: Structure and Function of Running Waters*. Chapman & Hall, London.

Armitage P.D., Cranston P.S. & Pinder L.C.V. (1995) *The Chironomidae: The Biology and Ecology of Non-biting Midges*. Chapman Hall, New York.

Bagan H., Wang Q., Watanabe M., Yang Y. & Ma J. (2005) Land cover classification from MODIS EVI times-series data using SOM neural network. *International Journal of Remote Sensing*, **26**, 4999–5012.

Bengtsson J., Engelhardt K., Giller P., Hobbie S., Lawrence D., Levine J., Vilà M. & Wolters V. (2002) Slippin' and slidin' between the scales: the scaling components of biodiversity-ecosystem functioning relations. In: *Biodiversity and Ecosystem Functioning – Synthesis and Perspectives* (Eds M. Loreau., S. Naeem & P. Inchausti.), pp. 209–220. Oxford University Press Inc, New York.

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.

Borcard D. & Legendre P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbor matrices. *Ecological Modelling*, **153**, 51–68.

Borcard D., Legendre P., Avois-Jacquet C. & Tuomisto H. (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, **85**, 1826–1832.

Burnham K.P. & Anderson D.R. (2002) *Model Selection and Multimodel Inference. A Practical Information-theoretical Approach*. Springer, New York.

Caley M.J. & Schluter D. (1997) The relationship between local and regional diversity. *Ecology*, **78**, 70–80.

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

Coffman W.P., De La Rosa C., Cummins K.W. & Wilzbach M.A. (1992) Species richness in some Neotropical (Costa Rica) and Afrotropical (West Africa) lotic communities of Chironomidae (Diptera). *Netherlands Journal of Aquatic Ecology*, **26**, 229–237.

Correia L.C.S. & Trivinho-Strixino S. (2007) New species of *Chironomus* Meigen (Diptera: Chironomidae: Chironominae) from Brazil. *Zootaxa*, **1504**, 53–68.

Cranston P.S. (2000) Monsoonal tropical *Tanytarsus* van der wulp (Diptera: Chironomidae) reviewed: new species, life histories and significance as aquatic environmental indicators. *Australian Journal of Entomology*, **39**, 138–159.

Cranston P.S., Cooper P.D., Hardwick R.A., Humphrey C.L. & Dostine P.L. (1997) Tropical acid streams – the chironomid (Diptera) response in northern Australia. *Freshwater Biology*, **37**, 473–483.

De Souza M.L. & Moulton T.P. (2005) The effects of shrimps on benthic material in a Brazilian island stream. *Freshwater Biology*, **50**, 539–550.

Death R.G. & Joy M.K. (2004) Invertebrate community structure in streams of the Manawatu-Wanganui region, New Zealand: the roles of catchment versus reach scale influences. *Freshwater Biology*, **49**, 982–997.

Delettre Y.R. & Morvan N. (2000) Dispersal of adult aquatic Chironomidae (Diptera) in agricultural landscapes. *Freshwater Biology*, **44**, 399–411.

Diniz-Filho J.A.F. & Bini L.M. (2005) Modelling geographical patterns in species richness using eigenvector based spatial filters. *Global Ecology and Biogeography*, **14**, 177–185.

Diniz-Filho J.A.F., Bini L.M. & Hawkins B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, **12**, 53–64.

Diniz-Filho J.A.F., Rangel T.F.L.V.B. & Bini L.M. (2008) Model selection and information theory in geographical ecology. *Global Ecology and Biogeography*, **17**, 479–488.

Feld C.K. & Hering D. (2007) Community structure or function: effects of environmental stress on benthic macroinvertebrates at different spatial scales. *Freshwater Biology*, **52**, 1380–1399.

Ferrington L.C. Jr (2008) Global diversity of non-biting midges (Chironomidae; Insecta-Diptera) in freshwater. *Hydrobiologia*, **595**, 447–455.

Fittkau E.J. (2001) Amazonian Chironomidae (Diptera, Chironomidae): a contribution to chironomid research in the neotropics. *Amazoniana-Limnologia et Oecologia Regionalis Systemae Fluminis Amazonas*, **16**, 313–323.

Frissel C.A., Liss W.J., Warren C.E. & Hurley M.D. (1986) A hierarchical framework for stream habitat classification: viewing streams in the watershed concept. *Environmental Management*, **10**, 199–214.

Griffith D.A. & Peres-Neto P.R. (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.

Hawkins C.P. & Norris R.H. (2000) Effects of taxonomic resolution and use of the subsets of the fauna on the performance of RIVPACS-type models. In: *Assessing the Biological Quality of Fresh Waters. RIVPACS and Other Techniques* (Eds J.F. Wright, D.W. Sutcliffe & M.T. Furse), pp. 217–228. Freshwater Biological Association, Ambleside.

Heino J., Muotka T., Mykrä H., Paavola R., Hämäläinen H. & Koskenniemi E.S.A. (2003) Defining macroinvertebrate assemblage types of headwater streams: Implications for bioassessment and conservation. *Ecological applications*, **13**, 842–852.

Heino J., Mykra H. & Kotanen J. (2008) Weak relationships between landscape characteristics and multiple facets of stream macroinvertebrate biodiversity in a boreal drainage basin. *Landscape Ecology*, **23**, 417–426.

Heino J. & Soininen J. (2007) Are higher taxa adequate surrogates for species-level assemblage patterns and species richness in stream organism? *Biological Conservation*, **137**, 78–89.

Helson J.E., Williams D.D. & Turner D. (2006) Larval chironomid community organization in four tropical rivers: human impacts and longitudinal zonation. *Hydrobiologia*, **559**, 413–431.

Hynes H.B.N. (1970) *The Ecology of Running Waters*. University of Toronto Press, Toronto.

Johnson R.K. (2007) Environmental Assessment meets Landscape Ecology meets Land use Planning. *Freshwater Biology*, **52**, 907–907.

Johnson J.B. & Omland K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology and Evolution*, **19**, 101–108.

Johnson R.K., Furse M.T., Hering D. & Sandin L. (2007) Ecological relationships between stream communities and spatial scale: implications for designing catchment-level monitoring programmes. *Freshwater Biology*, **52**, 939–958.

King J.R. & Jackson D.A. (1999) Variable selection in large environmental data sets using principal components analysis. *Environmetrics*, **10**, 67–77.

Kronka F.J.N., Nalon M.A., Matsukuma C.K. et al. (2005) *Inventário florestal da vegetação do Estado de São Paulo*. Secretaria do Meio Ambiente/Instituto Florestal. Imprensa Oficial Press, São Paulo. p. 200.

Legendre P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.

Legendre P. (2005) Species associations: the Kendall coefficient of concordance revisited. *Journal of Agricul-*

tural, Biological, and Environmental Statistics, **10**, 226–245.

Legendre P. & Gallagher E.D. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.

Levin S.A. (1992) The problem of pattern and scale in ecology. *Ecology*, **73**, 1943–1967.

McGarigal K., Cushman S.A., Neel M.C. & Ene E. (2002) *Fragstats: Spatial Pattern Analysis Program for Categorical Maps*. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at: http://www.umass.edu/landeco/research/fragstats/fragstats.html.

Metzger J.P. (2008) Landscape ecology: perspectives based on the 2007 IALE world congress. *Landscape Ecology*, **23**, 501–504.

Metzger J.P., Ribeiro M.C., Ciocheti G. & Tambosi L.R. (2008) Uso de índices de paisagem para a definição de ações de conservação e restauração da biodiversidade do Estado de São Paulo. In: *Diretrizes para Conservação e Restauração da Biodiversidade no Estado de São Paulo* (Eds R.R. Rodrigues, C.A. Joly, M.C.W. Brito, A. Paese, J.P. Metzger, L. Casatti, M.A. Nalon, N. Menezes, N.M. Ivanauskas, V. Bolzani & V.L.R. Bononi), pp. 120–127. Secretaria do Meio Ambiente & FAPESP, São Paulo.

Nijboer R.C., Verdonschot P.F.M. & Van Der Werf D.C. (2005) The use of indicator taxa as representatives of communities in bioassessment. *Freshwater Biology*, **50**, 1427–1440.

Oden N.L. (1984) Assessing the significance of a spatial correlogram. *Geographical Analysis*, **16**, 1–16.

Odum E.P. (1971) *Fundamentals of Ecology*, 3rd edn. W.B. Saunders Company, Philadelphia.

Peres-Neto P.R., Jackson D.A. & Somers K.M. (2003) Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology*, **84**, 2347–2363.

Peres-Neto P.R., Jackson D.A. & Somers K.M. (2005) How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, **49**, 974–997.

Petersen R.C. Jr (1992) The RCE: a Riparian, Channel, and Environmental Inventory for small streams in the agricultural landscape. *Freshwater Biology*, **27**, 295–306.

Pinder L.C.V. (1986) Biology of freshwater Chironomidae. *Annual Review of Entomology*, **31**, 1–23.

Poff N.L. (1997) Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society*, **16**, 391–409.

Poff N.L., Olden J.D., Vieira N.K.M., Finn D.S., Simmons M.P. & Kondratieff B.C. (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.

Pyne M.I., Rader R.B. & Christensen W.F. (2007) Predicting local biological characteristics in streams: a comparison of landscape classifications. *Freshwater Biology*, **52**, 1302–1321.

Rabeni C.F. & Wang N. (2001) Bioassessment of streams using macroinvertebrates: are the Chironomidae necessary? *Environmental Monitoring and Assessment*, **71**, 177–185.

Rangel T.F.L.V.B., Diniz-Filho J.A.F. & Bini L.M. (2006) Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography*, **15**, 321–327.

Raunio J., Paavola R. & Muotka T. (2007) Effects of emergence phenology, taxa tolerances and taxonomic resolution on the use of the Chironomid Pupal Exuvial Technique in river biomonitoring. *Freshwater Biology*, **52**, 165–176.

Ribeiro M.C., Metzger J.P., Ponzoni F.J., Martensen A.C. & Hirota M.M. (2009) Brazilian Atlantic forest: how much is left and how the remaining forest is distributed? Implications for conservation. *Biological Conservation*, **142**, 1141–1153.

Richards C., Johnson L.B. & Host G.E. (1996) Landscape scale influences on stream habitats and biota. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 295–311.

Roque F.O. & Trivinho-Strixino S. (2007) Spatial distribution of chironomid larvae in low-order streams in southeastern Brazilian Atlantic Forest, a multiple scale approach. In: *Contributions to the Systematics and Ecology of Aquatic Diptera – A Tribute to Ole A. Sæther* (Ed. T. Andersen), pp. 255–264. The Caddis Press, Columbus, OH.

Roque F.O., Corbi J.J. & Trivinho-Strixino S. (2000) Considerações sobre a utilização de larvas de Chironomidae (Diptera) na avaliação da qualidade da água de córregos do Estado de São Paulo. In: *Ecotoxicologia: perspectivas para o século XXI*. (Eds E.L.G. Espíndola, C.M.R.B. Paschoal, O. Rocha, M.B.C. Bohrer & A.L.O. Neto), pp. 115–126. São Carlos: Rima Artes e Textos, São Paulo.

Roque F.O., Trivinho-Strixino S., Milan L. & Leite J.G. (2007) Chironomid species richness in low-order streams in the Brazilian Atlantic Forest: a first approximation through a Bayesian approach. *Journal of the North American Benthological Society*, **26**, 221–231.

Rosenberg D.M. (1992) Freshwater biomonitoring and Chironomidae. *Aquatic Ecology*, **26**, 101–122.

Silver P., Wooster D. & Palmer M.A. (2004) Chironomid responses to spatially structured, dynamic, streambed landscapes. *Journal of the North American Benthological Society*, **23**, 69–77.

Sokal R.R. & Oden N.L. (1978a) Spatial autocorrelation in biology. 1. Methodology. *Biological Journal of the Linnean Society*, **10**, 199–228.

Sokal R.R. & Oden N.L. (1978b) Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society*, **10**, 229–249.

Statzner B., Hildrew A.G. & Resh V.H. (2001) Species traits and environmental constraints: entomological research and the history of ecological theory. *Annual Review of Entomology*, **46**, 291–316.

Stephens P.A., Buskirk S.W. & del Rio C.M. (2007) Inference in ecology and evolution. *Trends in Ecology & Evolution*, **22**, 192–197.

Thienemann A. (1954) Ein drittes biozonotisches Grundprinzip. *Archiv fur Hydrobiologie*, **49**, 421–422.

Thioulouse J., Chessel D. & Champely S. (1995) Multivariate analysis of spatial patterns: a unified approach to local and global structures. *Environmental and Ecological Statistics*, **2**, 1–14.

Townsend C.R. & Hildrew A.G. (1994) Species traits in relation to a habitat templet for river systems. *Freshwater Biology*, **31**, 265–275.

Townsend C.R., Arbuckle C.J., Crowl T.A. & Scarsbrook M.R. (1997) The relationship between land use and physiochemistry, food resources and macroinvertebrate communities in tributaries of the Taieri River, New Zealand: a hierarchically scaled approach. *Freshwater Biology*, **37**, 177–192.

Townsend C.R., Dolédec S., Norris R.H., Peacock K.A. & Arbuckle C.A. (2003) The influence of scale and geography on relationships between stream community composition and landscape variables: description and prediction. *Freshwater Biology*, **48**, 768–785.

Umetsu F., Metzger J.P. & Pardini R. (2008) Importance of estimating matrix quality for modeling species distribution in complex tropical landscapes: a test with Atlantic forest small mammals. *Ecography*, **31**, 359–370.

UNFCCC (1992) *United Nations Framework Convention on Climate Change*. United Nations, New York.

Vinson M.R. & Hawkins C.P. (1998) Biodiversity of stream insects: variation at local, basin, and regional spatial scales. *Annual Review of Entomology*, **43**, 271–293.

Wagner H.H. & Fortin M.J. (2005) Spatial analysis of landscape: concepts and statistics. *Ecology*, **86**, 1975–1987.

Wiens J.A. (2002) Riverine landscapes: taking landscape ecology into the water. *Freshwater Biology*, **47**, 501–515.

Williams P.H., Margules C.R. & Hilbert D.W. (2002) Data requirements and data sources for biodiversity priority area selection. *Journal of Biosciences*, **27**, 327–338.