

## Evaluating subsampling approaches and macroinvertebrate taxonomic resolution for wetland bioassessment

RYAN S. KING<sup>1</sup> AND CURTIS J. RICHARDSON

Duke University Wetland Center, Nicholas School of the Environment and Earth Sciences, Box 90328,  
Durham, North Carolina 27708 USA

**Abstract.** Methods for wetland bioassessment using macroinvertebrates are not well developed. Two of the most controversial issues in stream bioassessment, subsampling and taxonomic resolution, have yet to be quantitatively addressed for wetlands. Using a multivariate approach, we evaluated the efficacy of family-, genus-, and species-level assemblage data in reflecting the environment and distinguishing impaired sites from the reference condition. We used 5 basic levels of subsampling (100-, 200-, and 300-organism fixed counts; 10% and 25% fixed areas), an integrated subsample requiring a minimum fixed count and fixed area (100&10%), and 100-count and 10%-area subsamples coupled with a supplementary large-rare (LR) search. Data were obtained from 1.5-m<sup>2</sup> composite samples collected from 126 plots along a 10-km-long eutrophication gradient in the Florida Everglades. Our results suggest that effectiveness of subsampling depended more upon the minimum number of individuals retained than minimum area or proportion of the sample picked. Fixed-area subsamples were generally less efficient than fixed counts, with 200- and 300-individual fixed counts resulting in significantly greater assemblage–environment relationships and much higher accuracy in detecting impairment than 10% fixed area, despite averaging similar numbers of individuals. The greatest improvement with increasing subsample size was observed between fixed counts of 100 and 200 individuals; detecting impairment, in particular, was not markedly improved with subsample sizes >200 individuals. Supplementing subsamples with a LR search resulted in only very slight improvements in assemblage–environment relationships, but was effective in improving prediction accuracy, particularly for family-level data. However, family-level assemblage–environment relationships and abilities to detect impairment were inferior to genus- and species-level data, regardless of subsample size. Species-level data performed best, primarily because of the large proportion (>20%) of total species belonging to Chironomidae. The potential importance of Chironomidae to wetland bioassessment was further revealed through an evaluation of a tiered-taxonomic approach, which showed that non-Chironomidae family-level data tiered with species-level Chironomidae data produced results very similar to those obtained using genus- or species-level data exclusively. Our results suggest that fixed counts  $\geq 200$  or integrated fixed-area/fixed-count approaches that consistently obtain a minimum of 200 individuals should be considered as minimum subsample sizes for wetlands. We additionally advocate LR searches and recommend genus- or species-level taxonomy, particularly for the Chironomidae.

**Key words:** biological assessment, biological monitoring, fixed-count subsampling, fixed-area subsampling, taxonomic sufficiency, large-rare search, multivariate approach, Chironomidae, Everglades.

Bioassessment has become a widely accepted technique for monitoring water quality and ecological health of aquatic systems (Rosenberg and Resh 1993). Attributes of macroinvertebrate assemblages, in particular, provide considerable information regarding levels and sources of impairment imposed by human influence (e.g., Karr and Chu 1997). Bioassessment is especially effective in lotic systems and is used to monitor environmental quality in streams throughout the world (e.g., Reynoldson et al. 1995, Zamora-

Muñoz and Alba-Tercedor 1996, Bailey et al. 1998, Barbour et al. 1999, Moss et al. 1999, Smith et al. 1999). Until recently, however, the use of biota to assess ecological condition of lentic habitats like wetlands had not received much attention (USEPA 1997a). In the USA, several states (e.g., Apfelbeck 1999, Gernes and Helgen 1999) along with the US Environmental Protection Agency (Danielson 1998) have recognized the need for biologically grounded wetland assessment methods. Most wetland assessment techniques in use today are based on functional indicators (e.g., Brinson and Rheinhardt 1996) that do not explicitly measure biological condition despite the mandate of Section 101(a) of the

<sup>1</sup> Present address: Smithsonian Environmental Research Center, Box 28, Edgewater, Maryland 21037 USA. E-mail: king@serc.si.edu

Clean Water Act to restore and maintain the chemical, physical, and biological integrity of the USA's waters, which include wetlands. Such inconsistency with federal legislation has led to dissatisfaction with current wetland assessment methods (Kusler and Niering 1998) and a call for the development of methods that incorporate biological components, like macroinvertebrate assemblages, into assessment protocols (USEPA 1997a, King et al. 2000).

Although interest in wetland bioassessment is currently high, no accepted assessment protocols for wetlands have been developed and published like those that exist for streams (e.g., Barbour et al. 1999). In addition to poorly known sensitivities to anthropogenic stressors (Batzer and Wissinger 1996) and relatively few established metrics of human influence (Lemly and King 2000), wetland macroinvertebrate assemblages present difficulties in sampling and sample processing that are less prevalent in lotic bioassessments. In particular, wetlands are usually moderately to heavily vegetated, so large quantities of coarse particulate organic matter (CPOM) are often present in macroinvertebrate samples collected with active sampling methods (Turner and Trexler 1997). Picking macroinvertebrates from CPOM-rich samples is laborious because small specimens are difficult to see and subsequently require time-consuming sorting procedures (FDEP 1996). Thus, cost-efficiency in sample processing looms as a potential threat to wetland bioassessment and likely will dictate a need for subsampling of whole samples.

Subsampling is actively used in stream bioassessments and is a standard operating procedure within the USEPA's Rapid Bioassessment Protocols (RBPs; Barbour et al. 1999). Nevertheless, subsampling is still a source of much contention among managers and biologists (Barbour and Gerritsen 1996, Courtemanch 1996, Vinson and Hawkins 1996, Grown et al. 1997, Walsh 1997, Somers et al. 1998, Sovell and Vondracek 1999, Doberstein et al. 2000). Much of this debate has arisen from the use of fixed-count subsampling, a method in which random cells in a sorting pan are picked until a target number of organisms (e.g., 100) is obtained. Fixed-count subsampling, particularly the 100-organism count recommended by the original RBPs of Plafkin et al. (1989), has been criticized; some claim that it provides unstable estimates of taxon richness and destroys any consistency

of areal sample size (Courtemanch 1996). Others have demonstrated that metric values calculated from fixed counts are highly variable and perhaps misleading when <300 individuals are obtained (Doberstein et al. 2000). However, still others have maintained that the cost-saving benefits of fixed-count processing outweigh the potential loss of information (Barbour and Gerritsen 1996), with some studies showing that even 100-organism bioassessments were able to distinguish sites of differing ecological impairment (Grown et al. 1997, Somers et al. 1998) and produce stable values for metrics when compared to larger counts (Sovell and Vondracek 1999).

Countering the fixed-count approach, Courtemanch (1996) recommended that fixed-area (fixed-fraction) subsampling or fixed-area whole samples should be used, although few studies have compared this approach to fixed counts. In one example, Walsh (1997) compared fixed counts of 200 and 300 and fixed areas of 10% and 25% and concluded that fixed-area subsampling was less efficient than fixed counts because of the high variability in the number of individuals identified among samples. However, Walsh (1997) additionally proposed a flexible subsampling regime that required a minimum fixed number of organisms as well as a fixed proportion of the sample be picked; this approach improved both precision and accuracy in the separation of stream sites into classes of human influence. Such an integrated approach is attractive for wetland bioassessment because of the typically variable densities of macroinvertebrates among wetlands (reviewed by Batzer et al. 1999). Samples from unproductive but high-quality wetlands may only have low numbers in the entire sample and would presumably be underrepresented by a fixed-area subsample or a small, quantitative whole sample. Similarly, wetlands with high macroinvertebrate densities but with a healthy, rich assemblage could also produce a misleading bioassessment if based on a small fixed count because a very small proportion of the entire sample would have been sorted (Courtemanch 1996, Walsh 1997).

An additional factor in subsampling is the inclusion of large, rare taxa. Neither fixed-count nor fixed-area subsampling adequately address the omission of these taxa (Courtemanch 1996, Vinson and Hawkins 1996, Walsh 1997), which may provide valuable signals of environmental status (Faith and Norris 1989, Cao et al. 1998).

Courtemanch (1996) and Vinson and Hawkins (1996) were the first to suggest that an additional large-rare (LR) organism search should accompany subsampling. The LR search, in which large-bodied, uncommon taxa are exhaustively removed from the sample either before or after subsampling, eliminates the bias against selection of these ecologically important taxa that are often missed (Vinson and Hawkins 1996). To date, only Vinson and Hawkins (1996) and Gerritsen et al. (2000) have added a LR search to subsamples, but they did not evaluate the ability of the LR search to improve accuracy in bioassessment. No study has compared the effects of fixed-count, fixed-area, integrated, or LR subsampling approaches in wetlands.

A second, separate issue in wetland bioassessment is taxonomic resolution. A wide range in views exists on taxonomic levels sufficient for accurate bioassessments of both freshwater and marine habitats (e.g., Bailey et al. 2001, Lenat and Resh 2001). Several studies have concluded that genus- or species-level taxonomy of macroinvertebrates contributed little additional signal of pollution when compared to the family level (e.g., Wright et al. 1995, Bowman and Bailey 1997, Marchant et al. 1997, Olsgard et al. 1998, Urkiaga-Alberdi et al. 1999, Hewlett 2000). Some have even suggested that higher taxonomic levels may be more appropriate than species-level data because species show a greater response to natural environmental variation, which can contribute to noise and mask the effects of human activity (Warwick 1993). However, the general trend in bioassessment is to identify taxa to the lowest possible unit, usually genus or species. In wetlands, species-level taxonomy may be time intensive and cost prohibitive because of the great abundance and diversity of Chironomidae and Oligochaeta (particularly Naididae), which can be difficult to identify and usually require slide-mounting and examination using compound microscopy. Conversely, coarse taxonomy could result in a considerable loss of information when considering families like the Chironomidae, which often represent a substantial proportion of the total number of species in wetlands (e.g., Wrubleski 1987, Batzer et al. 1999). A possible compromise is a tiered-taxonomic approach, as suggested by Bailey et al. (2001), in which taxa known to be sensitive primarily at a genus or species level are identified as such, but remaining taxa are iden-

tified only to family. However, no study has evaluated the merit of this approach.

The objective of our study was to quantitatively address the implications of subsampling and taxonomic resolution in wetland bioassessment. We de-emphasized the focus on taxon richness, which has been studied extensively elsewhere in this context (Barbour and Gerritsen 1996, Courtemanch 1996, Vinson and Hawkins 1996, Larsen and Herlihy 1998, Doberstein et al. 2000) and was shown to be a poor metric in the area chosen for this study (King 2001). Instead, we used a multivariate approach to address two questions: Do differing levels of subsampling and taxonomic resolution affect 1) the magnitude of assemblage–environment relationships, and 2) the ability to distinguish impaired sites from the reference condition? We also evaluated the importance of identifying Chironomidae beyond family using a tiered-taxonomic approach (Bailey et al. 2001), and used actual laboratory subsampling procedures rather than simulated subsamples generated by computers (Walsh 1997, Doberstein et al. 2000). Our goal was to recommend the least labor-intensive subsampling strategy and taxonomic level that best represented the wetland environment.

## Methods

### Study area and sampling design

We sampled in Water Conservation Area 2A (WCA-2A) in the northern Everglades (Fig. 1). WCA-2A is a 44,000-ha diked wetland landscape, with water-control structures regulating inflow and outflow of surface water. Inflow primarily occurs along the northern levee through 3 water-control structures (S10-A, C, and D) on the Hillsboro Canal, a conduit for outflow from Lake Okeechobee and nutrient-rich runoff from the Everglades Agricultural Area (EAA; Fig. 1). The Everglades was historically P-limited, so inflow from the Hillsboro Canal has induced a steep longitudinal eutrophication gradient in WCA-2A, primarily because of excessive inputs of P (SFWMD 1992). Three relatively distinct zones of impact have been described along this gradient (Fig. 1): 1) a highly impacted zone ~0 to 3 km downstream of the canal inflow structures, where surface water and soils are heavily enriched with P, and vegetation characteristic of the pristine Everglades has been mostly re-

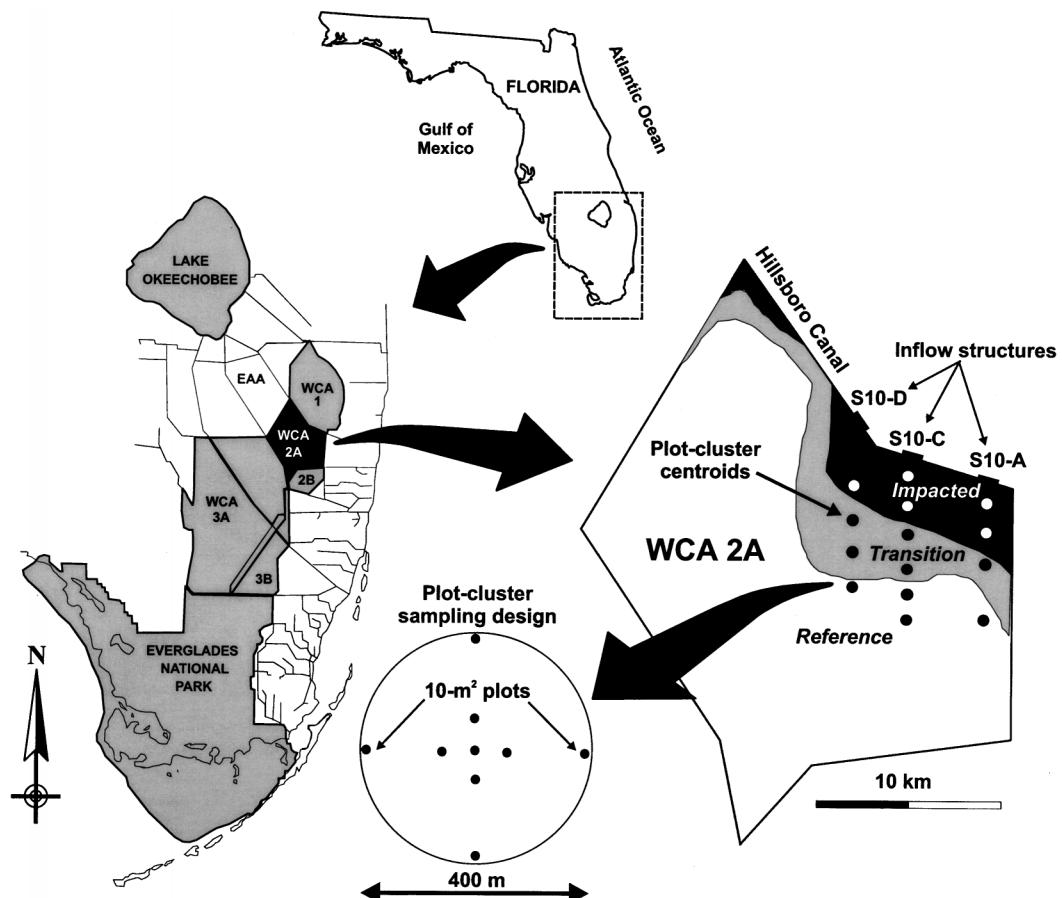


FIG. 1. Map of south Florida showing the location of Water Conservation Area 2A (WCA-2A); impacted, transition, and reference landscape zones; locations of S-10 water-control structures; centroids of sampling clusters; and plot-cluster sampling design. EAA = Everglades Agricultural Area.

placed by dense stands of cattail (*Typha domin-gensis* (Pers.)) and other invasive macrophytes, 2) a transition zone that ranges from 3 to 7 km from the canal, where P concentrations diminish but remain elevated, and vegetation is a mix of cattail, sawgrass (*Cladium jamaicense* Crantz), and infrequent open-water slough habitats, and 3) a relatively unimpacted reference zone >7 km from the canal that exhibits water and soil chemistry representative of the pristine Everglades, with vegetation structured as a mosaic of sawgrass stands laced with open-water sloughs (e.g., McCormick et al. 1996, Obeysekera and Rutchey 1997, Wu et al. 1997, SFWMD 2000, Vaithianathan and Richardson 1999).

Previous to our study, three 10-km long transects were established, each aligned with 1 of

the S-10 inflow structures and parallel to the eutrophication gradient (Fig. 1; Richardson et al. 2000). Six long-term sampling stations were marked along each transect, starting ~1.5 km from the canal and spaced at 1.5-km intervals. We selected 14 of these stations as centroids for our sampling, with all 6 selected from the C-transect, and random draw of 4 of the 6 from each of the A- and D-transects (Fig. 1). In aggregate, 5 stations were considered *impacted*, 5 *transition*, and 4 *reference* (Fig. 1).

We chose a landscape approach that would include pattern and scale of habitat as sources of variation rather than stratifying our study design by habitats or compositing across multiple habitats as commonly done in similar studies. We established a stratified-cluster sampling de-

TABLE 1. Physical and chemical characteristics of impacted, transition, and reference zones in Water Conservation Area (WCA)-2A of the Everglades. Mean ( $\pm 1$  SD) values were based on a single measurement collected at each of 126 plots (distance from canal, sediment chemistry, water depth<sup>a</sup>) in 1998, or multiple measurements of 14 plot-cluster centroids (water chemistry<sup>b</sup>) from 1995 to 1998.

Variable	n	Impacted	Transition	Reference
Distance from canal inflow structures (m)	126	2495 (869)	5541 (914)	9050 (924)
Total P (sediment, mg/kg)	126	1430 (172)	1203 (181)	578 (151)
Total Na (sediment, mg/kg)	126	3058 (160)	2900 (173)	2165 (113)
Total N (surface water, $\mu\text{g/L}$ )	495	2098 (765)	1867 (556)	1605 (514)
$\text{NO}_3\text{-N}-\text{NO}_2\text{-N}$ (surface water, $\mu\text{g/L}$ )	476	11.4 (30.9)	9.4 (12.5)	16.3 (29.5)
$\text{NH}_4\text{-N}$ (surface water, $\mu\text{g/L}$ )	475	67.7 (98.9)	31.0 (24.0)	39.9 (58.1)
Total P (surface water, $\mu\text{g/L}$ )	633	78.4 (74.3)	27.4 (26.4)	9.9 (5.4)
Soluble reactive P (surface water, $\mu\text{g/L}$ )	493	25.7 (29.5)	6.7 (5.7)	5.4 (5.3)
Cl (surface water, mg/L)	321	128.4 (67.1)	116.2 (38.7)	112.6 (39.5)
Total dissolved Ca (surface water, mg/L)	351	72.9 (17.9)	70.1 (18.7)	63.4 (18.8)
Total dissolved Na (surface water, mg/L)	351	98.8 (39.7)	89.5 (26.5)	78.5 (25.6)
Water depth (1981–1998 <sup>d</sup> , cm)	126	29.0 (8.7)	32.3 (9.6)	31.2 (11.4)

<sup>a</sup> 5 centroids and 45 plots for impacted and transition zones; 4 centroids and 36 plots for reference zone

<sup>b</sup> n = 3 grab samples per centroid per date

<sup>c</sup> Sediment samples were composites of three 2.75-cm diameter cores/plot, top 5 cm of sediment

<sup>d</sup> Based on a mean of 3 measurements/plot, corrected for water-level changes among locations during sampling period of 20 to 29 October 1998, and averaged for period of 1981 to 1998 (temporal extent of data set) using WCA-2A hydrological model developed for all 14 plot-cluster centroids by Romanowicz and Richardson (1997)

sign, which allowed us to consider fine, local-scale variation (as small as 50 m) as well as broad, landscape-scale variation (>10 km) in environmental and assemblage variables (Fortin et al. 1989, Urban 2000). A single plot at each of the 14 stations served as a plot-cluster centroid. Eight additional plots were marked in a constellation, with 4 plots placed at 50-m distances from centroids and 4 others at 200-m distances, each in the 4 cardinal directions. Thus, separation distances among plots within clusters ranged from 50 to 400 m, with a total of 9 plots/cluster and 126 plots across the landscape (Fig. 1). This design spanned virtually all local- and landscape-scale habitat types, water depths, and other environmental conditions within the reference condition and impaired zones, and thus incorporated multiple sources of variation that could influence wetland bioassessment. The constellation design and specific separation distances also allowed these plots to be found and resampled in the future.

We chose a plot area of 10 m<sup>2</sup>, large enough to integrate across microhabitat-specific effects and thus reduce noise, but not large enough to cross distinct patches of vegetation (Fortin et al. 1989). Plots were semicircular to allow sampling from the perimeter and minimize disturbance.

#### Field sampling

We measured a suite of physical and chemical variables at each plot or plot-cluster centroid to characterize the environment and corroborate impact-zone classifications (Table 1). Surface-water chemistry was measured as part of a long-term monitoring study at plot-cluster centroids on a quarterly basis from January 1995 to October 1998. All other variables were measured at every plot from 20 to 29 October 1998. Chemical analyses were done according to standard methods (APHA 1992). Details of sampling methods for environmental variables are presented in King (2001).

We based macroinvertebrate sampling on a modification of protocols used by the state of Florida (FDEP 1996; FDEP SOP #BA-7) and the USEPA (USEPA 1997b, Barbour et al. 1999) for bioassessment. We used a D-framed dip net (0.3-m wide, 500- $\mu\text{m}$  mesh) to collect 10 sweeps of 0.5-m length within each plot (total area 1.5 m<sup>2</sup>). Sweeps were allocated uniformly across the plot. Sweeps were collected by quickly jabbing the net frame onto the wetland bottom and forcefully sweeping across the surface sediments, vegetation, and metaphytic mats. A sweeping process was repeated rapidly twice

over the same area to capture dislodged organisms (USEPA 1997b, Maxted et al. 2000). Contents of all 10 sweeps were composited into a 500- $\mu\text{m}$  mesh sieve bucket, rinsed to remove fine particulates, placed in 4-L heavy duty storage bags, and preserved in 5% (v/v) buffered formalin stained with rose bengal. Macroinvertebrates were sampled by the same individual (RSK) to maintain consistency across all plots.

#### *Sample processing and subsampling*

Macroinvertebrate subsampling followed FDEP (1996; FDEP SOP #BA-8) and Barbour et al. (1999). Samples were rinsed and homogenized in a 500- $\mu\text{m}$  mesh sieve and large pieces of CPOM were discarded. Sieve contents were placed in a 20-cm wide  $\times$  45-cm long subsampling pan, and gently spread evenly throughout. The subsampling pan was gridded with numbered 5-cm square cells (36 cells total). Cell boundaries were grooved into the pan to facilitate removal of material from specific cells. Cells were selected for subsampling using a random numbers table. Material (detritus and invertebrates) was removed from cells primarily using forceps, but remaining material was collected using water from a rinse bottle and a wide-tipped plastic pipette. Small amounts of rinse water were pipetted from within the cell boundaries until all visible material had been removed.

Individual cell contents were transferred to a petri dish marked into  $\frac{1}{8}$  sections. One  $\frac{1}{8}$ -cell subsample of material was removed at a time, placed into a 2<sup>nd</sup> petri dish, and a small amount of water was added to suspend all contents. Macroinvertebrates subsequently were picked from the subsample using a stereomicroscope at 10 $\times$  magnification. The process was repeated until a target area or number of individuals was obtained. Subsample areal fraction was used to convert macroinvertebrate numerical abundance into density (no./ $\text{m}^2$ ) for each subsample based on the total sample area (1.5  $\text{m}^2$ ) (FDEP SOP #BA-8).

We selected 3 fixed-count (100, 200, and 300 individuals) and 2 fixed-area (10% and 25%) levels of subsampling for comparison. We chose fixed counts and fixed areas most commonly used in other bioassessment studies. We recognized that evaluations of fixed areas, by themselves, may be of limited use to biologists be-

cause few have agreed on a standard sample size to be used (e.g., Courtemanch 1996, Larsen and Herlihy 1998). However, evaluated in the context of average numbers of individuals per subsample and average proportions of the total sample sorted, these fixed-area subsamples were similar to the fixed-count subsamples and allowed for valid comparisons among approaches.

Upon reaching a specified number of individuals or area for a respective subsample level, specimens were placed in a vial containing 70% ethanol and total area, time required to complete, and number of individuals were noted. Larger subsamples (e.g., 300 individuals) were actually an accumulation of specimens stored in several vials, each representing a previous stopping point for other subsamples.

We implemented a supplementary LR search as defined by Courtemanch (1996) once 300 individuals and 25% of the total sample were subsampled. However, rather than pick all LR taxa from a sample before subsampling, as recommended by Courtemanch (1996), we picked remaining LR taxa after all subsampling was completed because to remove them prior to subsampling would have altered the composition of subsamples and prevented a valid assessment of the use of the LR search as a supplementary procedure. We defined *a priori* all LR taxa so that individuals included as part of a larger subsample (e.g., 25%) could be added into the pool of LR individuals for smaller subsamples that included the LR search (e.g., 100+LR). For example, a 100+LR subsample might only represent 5% of the total sample area for the fixed-count component. Subsequently, some LR taxa could be contained in the following subsamples of 200, 10%, 300, and 25% and would need to be counted in the final tally of additional LR organisms to be added to the 100+LR subsample. Thus, any LR taxa in the 200, 10%, 300, and 25% subsamples would have to be added to the remaining LR search for the 100+LR subsample to be accurate and valid. We classified large mollusks, hemipterans, hirudineans, coleopterans, decapods, all anisopteran odonates, and a few miscellaneous large taxa as LR taxa. We calculated densities for LR taxa based on the total number of individuals/sample, not the fractional area of individual subsamples in which LR taxa were supplemented.

We identified all macroinvertebrates to the

lowest possible taxon, usually species, except for copepods (order) and nematodes (phylum). Morphospecies designations were used when taxa were obviously unique but not identifiable as a species (Oliver and Beattie 1996). All species identifications were verified by expert taxonomists (see Acknowledgements).

#### *Data analysis*

**Assembly of data sets.**—We assembled macroinvertebrate data sets using the 5 basic levels of subsampling (100, 200, 300, 10%, and 25%), an integrated subsample requiring a minimum fixed count and fixed area in the same subsample (100&10%), and a fixed count (100) and fixed area (10%) subsample supplemented with the LR search. Data sets also were assembled using 3 levels of taxonomic resolution (family, genus, and species) for each subsampling level, thus totaling 24 sets. Each level of taxonomy connoted the lowest level achieved for most identifications. Data were densities ( $\text{no}/\text{m}^2$ ) of each taxon for each of the 126 plots sampled.

We evaluated the importance of identifying Chironomidae beyond family level by constructing 3 tiered data sets: 1) non-Chironomidae family-level data tiered with species-level Chironomidae data, 2) non-Chironomidae genus-level data tiered with family-level Chironomidae data, and 3) non-Chironomidae species-level data tiered with family-level Chironomidae data. These tiered sets were compared with family-, genus-, and species-level data sets. A representative midsized subsample (200 count+LR) was used for these comparisons.

**Effects of subsampling and taxonomic resolution on assemblage–environment relationships.**—We used a continuous-variable, gradient approach to relate assemblage composition to the environment. Here, we compared the magnitude of assemblage–environment relationships among subsampling approaches and levels of taxonomic resolution using the multivariate Mantel test (Mantel 1967), which measures the correlation between distance matrices. Increasing magnitude in Mantel  $r$ , the test statistic, reflects a stronger correlation. Mantel  $r$  typically ranges from 0.1 to 0.3 for assemblage–environment relationships that are ecologically significant and infrequently exceeds 0.5 because the analysis is based on the full rather than reduced dimensionality (e.g., ordination-axis scores) in the as-

semblage data (e.g., Leduc et al. 1992, Sanderson et al. 1995, Foster et al. 1999). We selected distance from canal inflow structures (hereafter, Canal) as a predictor of macroinvertebrate assemblage composition because 1) it was a surrogate for a wide range of biogeochemical, hydrological, and habitat-structural variables that substantially change along this eutrophication gradient (Table 1), and 2) it was the best predictor of biological changes in this study area (King 2001). Canal (m) was converted to a distance matrix using Euclidean distance, whereas assemblage matrices used Bray–Curtis dissimilarity as the distance metric (Legendre and Legendre 1998). Bray–Curtis dissimilarity was selected because it is one of the most robust and ecologically interpretable distance metrics available (e.g., Faith et al. 1987, Legendre and Anderson 1999, Hawkins and Norris 2000). All macroinvertebrate density data were  $\log_{10}(x + 1)$  transformed prior to conversion to distance matrices to give greater weight to less-abundant taxa (Legendre and Legendre 1998).

We estimated 95% confidence intervals (CI) for each test statistic using bootstrapping, a resampling method (Manly 1997), rather than qualitatively comparing the magnitude of Mantel  $r$  statistics among subsamples and taxonomic levels. We resampled (with replacement) distance matrices at a level of 90%, with 1000 resamples (Manly 1997). Mantel  $r$  statistics were considered significantly different if 95% CIs did not overlap (Manly 1997, Johnson 1999). We also evaluated whether Mantel  $r$  statistics were significantly different from 0 ( $p \leq 0.05$ ) using 10,000 random permutations (Manly 1997); however, this test was merely an antecedent to the more relevant comparison of uncertainty (95% CI) among assemblage–environment correlations (Suter 1996, Germano 1999, Johnson 1999). Mantel tests and bootstrapping were done using S-Plus 5.0 for Unix (Mathsoft, Inc., Seattle, Washington, USA).

**Effects of subsampling and taxonomic resolution on distinguishing impaired locations from the reference condition.**—We used a discrete-variable, reference-site approach to assess accuracy among subsampling approaches and levels of taxonomy in distinguishing impaired plots from the reference condition. We used the BEthnic Assessment of SedimenT (BEAST) multivariate predictive model (Reynoldson et al. 1995, 1997, 2001) to contrast assemblage composition at in-

dividual test sites with the reference condition. In the BEAST approach, several steps are typically used to assign reference- and test-site data into appropriate classes (Reynoldson et al. 2001). The last step in the model is construction of a Gaussian bivariate probability ellipse (Altman 1978, Owen and Chmielewski 1985) around reference sites in nonmetric or hybrid multidimensional scaling ordination space to assess whether assemblage composition at a test site lies outside of what would be expected within the reference condition. However, we only had 1 reference stratum and a substantial amount of published information from our study area to help assign plots as either reference or impaired (transition and impacted) *a priori*. Thus, we only used the last step in the BEAST approach, the test-site assessment, to evaluate the accuracy of subsampling approaches and taxonomic levels in correctly distinguishing impaired plots from the reference condition. However, rather than including test sites along with reference sites in the same ordination (Reynoldson et al. 1997), we used predictive-mode nonmetric multidimensional scaling (nMDS) to predict the location of impaired sites in ordination space using a calibration ordination based solely on the reference condition for each subsampling approach and taxonomic level (McCune et al. 1997a, 1997b, B. McCune and M. J. Mefford. 1999. Multivariate analysis of ecological data, version 4.09, user's guide, MjM Software, Gleneden Beach, Oregon, USA). The advantage of using predictive-mode nMDS was that locations of transition and impacted plots were independently projected into calibration ordinations without affecting the position of the reference plots in ordination space.

We used Bray–Curtis dissimilarity as the distance metric, and  $\log_{10}(x + 1)$  transformed all abundances. We initially evaluated the most appropriate dimensionality for the calibration data sets by examining stress coefficients (a metric of agreement between compositional dissimilarities and distance among points in nMDS space) for 1 to 3 dimensions. We selected a 2-dimensional solution for all calibration ordinations because an additional dimension only slightly reduced stress (McCune et al. 1997a, 1997b). We used 100 iterations for each nMDS run, and random starting coordinates. We used 10 runs for each data set to ensure that the solution was stable and represented a configuration with the

best possible fit (McCune and Mefford 1999). We then used calibration configurations to predict the locations of each impaired plot in the same nMDS ordination space. Last, we assessed accuracy in detecting impairment using a 95% probability ellipse constructed around reference plots ( $n = 36$ ); we considered transition ( $n = 45$ ) and impacted ( $n = 45$ ) plots outside of this probability ellipse to be accurately assessed as impaired (Reynoldson et al. 1997, 2001). For greater precision, we limited extrapolation of predicted locations to 50% of the lengths of nMDS Axes 1 and 2 (McCune and Mefford 1999)—these extrapolation limits were well beyond the boundaries of probability ellipses encompassing the reference data. Calibration and predictive-mode nMDS ordinations were done using PC-ORD 4.09 (MjM Software, Gleneden Beach, Oregon, USA), whereas probability ellipses were calculated using Statistica 5.5 (Statsoft, Inc., Tulsa, Oklahoma, USA).

## Results

### *Subsample characteristics and taxonomic structure*

Numbers of individuals showed tremendous variation among subsamples using the fixed-area approach (Table 2). Although 10% area averaged over twice the number of individuals as the 100 count, it produced as few as 23 individuals in 1 subsample, and had <100 individuals 27% of the time. Similarly, the 25% area averaged nearly twice the number of individuals as the 300 count despite averaging a similar % of the total sample subsampled. The LR search added an average of as many as 19 individuals to subsamples.

Sorting times mirrored the % of total sample subsampled rather than number of individuals picked (Table 2). LR searches added an average of 23 min (100+LR) to sorting time.

Over 78,000 individuals were identified across all 126 plots. A total of 93 families, 181 genera, and 252 species were identified. Coleopterans, dipterans, gastropods, odonates, and oligochaetes were the most diverse of the major taxonomic groups, and contributed most to the differences among the number of families, genera, and species identified. Chironomidae was the most diverse family, represented by 30 and 51 genera and species, respectively.

The LR search added as many as 4 families, 9 genera, and 16 species, cumulatively, to any 1 lev-

TABLE 2. Comparison of selected properties of the 8 subsampling approaches evaluated for wetland bioassessment. LR = large-rare organisms. – = not applicable.

Subsample	% Of total sample Mean ( $\pm 1$ SD)	No. of individuals		Sorting time (min) Mean ( $\pm 1$ SD)
		Mean ( $\pm 1$ SD)	Range	
100 count	7.7 (6.4)	102.7 (4.7)	92–118	94.2 (55.3)
200 count	15.2 (12.7)	203.7 (7.3)	191–224	156.1 (100.0)
300 count	22.6 (18.3)	304.6 (10.9)	283–326	206.7 (117.8)
10% area	10.0 (–)	230.0 (178.3)	23–1036	94.7 (51.5)
25% area	25.0 (–)	573.1 (440.3)	62–2558	250.7 (136.2)
100 count+LR <sup>a</sup>	– (–)	121.3 (14.3)	100–214	117.0 (54.0)
10% area+LR	– (–)	247.7 (177.7)	36–1044	117.2 (51.7)
100 count and 10% area	11.6 (4.3)	238.2 (170.9)	97–1036	118.0 (64.5)

<sup>a</sup> Subsamples containing the LR component were picked completely for all LR taxa (100% of sample area) in addition to the fixed-count or fixed-area component

el of subsampling (100 vs 100+LR). Frequencies of occurrence of many LR taxa increased as much as a factor of 10 by implementing the LR search, with the 100-count subsample performing the poorest of all in capturing LR taxa (Table 3).

#### Assemblage–environment relationships

Mantel *r* statistics were significantly different from 0 ( $p \leq 0.0001$ ), regardless of subsample or taxonomic level. However, the magnitude of these assemblage–environment correlations varied significantly (95% CI) among subsamples and taxonomic levels (Fig. 2). In particular, the greatest increase in assemblage–environment relationships with increasing subsample size was observed between 100 and 200 counts—100 performed significantly worse than 200, whereas 200 was not different from 300, regardless of taxonomic level.

The magnitude of Mantel *r* values suggested that fixed areas generally were less efficient than fixed counts. Fixed counts of 200 and 300 indi-

viduals produced significantly greater Mantel *r* values than 10% area, despite averaging similar numbers of individuals (Fig. 2). Similarly, 25% area assemblage–environment relationships were not significantly greater than the less labor-intensive 300 count at the genus and species levels.

Evaluation of the integrated subsampling approach (100 count and 10% area) showed significant increases over the respective basic fixed-area and fixed-count components, but these differences depended upon level of taxonomy. At genus and species levels, 100&10% assemblage–environment relationships were significantly greater than 100-count or 10%-area subsamples (Fig. 2). However, at the family level, the 100&10% correlation was only significantly greater than that of the 100 count.

Adding the LR search to 100-count and 10%-area subsamples resulted in very slight increases in the strength of assemblage–environment relationships for all 3 levels of taxonomy (Fig. 2). LR taxa significantly increased the Mantel *r* value for 100-count data at the family level.

TABLE 3. Comparison of the frequency of occurrence (%) of 5 representative large-rare (LR) taxa among fixed-count and fixed-area subsampling methods with respect to the whole-sample frequency, as estimated by the LR search ( $n = 126$ ).

Taxon	Subsample					
	100	200	300	10%	25%	Whole
<i>Belostoma testaceum</i>	0.8	2.4	4.0	2.4	4.0	9.5
<i>Celithemis eponina</i>	2.4	4.0	4.0	62.4	5.6	11.1
<i>Palaemonetes paludosus</i>	30.2	41.3	47.6	38.1	57.1	65.9
<i>Pelocoris femoratus</i>	8.7	16.7	25.4	10.3	31.0	57.9
<i>Procambarus fallax</i>	7.1	16.7	22.2	13.5	27.0	51.6

TABLE 4. Comparison of the ability of differing levels of subsampling and taxonomic resolution to distinguish impaired plots from the reference condition. Accuracy (no. of plots (% in parentheses)) was estimated using predicted scores of impaired plots in nonmetric multidimensional scaling (nMDS) ordination space based on reference-plot data ( $n = 36$ ). Impaired plots falling inside 95% confidence ellipses constructed around reference plots were considered misclassified. Results are shown by impact strata (transition, impacted) and in aggregate (total). Subsamples are listed by increasing mean number of individuals.

Subsample	Stress <sup>a</sup>	Accuracy		
		Transition ( $n = 45$ )	Impacted ( $n = 45$ )	Total ( $n = 90$ )
<i>Family</i>				
100	0.2051	12 (27)	30 (67)	42 (47)
100+LR	0.1960	13 (29)	34 (76)	47 (52)
200	0.1805	17 (38)	40 (89)	57 (63)
10%	0.2067	1 (2)	32 (71)	33 (37)
100&10%	0.1972	13 (29)	39 (87)	52 (58)
10%+LR	0.1957	8 (18)	40 (89)	48 (53)
300	0.1747	22 (49)	40 (89)	62 (69)
25%	0.1880	25 (56)	38 (84)	63 (70)
<i>Genus</i>				
100	0.1816	29 (64)	45 (100)	74 (82)
100+LR	0.1732	31 (69)	45 (100)	76 (84)
200	0.1444	36 (80)	45 (100)	81 (90)
10%	0.1603	17 (38)	43 (96)	60 (67)
100&10%	0.1576	30 (67)	45 (100)	75 (83)
10%+LR	0.1586	19 (42)	45 (100)	64 (71)
300	0.1320	34 (76)	45 (100)	79 (88)
25%	0.1355	36 (80)	45 (100)	81 (90)
<i>Species</i>				
100	0.1766	30 (67)	45 (100)	75 (83)
100+LR	0.1756	31 (69)	45 (100)	76 (84)
200	0.1448	37 (82)	45 (100)	82 (91)
10%	0.1537	22 (49)	43 (96)	65 (72)
100&10%	0.1533	30 (67)	45 (100)	75 (83)
10%+LR	0.1520	24 (53)	45 (100)	69 (77)
300	0.1313	38 (84)	45 (100)	83 (92)
25%	0.1361	41 (91)	45 (100)	86 (96)

<sup>a</sup> Metric of agreement between compositional dissimilarities and distance among points in 2-dimensional nMDS ordination of reference-plot data

Differences in the magnitudes of assemblage-environment correlations were more apparent among levels of taxonomic resolution than among subsampling approaches. The family level was significantly inferior to both genus- and species-level data, regardless of subsampling approach (Fig. 2). Species-level data

showed significantly stronger relationships to the environment than the genus level, although 95% CIs were only marginally separated within each level of subsampling.

#### Accuracy in detecting impairment

Results from the BEAST approach indicated that accuracy in detecting impairment varied among levels of subsampling and taxonomy (Table 4, Fig. 3). For all 3 taxonomic levels, the largest increase in accuracy occurred between the 100 and 200 fixed-count subsamples. Overall accuracy for species-level data increased from 83% to 91%, whereas accuracy in detecting subtle, transitional impairment increased from 67% to 82% (Table 4). Increasing fixed counts from 200 to 300 only improved total accuracy from 91% to 92% at the species level, whereas total accuracy decreased from 90% to 88% at the genus level. Family-level total accuracy went from 47% to 63% from 100 to 200 counts and only increased to 69% at the 300 count. All genus- and species-level fixed counts were 100% accurate in distinguishing impacted-zone plots from the reference condition (Fig. 3).

The 10% fixed-area subsample exhibited the poorest accuracy of all procedures. Its overall accuracies ranged from as low as 37% for family-level data to a maximum of 72% at the species level (Table 4). The 10%-area subsample was the only subsample to misclassify impacted plots at the genus and species levels, and was ineffective in distinguishing transition plots from the reference condition—only 1 out of 45 transition plots was diagnosed as impaired using family-level data. The 25%-area subsample was equivalent to or slightly more accurate than 200 and 300 counts, although improvements were modest considering that 25% area averaged nearly 300 and 400 more organisms than the 200 and 300 counts, respectively.

The 100&10% area subsamples resulted in only slight to no improvement in accuracy over the 100 fixed count at genus and species levels, but notably increased overall accuracy over the 10% fixed-area subsample (e.g., from 67% to 83% for genus-level data) (Table 4). The LR search also contributed more to improving accuracy of the 10% fixed-area than of the 100-count subsample (e.g., an overall increase from 37% to 53% for family-level data).

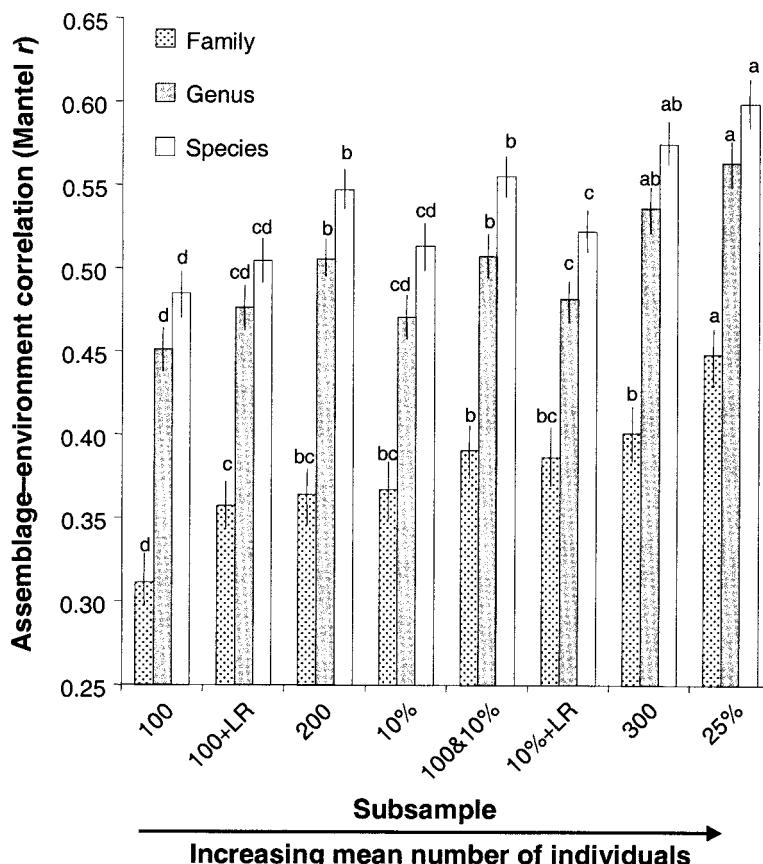


FIG. 2. Assemblage–environment correlations for each subsampling approach and level of taxonomic resolution, as estimated using Mantel tests. Significant differences in the magnitude of Mantel  $r$  values (bootstrapped 95% CI, error bars) among subsamples within taxonomic levels are indicated by the lower-case letters; Mantel  $r$  values with the same letters were not different. Among taxonomic levels, Mantel  $r$  values with overlapping 95% CI were not significantly different (all Mantel  $r$  values differed among the 3 levels of taxonomy within each level of subsampling).

Taxonomic resolution had a greater effect on accuracy than did subsampling (Table 4, Fig. 3). Family-level data were only 67% to 89% accurate in distinguishing impacted-zone plots from the reference condition (Table 4). Family-level resolution was also poor in detecting transitional impairment, with accuracies as low as 2% and no greater than 56%. In contrast, genus and species levels were able to detect impacted-zone plots with almost 100% accuracy, even using 100-count subsamples. Overall, accuracy of genus and species levels was similar, with species-level data slightly better in distinguishing transition-zone plots from the reference condition.

#### Tiered taxonomy using Chironomidae

Chironomidae may have been largely responsible for the observed disparity among taxonomic levels. Results from the tiered-taxonomic analysis revealed that tiering family-level data with species-level Chironomidae data yielded assemblage–environment correlations that were not different from those obtained by identifying all taxa to genus or species (Fig. 4). Conversely, leaving Chironomidae identifications at just the family level but identifying other taxa to genus or species produced significantly worse assemblage–environment correlations than that of genus, species, and tiered family/Chironomidae-

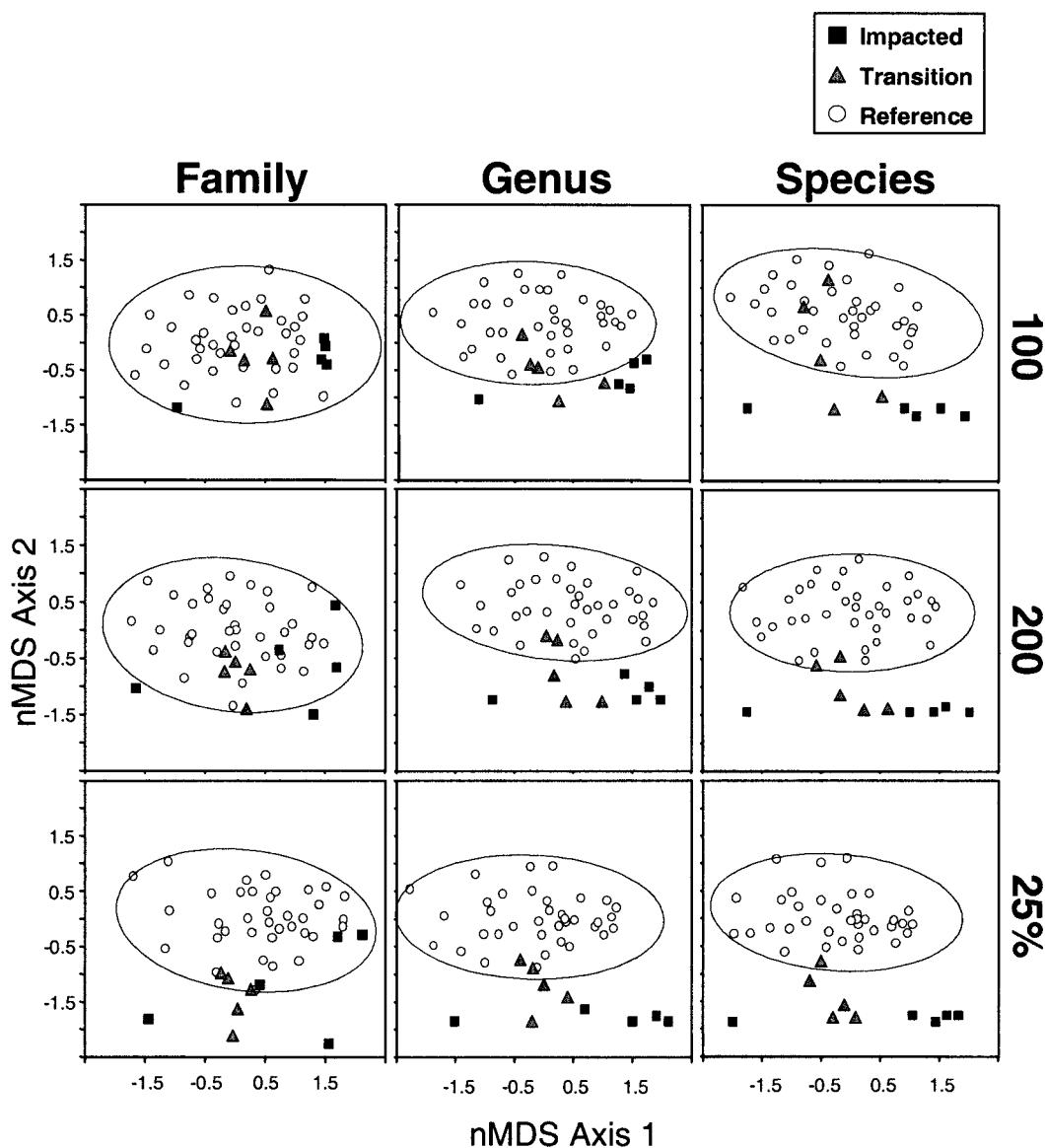


FIG. 3. Comparison of the abilities of 100-count, 200-count, and 25%-area subsamples and family, genus, and species levels of taxonomic resolution to distinguish impaired plots from the reference condition. Shown are nonmetric multidimensional scaling (nMDS) ordinations of macroinvertebrate assemblage composition, with 95% probability ellipses constructed around reference plots. For clarity, predicted locations of only 5 transition and 5 impacted plots are projected onto each ordination; these plots were the same for each ordination.

to-species data (Fig. 4). Total accuracy of the tiered family/Chironomidae-to-species data was 88%, compared to 90% and 91% for genus and species-level data, respectively, whereas Chironomidae-to-family data tiered with genus and species data yielded accuracies of only 68% and 78%, respectively (Table 5).

## Discussion

### Evaluating subsampling approaches

Our results showed that subsampling significantly influenced the ability of bioassessments to represent the environment and detect impairment in wetlands. Number of individuals,

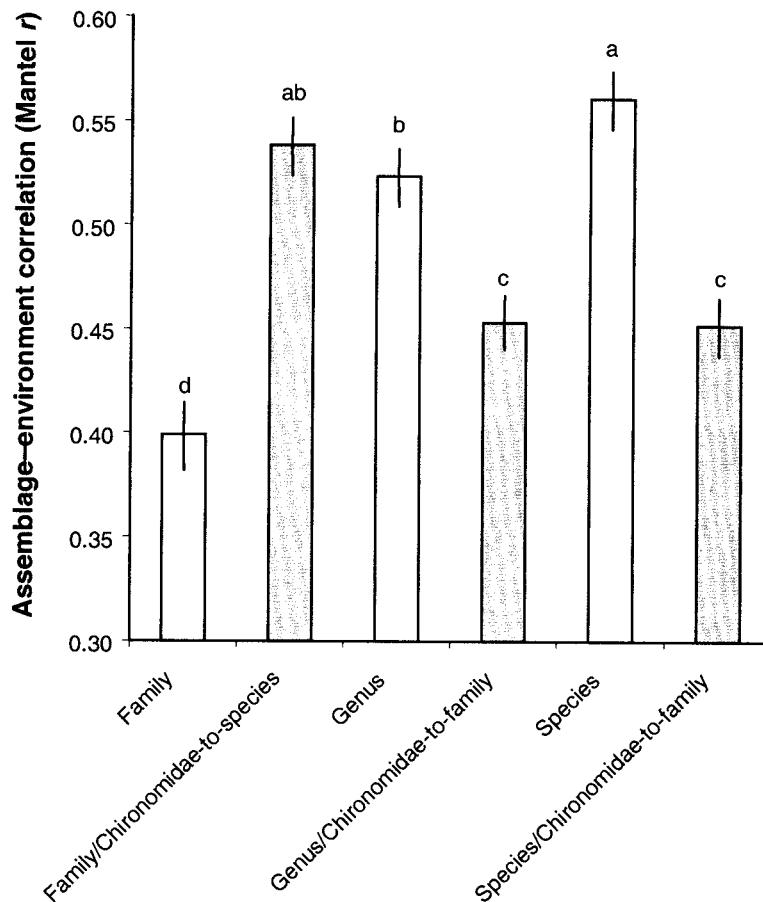


FIG. 4. Assemblage–environment correlations for tiered-taxonomic levels using Chironomidae and family-, genus- and species-level taxonomy, as estimated using Mantel tests. Significant differences in the magnitude of Mantel *r* values (bootstrapped 95% CI, error bars) among taxonomic levels are indicated by the lower-case letters; Mantel *r* values with the same letters are not different.

subsample area, and LR taxa all proved to be important considerations in subsampling. Although no subsampling approach was clearly best, there was variation in the gains afforded by each approach and respective level of effort. The 100 count performed worse than all other subsamples except for the 10% area. Assemblage–environment relationships were significantly weaker for 100-count data than other fixed-count subsamples. Perhaps most disturbing was the relatively poor accuracy of the 100 count in detecting subtle, transition-zone impairment. This result has serious implications for wetland bioassessment because 100-organism fixed counts have become widely used among state (e.g., FDEP 1996), regional (e.g.,

Maxted et al. 2000), and federal (e.g., Plafkin et al. 1989) protocols for streams in the USA. Other studies have concluded that the 100 count can be equally informative as larger counts for lakes and streams (Barbour and Gerritsen 1996, Grown et al. 1997, Somers et al. 1998, Sovell and Vondracek 1999), but these studies only considered the ability of subsamples to discriminate differences in *mean* values of specific metrics among  $\geq 2$  levels of impairment. None of these studies statistically compared the magnitude of a test statistic or evaluated accuracy among subsamples. Rather, statistical significance of analyses of variance (ANOVAs) or analyses of similarities (ANOSIMs) was used in these studies as the evaluation tool. The problem

TABLE 5. Comparison of the ability of tiered-taxonomic levels and family, genus, and species levels to distinguish impaired plots from the reference condition. Accuracy (no. of plots (% in parentheses)) was estimated using predicted scores of impaired plots in nonmetric multidimensional scaling (nMDS) ordination space based on reference-plot data ( $n = 36$ ). Impaired plots falling inside 95% confidence ellipses constructed around reference plots were considered misclassified. Results are shown by impact strata (transition, impacted) and in aggregate (total). Comparisons were made using the 200 + LR (large-rare) subsample.

Taxonomic level	Stress <sup>a</sup>	Accuracy		
		Transition ( $n = 45$ )	Impacted ( $n = 45$ )	Total ( $n = 90$ )
Family	0.1791	18 (40)	40 (89)	58 (64)
Family/Chironomidae-to-species	0.1568	34 (76)	45 (100)	79 (88)
Genus	0.1429	36 (80)	45 (100)	81 (90)
Genus/Chironomidae-to-family	0.1593	19 (42)	42 (93)	61 (68)
Species	0.1436	37 (82)	45 (100)	82 (91)
Species/Chironomidae-to-family	0.1651	27 (60)	43 (96)	70 (78)

<sup>a</sup> Metric of agreement between compositional dissimilarities and distance among points in 2-dimensional nMDS ordination of reference-plot data

with this approach is that considering only the mean among replicates from different levels of impairment does not adequately assess how increasing size of subsamples improves the strength of relationships, increases the accuracy in detecting impairment, or estimates the risk of misclassifying *individual* sites (e.g., Suter 1996, Doberstein et al. 2000). Particularly with large sample sizes, means can be minimally different with much overlap in values among classes of impairment yet still produce statistically significant differences (Germano 1999, Johnson 1999). Accuracy in detecting impairment, as estimated using the multivariate BEAST approach, is a more appropriate measure of subsampling effectiveness because each plot or site is considered individually, as in real bioassessments (Reynoldson et al. 1997). One single sample, often a single- or multihabitat composite, is typically collected from a site and solely is used to assign a quality rating to that site, rather than using means of multiple replicates within a site (e.g., Barbour et al. 1999, Maxted et al. 2000). Thus, comparing means of replicates from within sites or among sites is misleading and inconsistent with how most bioassessment programs generate data and assess environmental condition.

Barbour et al. (1999) recently modified the original RBP of Plafkin et al. (1989) by increasing the recommended fixed-count subsample from 100 to 200 individuals. Our data support that modification. The greatest improvement in assemblage–environment relationships with in-

cremental increases in subsample size was from a fixed count of 100 to 200 individuals. Although more costly, this increase in effort seems justified given a significant increase in the ability of 200 counts to reflect the environment. This 100-individual increase also markedly improved accuracy in detecting transitional impairment for genus- and species-level data. Alternatively, increasing fixed counts to 300 did not significantly improve assemblage–environment correlations when compared to 200. Moreover, 300 counts scarcely improved accuracy in detecting impairment over 200 counts. Thus, our data are not entirely conclusive regarding whether counts >200 are warranted given the proportional increases in costs. However, our data suggest that counts of <200 individuals may degrade the accuracy of bioassessment in wetlands.

The evaluation of fixed-area subsampling suggested that it may be less efficient than the fixed-count approach. We anticipated that fixed-area subsamples would have an advantage over fixed counts for samples with disproportionately high densities of one or a few taxa because this situation would result in a small proportion of the total sample picked and an under-representation of the rest of the assemblage. However, we also recognized that fixed counts always produce a fixed number of individuals regardless of organism densities, which could be an advantage over fixed areas if densities are low and a fixed area produces very few individuals (Walsh 1997). Our results suggested that sub-

sampling effectiveness depended more upon obtaining a minimum number of individuals than a minimum, standardized sample area. Although averaging more individuals than either 100 or 200 counts, the 10% area performed poorly because of a large proportion of subsamples that did not achieve either of these counts. Such low counts yielded insufficient estimates of composition and densities for certain plots, and subsequently resulted in poor assemblage-environment relationships and very low accuracies in comparison to 100 and 200 counts. The detrimental effect of insufficient counts was obvious by the sizable improvement in accuracies resulting from integrating a 100-count subsample with 10% area. Thus, it appears that for fixed-area subsampling alone to be reliable, an area of sufficient size to attain a minimum number of individuals should be selected; otherwise, a smaller fixed area needs to be integrated with a minimum fixed count, as recommended by Walsh (1997).

The need for a sufficient fixed area also has implications to sampling programs that do not use subsampling but pick samples exhaustively. Our data suggest that unless a sufficient sampling area is retained to produce a minimum threshold number of individuals, whole-sample processing has no advantage over subsampling and may even be less reliable because of scale dependence of samplers and intrinsic patchiness of macroinvertebrate assemblages. A large composite sample (e.g., 1.5 m<sup>2</sup>) that is homogenized, and then subsampled for a fixed area (e.g., 25% or ~0.4 m<sup>2</sup>) is more likely to reflect an actual wetland assemblage than a quantitatively taken whole sample of identical size (0.4 m<sup>2</sup>) because the latter has a greater probability of entirely missing patches of representative taxa. This situation is particularly true for LR taxa that have low densities but may be ecologically important (e.g., Courtemanch 1996, Vinson and Hawkins 1996, Walsh 1997, Cao et al. 1998). Ironically, Courtemanch (1996), who advocated a LR search if subsampling is used, recommended a quantitative whole sample that only yielded a median of 250 individuals. In our study, a median of 250 individuals would be obtained with ~15% of the total sample, or ~0.2 m<sup>2</sup>. All other factors being equal, this sample size would have missed the 5 representative LR taxa (Table 4) up to 6× more often than the 1.5-m<sup>2</sup> samples that were subsampled and supplemented with a LR

search. This estimate is probably low because frequencies of occurrence actually were calculated using a random subsample from a larger, homogenized sample rather than a small, more patch-sensitive whole sample as used by Courtemanch (1996). Clearly, collecting larger composite samples, either single- or multihabitat, has a distinct advantage over much smaller quantitative ones if characterizing assemblage composition and, particularly, detecting presence of LR taxa is of interest.

The importance of LR taxa to bioassessment has not received much quantitative investigation but has been the focus of discussion (Courtemanch 1996, Vinson and Hawkins 1996, Walsh 1997, Cao et al. 1998, Marchant 1999). Faith and Norris (1989) did not consider subsampling approaches or a LR search but showed that rare taxa (not necessarily just LR ones) positively contributed to assemblage-environment relationships in streams of Australia. Cao et al. (1998) concluded that rare taxa were not only important in bioassessment, but that their removal from data sets led to an unacceptable loss of ecological information. Others, however, have been skeptical of this assertion (Marchant 1999). Although we did not assess how removing rare taxa from subsamples affected bioassessment, we did evaluate the contribution of LR taxa to subsamples that otherwise would have misrepresented their abundance or completely missed them. Our results support the claims of Cao et al. (1998) because we were able to show that LR taxa had a positive effect on the ability of subsamples to reflect the environment and detect impairment. All assemblage-environment relationships were slightly stronger when LR taxa were supplemented compared to the same subsample without the LR search, although the relationships were only significantly stronger at a family level for the 100+LR subsample. LR taxa also improved accuracies in detecting impairment, particularly at the family level. Considering that sorting time for the LR search averaged, at most, 23 min and most taxa were easily identified because of their large size, the LR search seems to be a relatively simple, cost-effective means of improving accuracy in bioassessment.

#### *Importance of taxonomic resolution*

Our results suggest that taxonomic resolution may be a more important consideration than

subsampling in the development of a wetland bioassessment protocol. Family-level data yielded much weaker assemblage–environment relationships and lower accuracies in detecting impairment than those of genus- or species-level data. This result was somewhat surprising given that the general consensus among previous studies has been that coarse levels of taxonomy, such as family level, can be as sensitive as species data in detecting environmental impacts (Ferraro and Cole 1990, 1992, Wright et al. 1995, Marchant et al. 1997, Olsgard et al. 1998, Urkiaga-Alberdi et al. 1999, Bailey et al. 2001) and tend to be highly correlated to finer-resolution data (Bowman and Bailey 1997, Hewlett 2000). In some cases, family-level data have become the standard for bioassessment (e.g., Chessman 1995). However, a few studies have demonstrated that family-level data can indeed reduce accuracy in bioassessment. For example, Lenat and Resh (2001) showed that family-level data misclassified 40% of stream sites that should have received “excellent” water-quality designations, consequently underestimating their value. In addition, their study showed that family data overestimated stream quality at least 25% of the time, leading to inappropriate classification of “poor” quality sites as “fair” or “good”. Similarly, Hawkins et al. (2000) concluded that family-level predictive models were not sufficient to detect biological impairment in streams of California. Results from our study support the position that genus- or species-level data can improve the quality of bioassessments. Although our study is the 1<sup>st</sup> to evaluate taxonomic resolution for wetlands, our results indicate that family-level data may not be adequate for wetland bioassessment.

Fundamental differences in taxonomic structure may have contributed most to discrepancies between our wetland results and those from some stream or marine habitats. Streams in particular have many families of macroinvertebrates with genera or species that have similar environmental sensitivities. This situation is especially true for many EPT families (e.g., Hilsenhoff 1988, Eaton and Lenat 1991). Relatively little internal heterogeneity in tolerances consequently allows families to often provide as much or more signal regarding human influence as genera or species (e.g., Chessman 1995, Wright et al. 1995). However, wetlands do not have many EPT taxa or other coarse-level taxa that

have been documented as particularly sensitive to human activity. Families like Dytiscidae or Chironomidae represent large proportion of total species richness in wetlands, yet genera and species within these families exhibit a wide range of environmental tolerances (Wrubleski 1987, Hudson et al. 1990, Armitage et al. 1995, Epler 1996). In addition, high numbers of species per family for many taxa commonly found in wetlands suggests that extensive adaptive radiation has occurred in wetland habitats, a factor that will likely limit the utility of family-level data for wetland bioassessment (Hawkins et al. 2000, Lenat and Resh 2001). Chironomidae, in particular, exhibit profound adaptive radiation in wetlands and may have been largely responsible for variation among taxonomic levels in our study. Results from tiering family data with species-level Chironomidae data further supports this assertion; this tiered-taxonomic approach produced results equivalent to those obtained with genus and species data, exclusively. We found many chironomid genera that showed differing affinities for specific classes of impairment, and even multiple species within the same genus with different environmental tolerances (e.g., *Dicrotendipes* and *Tanytarsus*). Figure 5A illustrates that little signal was provided by presence or abundance of the family Chironomidae with increasing distance from canal inflow structures in WCA-2A. However, *Dicrotendipes* spp. (4 species) showed a clear unimodal response along this gradient, peaking in the transition zone (Fig. 5B). *Tanytarsus* spp. (6 species), on the other hand, did not exhibit a stress-or-response relationship with distance from the canal (Fig. 5B), yet 4 of its 6 species were associated strongly with either the impacted (*Tanytarsus* sp. F Epler, *Tanytarsus* sp. J Epler) or reference (*Tanytarsus* sp. R Epler, *Tanytarsus* sp. T Epler) zones (Fig. 5C). Given that chironomids contributed >20% of the total number of species collected, it is intuitive that family-level data alone failed to capture as much variation in the environment as did genus or species. Thus, our results suggest that identifications beyond family, particularly of Chironomidae, may be necessary for wetland bioassessment to be effective.

#### *Final considerations and recommendations*

Adequateness of subsampling approaches and levels of taxonomic identification is likely to

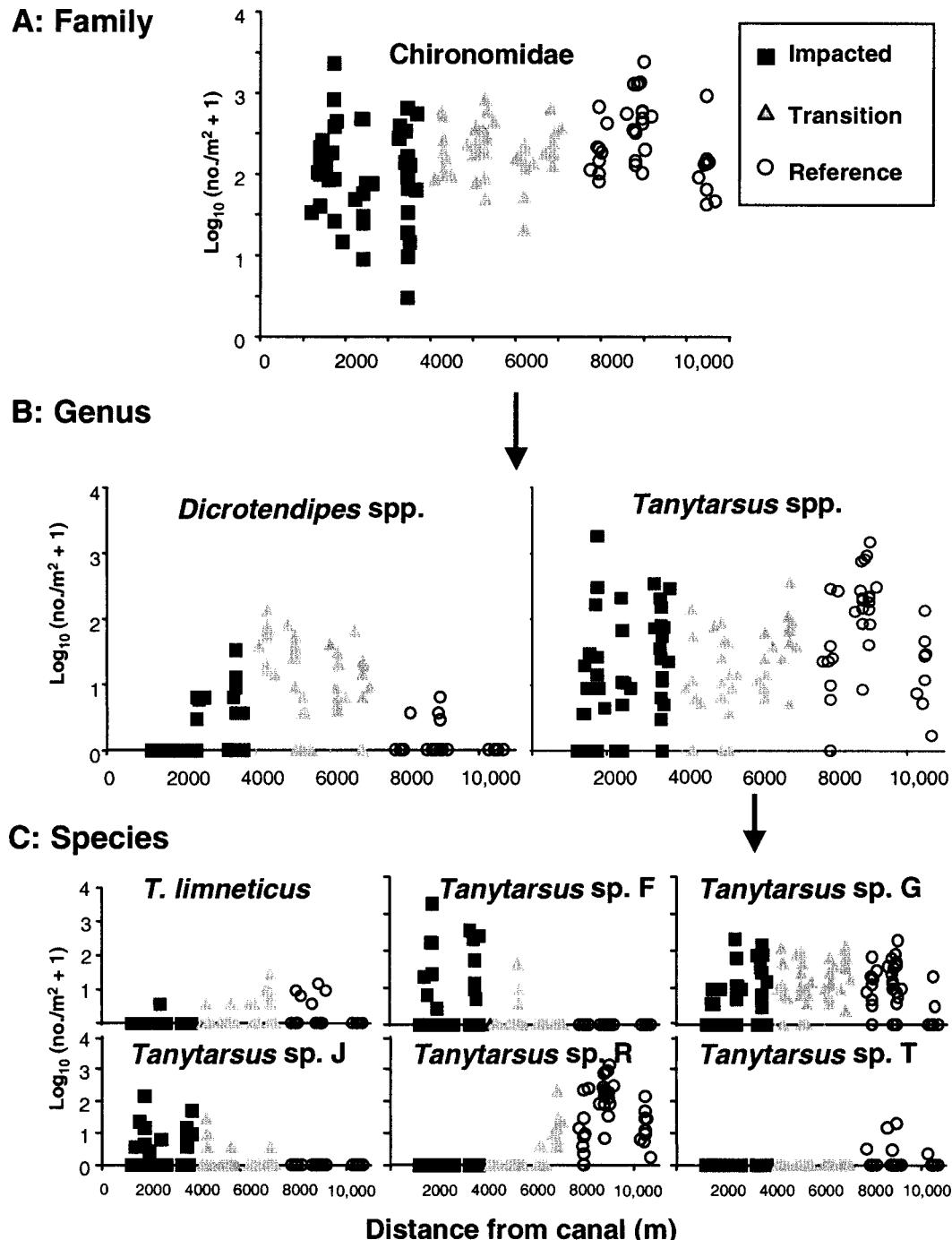


FIG. 5. The effect of differing levels of taxonomic resolution on the environmental signal provided by the Chironomidae. Scatterplots of density ( $\text{no./m}^2$ ) for all taxa in (A) Chironomidae, (B) 2 representative genera (*Dicrotendipes* and *Tanytarsus*), and (C) all 6 species within the genus *Tanytarsus* are shown as a function of distance from canal inflow structures in Water Conservation Area (WCA)-2A. Symbols indicate impacted, transition, and reference landscape zones.

depend upon the goals of a wetland biomonitoring project and available funding. Thus, several of the approaches evaluated in this study may be acceptable, depending on the specific needs of wetland biologists or managers. We recommend:

1) Fixed counts  $\geq 200$  may be the most cost-effective means of obtaining quality biological data. Counts  $< 200$  represented a significant reduction in data quality and cannot be recommended for wetland bioassessment.

2) If fixed-area subsamples (or whole samples) are to be used, we recommend an area sufficiently large to ensure that most samples obtain a minimum of 200 individuals. This approach would ameliorate concerns regarding taxa-area relationships and subsequent areal richness calculations for programs that focus on these attributes (e.g., Courtemanch 1996, Vinson and Hawkins 1996, Larsen and Herlihy 1998). Alternatively, an integrated fixed-area/fixed-count approach may be the best compromise. A smaller, cost-efficient fixed area could be used, but in cases when samples failed to collect a minimum number of individuals, a separate vial could be used to retain a remaining fixed-count component of the subsample. Areal richness metrics would be calculated using only the fixed-area subsample, and the remaining metrics and compositional dissimilarity would be based on the more robust integrated subsample.

3) We recommend a supplementary LR search, regardless of subsampling approach. This search could be more valuable in other wetland systems with greater numbers of potentially sensitive LR taxa, such as odonates, which were not particularly common or diverse in our study. Given the ease of sorting and identifying LR taxa, this search should not incur substantial costs to a wetland bioassessment program.

4) If either subsample size or taxonomic resolution have to be sacrificed, we recommend smaller subsamples with genus- or species-level data as the more judicious tradeoff. We recognize the difficulty for many programs in identifying taxa like chironomids to species, but given their dominance and diversity in wetlands, an attempt to meet this level seems warranted. In the USA, many state agencies have developed in-house keys for local chironomid taxa, and while many are not described as species, morphospecies designations serve equally well in allowing biologists to assign sensitivities and wa-

ter-quality ratings to highly speciose taxa such as Chironomidae (e.g., Epler 1995, NCDWQ 1997). At a minimum, genus-level identifications should be used, as keys to genus for virtually all common macroinvertebrate taxa in North America are ubiquitous and are relatively easy to use by trained biologists (e.g., Merritt and Cummins 1996). A final consideration may be to identify chironomids or other families that contain indicator taxa to genus or species but identify remaining individuals to family (Bailey et al. 2001). We recommend that tiered taxonomy be considered with caution because this approach requires a priori knowledge of taxa sensitivities, and could result in irretrievable loss of potentially useful information.

In conclusion, our study was conducted along a eutrophication gradient. Other types of stressors in other wetland systems could produce different results. For example, 100-count subsamples and family-level taxonomy might be sufficient to accurately detect impairment caused by toxic substances (e.g., metals, pesticides)—other stressors clearly need to be evaluated. However, nutrients represent the leading cause of impairment to lentic systems in the USA, accounting for at least 40% to 60% of observed impairments (USEPA 1998). Moreover, coarse taxonomic structure in the Everglades is quite similar to community composition of most other wetland systems. Thus, our results should be robust and applicable to a relatively large proportion of wetland ecosystems, particularly those threatened or impaired by nutrient enrichment.

#### Acknowledgements

We thank J. Johnson and L. Karppi for assistance in the field, D. Bressler, H. Jacobs, and L. Munger for invaluable help in subsampling, P. Heine, J. Rice, and W. Willis for conducting water and soil chemical analyses, D. Urban for providing study design suggestions, and T. Reynoldson for suggestions on using the BEAST approach. Species identifications were verified by J. Epler, M. Milligan, F. Thompson, A. Rasmussen, R. Maddocks, J. Daigle, and M. Larsen. The manuscript was improved by helpful comments from D. Lenat, C. Stow, T. Reynoldson, J. King, W. Clements, D. Rosenberg, and 2 anonymous reviewers. Funding was provided by the Everglades Agricultural Area, Environmental Protection District.

- ### Literature Cited
- ALTMAN, D. G. 1978. Plotting probability ellipses. *Applied Statistics* 27:347–349.
- APFELBECK, R. S. 1999. Development of biocriteria for wetlands in Montana. Montana Department of Environmental Quality, Helena, Montana. (Available from: Montana Department of Environmental Quality, 2290 Phoenix Avenue, Helena, Montana 59620 USA.)
- APHA (American Public Health Association). 1992. Standard methods for the evaluation of water and wastewater. 18<sup>th</sup> edition. American Public Health Association, Washington, DC.
- ARMITAGE, P., P. S. CRANSTON, AND L. C. V. PINDER (EDITORS). 1995. The Chironomidae: biology and ecology of non-biting midges. Chapman and Hall, London, UK.
- BAILEY, R. C., M. G. KENNEDY, M. Z. DERVISH, AND R. M. TAYLOR. 1998. Biological assessment of freshwater ecosystems using a reference condition approach: comparing predicted and actual benthic invertebrate communities in Yukon streams. *Freshwater Biology* 39:765–774.
- BAILEY, R. C., R. H. NORRIS, AND T. B. REYNOLDSON. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. *Journal of the North American Benthological Society* 20:280–286.
- BARBOUR, M. T., AND J. GERRITSEN. 1996. Subsampling of benthic samples: a defense of the fixed-count method. *Journal of the North American Benthological Society* 15:386–391.
- BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates, and fish. EPA 841-OB-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- BATZER, D. P., AND R. B. WISSINGER. 1996. Ecology of insect communities in nontidal wetlands. *Annual Review of Entomology* 41:75–100.
- BATZER, D. P., S. A. WISSINGER, AND R. B. RADER (EDITORS). 1999. Invertebrates in freshwater wetlands of North America: ecology and management. John Wiley and Sons, New York.
- BOWMAN, M. F., AND R. C. BAILEY. 1997. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? *Canadian Journal of Fisheries and Aquatic Sciences* 54:1802–1807.
- BRINSON, M. M., AND R. RHEINHARDT. 1996. The role of reference wetlands in functional assessment and mitigation. *Ecological Applications* 6:69–76.
- CAO, Y., D. D. WILLIAMS, AND N. E. WILLIAMS. 1998. How important are rare species in aquatic community ecology and bioassessment? *Limnology and Oceanography* 43:1403–1409.
- CHESSMAN, B. C. 1995. Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family level identification and a biotic index. *Australian Journal of Ecology* 20:122–129.
- COURTEMANCH, D. L. 1996. Commentary on the subsampling procedures used for rapid bioassessments. *Journal of the North American Benthological Society* 15:381–385.
- DANIELSON, T. J. 1998. Wetland bioassessment fact sheets. EPA 843-F-98-001. Office of Wetlands, Oceans, and Watersheds, Wetlands Division, US Environmental Protection Agency, Washington, DC.
- DOBERSTEIN, C. P., J. R. KARR, AND L. L. CONQUEST. 2000. The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology* 44:355–371.
- EATON, L. E., AND D. R. LENAT. 1991. Comparison of a rapid bioassessment method with North Carolina's qualitative macroinvertebrate collection method. *Journal of the North American Benthological Society* 10:335–338.
- EPLER, J. H. 1995. Identification manual for the Chironomidae of Florida. Florida Department of Environmental Protection, Tallahassee, Florida. (Available from: <http://www8.myflorida.com/environment/learn/science/laboratories/library/keys.html>)
- EPLER, J. H. 1996. Identification manual for the water beetles of Florida. Florida Department of Environmental Protection, Tallahassee, Florida. (Available from: <http://www8.myflorida.com/environment/learn/science/laboratories/library/keys.html>)
- FAITH, D. P., P. R. MINCHIN, AND L. BELBIN. 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69:57–68.
- FAITH, D. P., AND R. H. NORRIS. 1989. Correlation of environmental variables with patterns of distribution and abundance of common and rare freshwater macroinvertebrates. *Biological Conservation* 50:77–98.
- FDEP (Florida Department of Environmental Protection). 1996. Standard operating procedures manual—benthic macroinvertebrate sampling and habitat assessment methods: 1. Freshwater streams and rivers. Florida Department of Environmental Protection, Tallahassee, Florida. (Available from: Florida Department of Environmental Protection, 2600 Blair Stone Road, Tallahassee, Florida 32399-2900 USA.)
- FERRARO, S. P., AND F. A. COLE. 1990. Taxonomic level and sample size sufficient for assessing pollution impacts on the Southern California Bight macrobenthos. *Marine Ecology Progress Series* 67:251–262.
- FERRARO, S. P., AND F. A. COLE. 1992. Taxonomic level

- sufficient for assessing a moderate impact on macrobenthic communities in Puget Sound, Washington, USA. *Canadian Journal of Fisheries and Aquatic Sciences* 49:1184–1188.
- FORTIN, M.-J., P. DRAPEAU, AND P. LEGENDRE. 1989. Spatial autocorrelation and sampling design in plant ecology. *Vegetatio* 83:209–222.
- FOSTER, D. R., M. FLUET, AND E. R. BOOSE. 1999. Human or natural disturbance: landscape-scale dynamics of the tropical forests of Puerto Rico. *Ecological Applications* 9:555–572.
- GERMANO, J. D. 1999. Ecology, statistics, and the art of misdiagnosis: the need for a paradigm shift. *Environmental Reviews* 7:167–190.
- GERNES, M. C., AND J. C. HELGEN. 1999. Indexes of Biotic Integrity (IBI) for wetlands: vegetation and invertebrate IBIs. Report to US Environmental Protection Agency, #CD995525-01. Environmental Outcomes Division, Minnesota Pollution Control Agency, Minneapolis, Minnesota.
- GERRITSEN, J., M. T. BARBOUR, AND K. KING. 2000. Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. *Journal of the North American Benthological Society* 19:487–496.
- GROWNS, J. E., B. C. CHESSMAN, J. E. JACKSON, AND D. G. ROSS. 1997. Rapid assessment of Australian rivers using macroinvertebrates: cost and efficiency of 6 methods of sample processing. *Journal of the North American Benthological Society* 16: 682–693.
- HAWKINS, C. P., AND R. H. NORRIS. 2000. Performance of different landscape classifications for aquatic bioassessments: introduction to the series. *Journal of the North American Benthological Society* 19: 367–369.
- HAWKINS, C. P., R. H. NORRIS, J. N. HOGUE, AND J. W. FEMINELLA. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* 10: 1456–1477.
- HEWLETT, R. 2000. Implications of taxonomic resolution and sample habitat for stream classification at a broad geographic scale. *Journal of the North American Benthological Society* 19:352–361.
- HILSENHOFF, W. L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. *Journal of the North American Benthological Society* 7:65–68.
- HUDSON, P. L., D. R. LENAT, B. A. CALDWELL, AND D. SMITH. 1990. Chironomidae of the southeastern United States: a checklist of species and notes on biology, distribution, and habitat. Fish and Wildlife Research 7. US Department of the Interior, Washington, DC.
- JOHNSON, D. H. 1999. The insignificance of statistical significance testing. *Journal of Wildlife Management* 63:763–772.
- KARR, J. R., AND E. W. CHU. 1997. Biological monitoring and assessment: using multimetric indexes effectively. EPA 235-R97-001. University of Washington, Seattle, Washington.
- KING, R. S. 2001. Dimensions of invertebrate assemblage organization across a phosphorus-limited Everglades landscape. PhD Dissertation, Duke University, Durham, North Carolina.
- KING, R. S., K. T. NUNNERY, AND C. J. RICHARDSON. 2000. Macroinvertebrate assemblage response to highway crossings in forested wetlands: implications for biological assessment. *Wetlands Ecology and Management* 8:243–256.
- KUSLER, J., AND W. NIERING. 1998. Wetland assessment: have we lost our way? *National Wetlands Newsletter* 20:8–14.
- LARSEN, D. P., AND A. T. HERLIHY. 1998. The dilemma of sampling streams for macroinvertebrate richness. *Journal of the North American Benthological Society* 17:359–366.
- LEDUC, A., P. DRAPEAU, Y. BERGERON, AND P. LEGENDRE. 1992. Study of spatial components of forest cover using partial Mantel tests and path analysis. *Journal of Vegetation Science* 3:69–78.
- LEGENDRE, P., AND M. ANDERSON. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial experiments. *Ecological Monographs* 69:1–24.
- LEGENDRE, P., AND L. LEGENDRE. 1998. Numerical ecology. 2nd edition. Elsevier, Amsterdam, The Netherlands.
- LEMLY, A. D., AND R. S. KING. 2000. An insect-bacteria bioindicator for assessing detrimental nutrient enrichment in wetlands. *Wetlands* 20:91–100.
- LENAT, D. R., AND V. H. RESH. 2001. Taxonomy and stream ecology—The benefits of genus- and species-level identifications. *Journal of the North American Benthological Society* 20:287–298.
- MANLY, B. F. J. 1997. Randomization, bootstrap, and Monte Carlo methods in biology. 2<sup>nd</sup> edition. Chapman and Hall, London, UK.
- MANTEL, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27:209–220.
- MARCHANT, R. 1999. How important are rare species in aquatic community ecology and bioassessment? A comment on the conclusions of Cao et al. *Limnology and Oceanography* 44:1840–1841.
- MARCHANT, R., A. HIRST, R. H. NORRIS, R. BUTCHER, L. METZELING, AND D. TILLER. 1997. Classification and prediction of macroinvertebrate assemblages from running waters in Victoria, Australia. *Journal of the North American Benthological Society* 16:664–681.
- MAXTED, J. R., M. T. BARBOUR, J. GERRITSEN, V. PORETTI, N. PRIMROSE, A. SILVIA, D. PENROSE, AND R. RENFROW. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic ma-

- croinvertebrates. *Journal of the North American Benthological Society* 19:128–144.
- MCCORMICK, P. V., P. S. RAWLIK, K. LURDING, E. P. SMITH, AND F. H. SKLAR. 1996. Periphyton-water quality relationships along a nutrient gradient in the northern Florida Everglades. *Journal of the North American Benthological Society* 15:433–449.
- MCCUNE, B., J. P. DEY, J. E. PECK, D. CASSELL, K. HEIMAN, S. WILL-WOLF, AND P. N. NEITLICH. 1997a. Repeatability of community data: species richness versus gradient scores in large-scale lichen studies. *Bryologist* 100:40–46.
- MCCUNE, B., J. DEY, J. PECK, K. HEIMAN, AND S. WILL-WOLF. 1997b. Regional gradients in lichen communities of the southeast United States. *Bryologist* 100:145–158.
- MERRITT, R. W., AND K. W. CUMMINS (EDITORS). 1996. An introduction to the aquatic insects of North America. 3<sup>rd</sup> edition. Kendall/Hall, Dubuque, Iowa.
- MOSS, D., J. F. WRIGHT, M. T. FURSE, AND R. T. CLARKE. 1999. A comparison of alternative techniques for prediction of the fauna of running-water sites in Great Britain. *Freshwater Biology* 41:167–181.
- NCDWQ (North Carolina Division of Water Quality). 1997. Standard operating procedures, biological monitoring. North Carolina Department of Environment and Natural Resources, Raleigh, North Carolina. (Available from: Division of Water Quality, North Carolina Department of Environment and Natural Resources, 1621 Mail Service Center, Raleigh, North Carolina 29699-1621 USA.)
- OBEYSEKERA, J., AND K. RUTCHY. 1997. Selection of scale for Everglades landscape models. *Landscape Ecology* 12:7–18.
- OLIVER, I., AND A. J. BEATTIE. 1996. Invertebrate morphospecies as surrogates for species: a case study. *Conservation Biology* 10:99–109.
- OLSGARD, F., P. J. SOMERFIELD, AND M. R. CARR. 1998. Relationships between taxonomic resolution, macrobenthic community patterns and disturbance. *Marine Ecology Progress Series* 172:25–36.
- OWEN, J. G., AND M. A. CHMIELEWSKI. 1985. On canonical variates ellipses and the construction of confidence ellipses in systematic studies. *Systematic Zoology* 34:366–374.
- PLAFKIN, J. L., M. T. BARBOUR, K. D. PORTER, S. K. GROSS, AND R. M. HUGHES. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA/440/4-89/001. Office of Water, United States Environmental Protection Agency, Washington, DC.
- REYNOLDSON, T. B., R. C. BAILEY, K. E. DAY, AND R. H. NORRIS. 1995. Biological guidelines for freshwater sediment based on BEthnic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. *Australian Journal of Ecology* 20:198–219.
- REYNOLDSON, T. B., R. H. NORRIS, V. H. RESH, K. E. DAY, AND D. M. ROSENBERG. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* 16:833–852.
- REYNOLDSON, T. B., D. M. ROSENBERG, AND V. H. RESH. 2001. Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchment, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* 58:1395–1410.
- RICHARDSON, C. J., P. VAITHYANATHAN, R. J. STEVENSON, R. S. KING, C. A. STOW, R. G. QUALLS, AND S. S. QIAN. 2000. The ecological basis for a phosphorus (P) threshold in the Everglades: directions for sustaining ecosystem structure and function. Duke Wetland Center Publication 00-02. Duke University, Durham, North Carolina.
- ROMANOWICZ, E. A., AND C. J. RICHARDSON. 1997. Hydrologic investigation of Water Conservation Area 2A. Pages 12.1–12.29 in C. J. Richardson (editor). Effects of phosphorus and hydroperiod alterations on ecosystem structure and function in the Everglades. Duke Wetland Center Publication 97-05. Duke University, Durham, North Carolina.
- ROSENBERG, D. M., AND V. H. RESH (EDITORS). 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.
- SANDERSON, R. A., S. P. RUSHTON, A. J. CHERRILL, AND J. P. BYRNE. 1995. Soil, vegetation and space: an analysis of their effects on the invertebrate communities of a moorland in north-east England. *Journal of Applied Ecology* 32:506–518.
- SFWMD (South Florida Water Management District). 1992. Surface water improvement plan for the Everglades. Supporting Information Document, South Florida Water Management District, West Palm Beach, Florida. (Available from: South Florida Water Management District, 3301 Gun Club Road, West Palm Beach, Florida 33416 USA.)
- SFWMD (South Florida Water Management District). 2000. 2000 Everglades consolidated report. Supporting Information Document, South Florida Water Management District, West Palm Beach, Florida. (Available from: South Florida Water Management District, 3301 Gun Club Road, West Palm Beach, Florida 33416 USA.)
- SMITH, M. J., W. R. KAY, D. H. D. EDWARD, P. J. PAPAS, K. S. J. RICHARDSON, M. SIMPSON, D. J. CALE, P. H. J. HORWITZ, J. A. DAVIS, F. H. YUNG, R. H. NORRIS, AND S. A. HALSE. 1999. AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biology* 41:269–282.

- SOMERS, K. M., R. A. REID, AND S. M. DAVID. 1998. Rapid biological assessments: how many animals are enough? *Journal of the North American Benthological Society* 17:348–358.
- SOVELL, L. A., AND B. VONDRAČEK. 1999. Evaluation of the fixed-count method for Rapid Bioassessment Protocol III with benthic macroinvertebrate metrics. *Journal of the North American Benthological Society* 18:420–426.
- SUTER, G. W. 1996. Abuse of hypothesis testing statistics in ecological risk assessment. *Human and Ecological Risk Assessment* 2:331–347.
- TURNER, A. M., AND J. C. TREXLER. 1997. Sampling aquatic invertebrates from marshes: evaluating the options. *Journal of the North American Benthological Society* 16:694–709.
- URBAN, D. L. 2000. Using model analysis to design monitoring programs for landscape management and impact assessment. *Ecological Applications* 10:1820–1832.
- URKIAGA-ALBERDI, J., S. PAGOLA-CARTE, AND J. I. SAIZ-SALINAS. 1999. Reducing effort in the use of benthic bioindicators. *Acta Oecologica* 20:489–497.
- USEPA (US Environmental Protection Agency). 1997a. Wetlands: biological assessment methods and criteria development workshop. EPA 843-3-97-001. Office of Water, US Environmental Protection Agency, Washington, DC.
- USEPA (US Environmental Protection Agency). 1997b. Field and laboratory methods for macroinvertebrate and habitat assessment of low-gradient, nontidal streams. Mid-Atlantic Coastal Streams Workgroup, Environmental Services Division, Region 3, USEPA, Wheeling, West Virginia. (Available from: Environmental Services Division, Region 3, USEPA, 303 Methodist Building, 11<sup>th</sup> and Chapline Streets, Wheeling, West Virginia 26003 USA.)
- USEPA (US Environmental Protection Agency). 1998. National strategy for the development of regional nutrient criteria. EPA 822-R-98-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- VAITHIYANATHAN, P., AND C. J. RICHARDSON. 1999. Macrophyte species changes in the Everglades: examination along a eutrophication gradient. *Journal of Environmental Quality* 28:1347–1358.
- VINSON, M. R., AND C. P. HAWKINS. 1996. Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. *Journal of the North American Benthological Society* 15:392–399.
- WALSH, C. J. 1997. A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. *Marine and Freshwater Research* 48:241–248.
- WARWICK, R. M. 1993. Environmental impact studies on marine communities: pragmatical considerations. *Australian Journal of Ecology* 18:63–80.
- WRIGHT, I. A., B. C. CHESSMAN, P. G. FAIRWEATHER, AND L. J. BENSON. 1995. Measuring the impact of sewage effluent on the macroinvertebrate community of an upland stream: the effect of different levels of taxonomic resolution and quantification. *Australian Journal of Ecology* 20:142–149.
- WRUBLESKI, D. A. 1987. Chironomidae (Diptera) of peatlands and marshes in Canada. *Memoirs of the Entomological Society of Canada* 140:141–161.
- WU, Y., F. H. SKLAR, AND K. RUTCHEY. 1997. Analysis and simulations of fragmentation patterns in the Everglades. *Ecological Applications* 7:268–276.
- ZAMORA-MUÑOZ, C., AND J. ALBA-TERCEDOR. 1996. Bioassessment of organically polluted Spanish rivers, using a biotic index and multivariate methods. *Journal of the North American Benthological Society* 15:332–352.

Received: 24 January 2001

Accepted: 15 October 2001