EVALUATING SUBSAMPLING APPROACHES AND MACROINVERTEBRATE TAXONOMIC RESOLUTION FOR WETLAND BIOASSESSMENT

Ryan S. King and Curtis J. Richardson

Duke University Wetland Center, Durham, North Carolina

ABSTRACT

Two of the most controversial issues in stream bioassessment using macroinvertebrates, subsampling and taxonomic resolution, have yet to be quantitatively addressed for wetlands. We evaluated the efficacy of family-, genus-, and species-level assemblage data in reflecting environmental condition and classifying samples into strata of ecological impairment using 5 basic levels of subsampling (100-, 200-, and 300-organism fixed counts; 10% and 25% fixed areas), 5 integrated subsamples requiring a minimum fixed count and fixed area (100 & 10%, 200 & 10%, 300 & 10%, 200 & 25%, and 300 & 25%), and each of the 10 basic and integrated subsamples supplemented with a "large-rare" (LR) search. Data were obtained using 1.5-m^2 composite samples collected from 126 plots along a 10-km long eutrophication gradient in the Everglades. Our results suggest that effectiveness of subsampling depends 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 fixed counts resulting in significantly greater assemblage-environment relationships, lower misclassification rates, and more robust classification models than 10% fixed area despite averaging similar numbers of individuals. However, integrated subsamples significantly improved relationships and reduced misclass-ifications for all basic subsamples except 25%. The greatest marginal improvement with increasing subsample size was observed between fixed counts of 100 and 200; 100 counts performed significantly worse than all other subsampling approaches. Supplementing subsamples with a LR search always resulted in a slight improvement in assemblage-environment relationships, and was particularly important in reducing misclassification rates for family-level data. However, family-level assemblage-environment relationships and predictive capabilities were highly inferior to genus- and species-level data, regardless of subsample size. Species-level data performed best, largely due to the large proportion (> 20%) of total species belonging to Chironomidae. Our results suggest that fixed counts ³ 200 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 large-rare searches and highly recommend species-level taxonomy.

INTRODUCTION

Despite high interest in wetland biological assessment using macroinvertebrates, no accepted assessment protocols have been developed and published like those that exist for streams (e.g., Barbour et al. 1999). Wetland macroinvertebrate assemblages present difficulties in sampling and sample processing that are less prevalent in lotic bioassessments:

•Large amounts of course particulates in wetland samples can make sorting of organisms highly laborious and possibly cost prohibitive.

•An abundance of taxonomic groups that are challenging and time consuming to identify, such as Chironomidae.

Various approaches to subsampling large composite samples (e.g., fixed counts, fixed areas, and large-rare searches) and identification of organisms to coarse levels of taxonomic resolution (e.g., family) have been proposed as two potential ways of making wetland bioassessment cost effective. However, no wetland studies have addressed effects of these approaches on assessment accuracy.

Here, we build upon research conducted in streams and compare several approaches of subsampling and levels of taxonomic resolution in a comprehensive wetland study. The foundation of our study is that biota are a reflection of their environment, which is the basic premise behind bioassessment. **We posed two questions: Do differing levels of subsampling and taxonomic resolution affect**

•the magnitude of assemblage-environment relationships in wetlands?

•the ability to correctly classify wetland samples into distinct strata of ecological impairment?

To address these questions, we sampled along a 10-km eutrophication gradient in the Everglades, an area in which environmental conditions are well known. **Our goal was to recommend the subsampling strategy and taxonomic level that best represented the wetland environment, yet were the least labor intensive of the approaches.**

STUDY AREA AND SAMPLING DESIGN

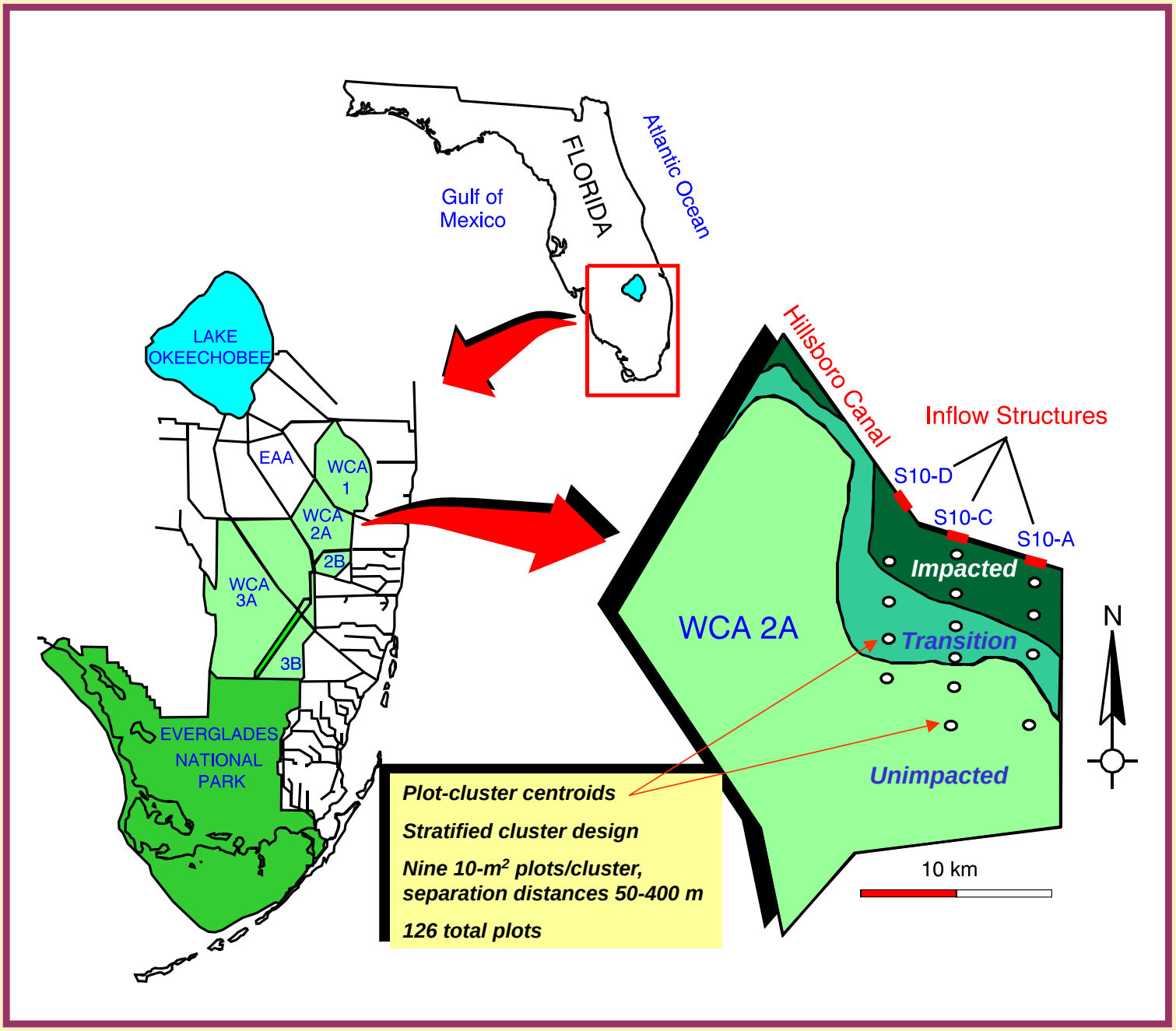

Figure 1. Map showing location of the study area and sampling design.

STUDY AREA CHARACTERISTICS

Table 1. Mean (± 1 SD) of several key spatial and environmental variables among impacted, transition, and unimpacted landscape zones. Canal inputs of phosphorus (P) were particularly important and responsible for many ecosystem changes.

		Impacted (n=45)	Transition (n=45)	Unimpacted (n=36)
Variable	Units	Mean (± SD)	Mean (± SD)	Mean (± SD)
Distance from canal	m	2500 (870)	5540 (910)	9050 (920)
Total P (sediment)	mg/kg	1430 (172)	1200 (181)	570 (150)
Total P (surface water)	mg/L	78.4 (74.3)	27.4 (26.4)	9.9 (5.4)
Dissolved sodium (surface water)	mg/L	98.8 (39.7)	89.5 (26.5)	78.5 (25.6)
Water depth (1981-1998)	cm	29.0 (8.7)	32.3 (9.6)	31.2 (11.4)
Dominant macrophyte vegetation		*Typha* (cattail); *Mikania* (vines); duckweed	*Cladium* (sawgrass) cattail; vines; water lily	sawgrass water lily, spike-rush; bladderwort

METHODS

Macroinvertebrate Sampling

Sampling based on protocols developed by FDEP (1996; SOP #BA-7) and USEPA (1997):

- D-framed dip net (0.3-m wide, 500-mm mesh)
- Ten 0.5-m length sweeps across each plot (1.5 m^2 total area)
- Contents of all 10 sweeps composited into 500-mm mesh sieve bucket, rinsed in field, and preserved in 5% (v/v) buffered formalin stained with rose bengal.

Sample Processing and Subsampling

We chose 20 combinations of subsampling based on 4 published approaches:

- Fixed counts (100, 200, or 300 organisms; Barbour and Gerritsen 1996, Barbour et al. 1999)
- Fixed areas (10% or 25% of total sample; Courtemanch 1996, Walsh 1997)
- Integrated subsample requiring minimum fixed count *and* fixed area (100&10%, 200&10%, 300&10%, 200&25%, 300&25%; Walsh 1997)
- Large-rare (LR) search supplementing each of the 10 fixed count, fixed area, and integrated subsamples (Courtemanch 1996, Vinson and Hawkins 1996).

Data sets were constructed for each subsampling approach and each level of taxonomic resolution (family, genus, and species).

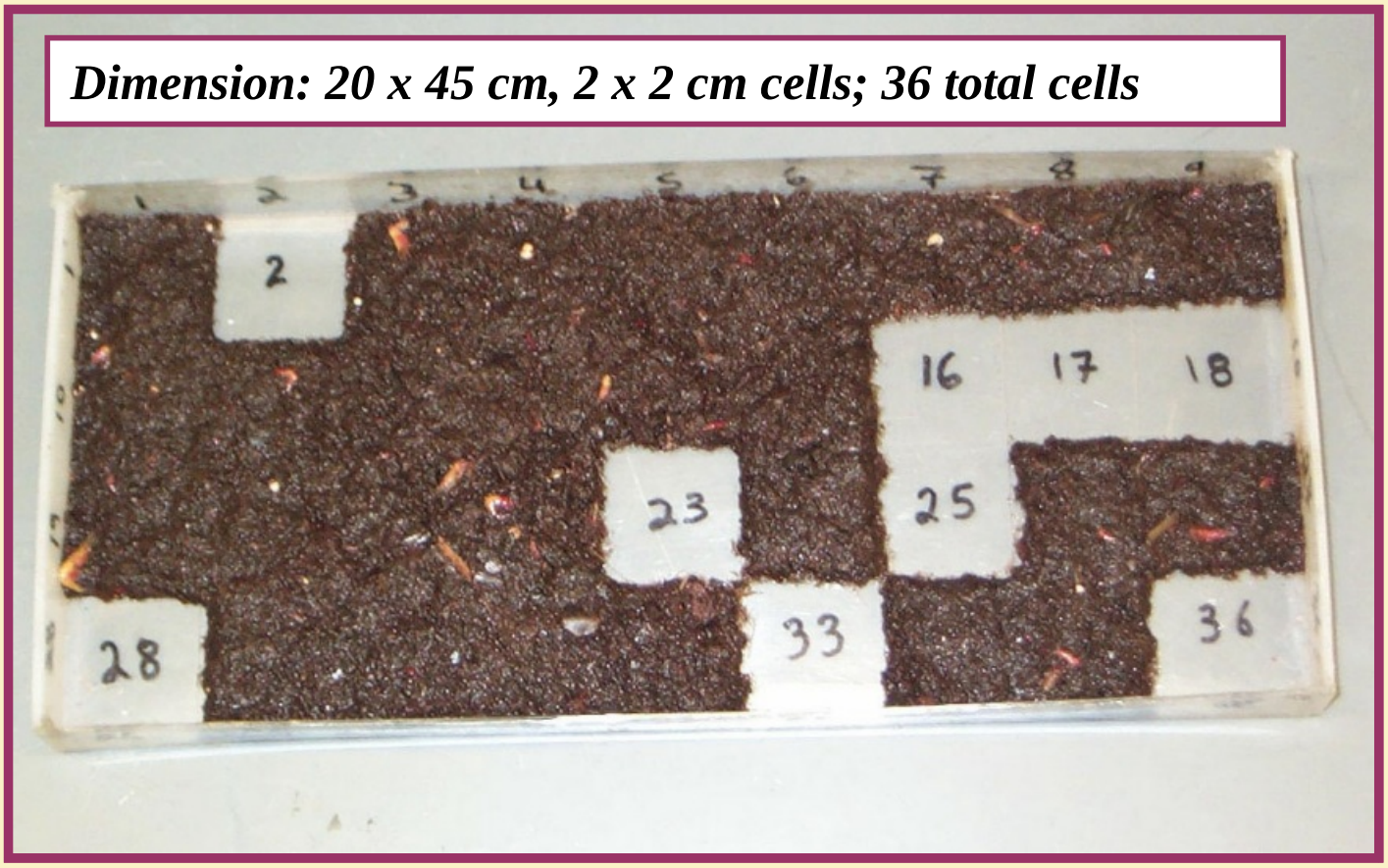

Figure 2. Photograph of a sample contained in the gridded pan used for subsampling. Samples were homogenized and spread evenly in pan. Cells were selected randomly and material within removed and transferred to a smaller gridded dish. Macroinvertebrates were sorted from organic matter under a stereomicroscope at 10X magnification. The process was repeated until a target area or number of individuals was obtained. Subsample areal fraction was used to convert numerical abundance into density (no./m^2) for each subsample based on total sample area (1.5 m^2). Method followed Florida DEP (1996; SOP #BA-8).

Table 2. Properties of macroinvertebrate subsamples (integrated subsamples not shown) taken from 1.5-m^2 dip-net composite samples (n=126).

	% of total sample		No. of individuals			No. + LR search		Sorting time (min)	
Sub-sample	Mean	± SD	Mean	± SD	Range	Mean	± SD	Mean	± SD
100 count	8	6	103	5	92-118	121	14	94	55
200 count	15	13	204	7	191-224	221	15	156	100
300 count	23	18	305	11	283-326	321	16	207	118
10% area	10	0	230	178	23-1036	248	178	95	52
25% area	25	0	573	440	62-2558	588	439	251	136

Data Analysis

We used complementary nonparametric statistical procedures to compare subsampling approaches and taxonomic levels:

- Assemblage-environment relationships were contrasted using Mantel's tests, a multivariate technique that relates ecological differences (assemblage data) to environmental differences. Magnitude of relationships were compared using 95% CI estimated by bootstrapping. Three important environmental variables were used independently as predictors of assemblage composition.
- Classification into impact strata was assessed using ordination (nonmetric multidimensional scaling, or nMDS) and classification tree analysis. Ordination was used to provide a visual assessment of separation of plots into the 3 impact zones based on assemblage data. Classification tree analysis evaluated the ability of data to correctly classify samples into impact strata.

ASSEMBLAGE-ENVIRONMENT RELATIONSHIPS

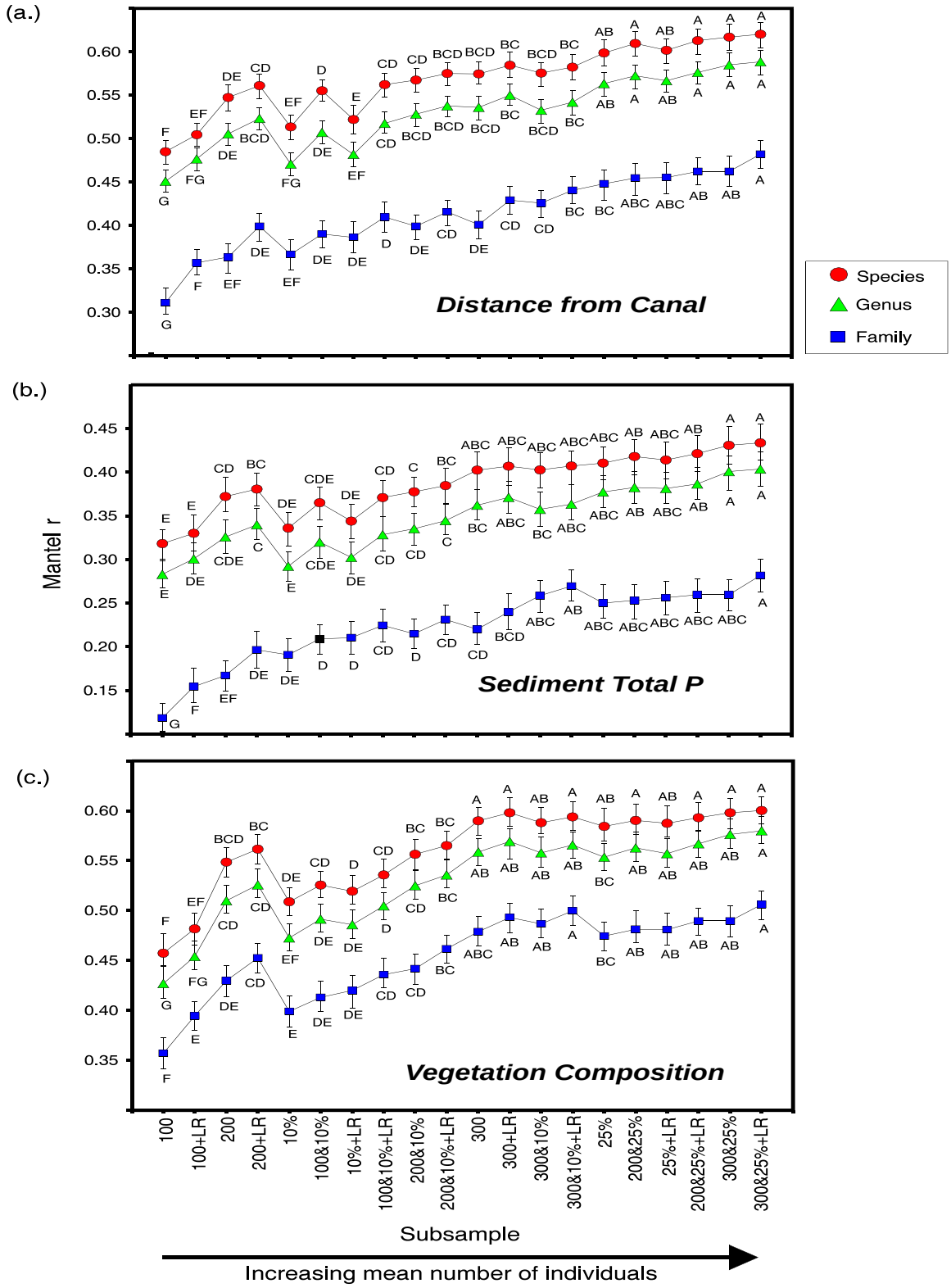

Figure 3. Results from Mantel's tests relating macroinvertebrate multivariate assemblage composition (Bray-Curtis dissimilarity) to environmental variables. Mantel *r* values indicate the magnitude of assemblage-environment relationships. Within taxonomic levels, subsamples with the same letters are not significantly different (95% CI). Among taxonomic levels, subsamples with overlapping confidence intervals are not different.

CLASSIFICATION INTO IMPACT STRATA

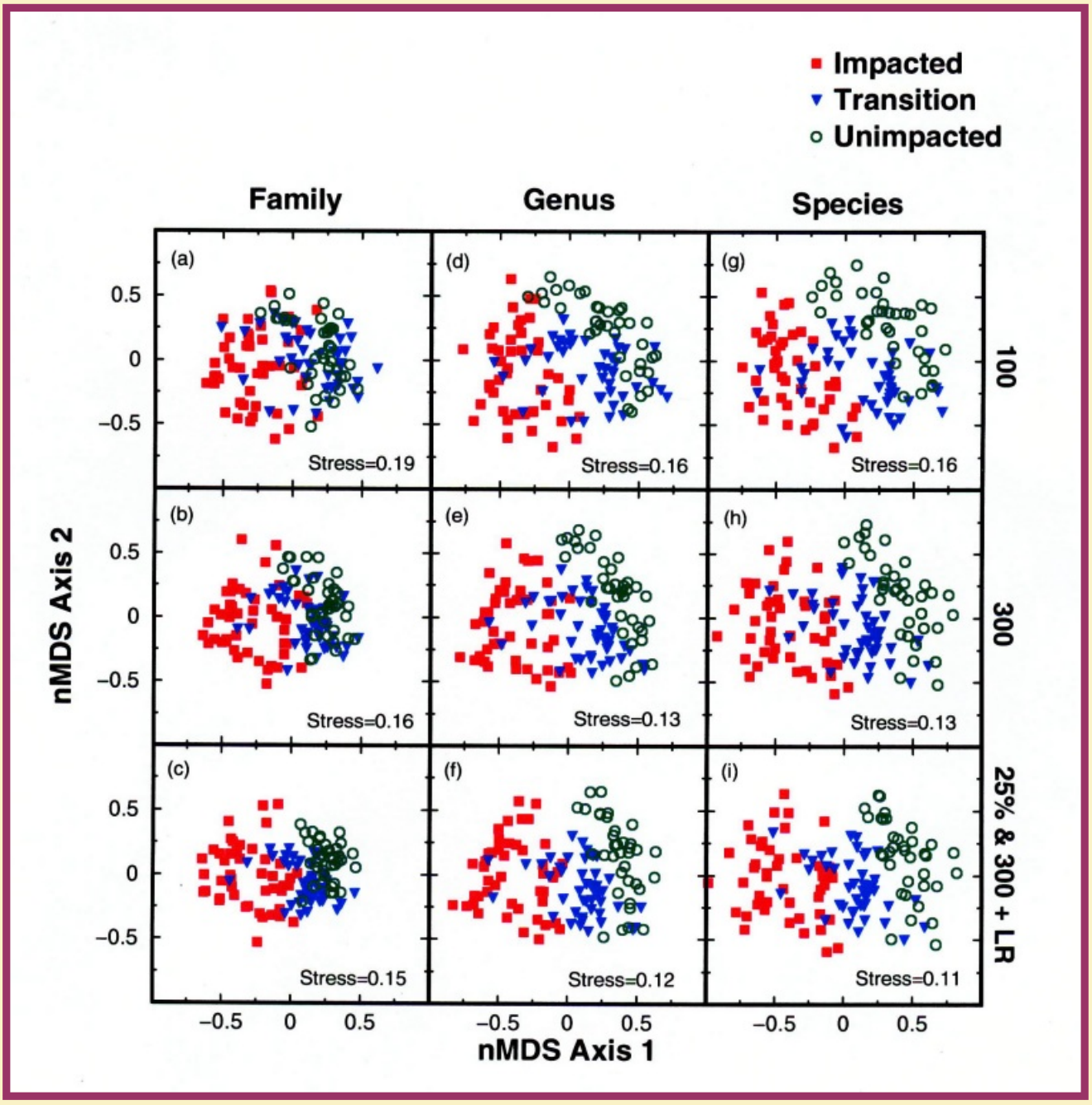

Figure 4. Nonmetric multidimensional scaling (nMDS) ordin-ations of macroinvertebrate assemblage data among family, genus, and species levels of taxonomic resolution and 100, 300, and 25% & 300 + LR levels of subsampling. Note the compression and extensive overlap of plots among impact zones with family-level data and/or 100-count subsamples.

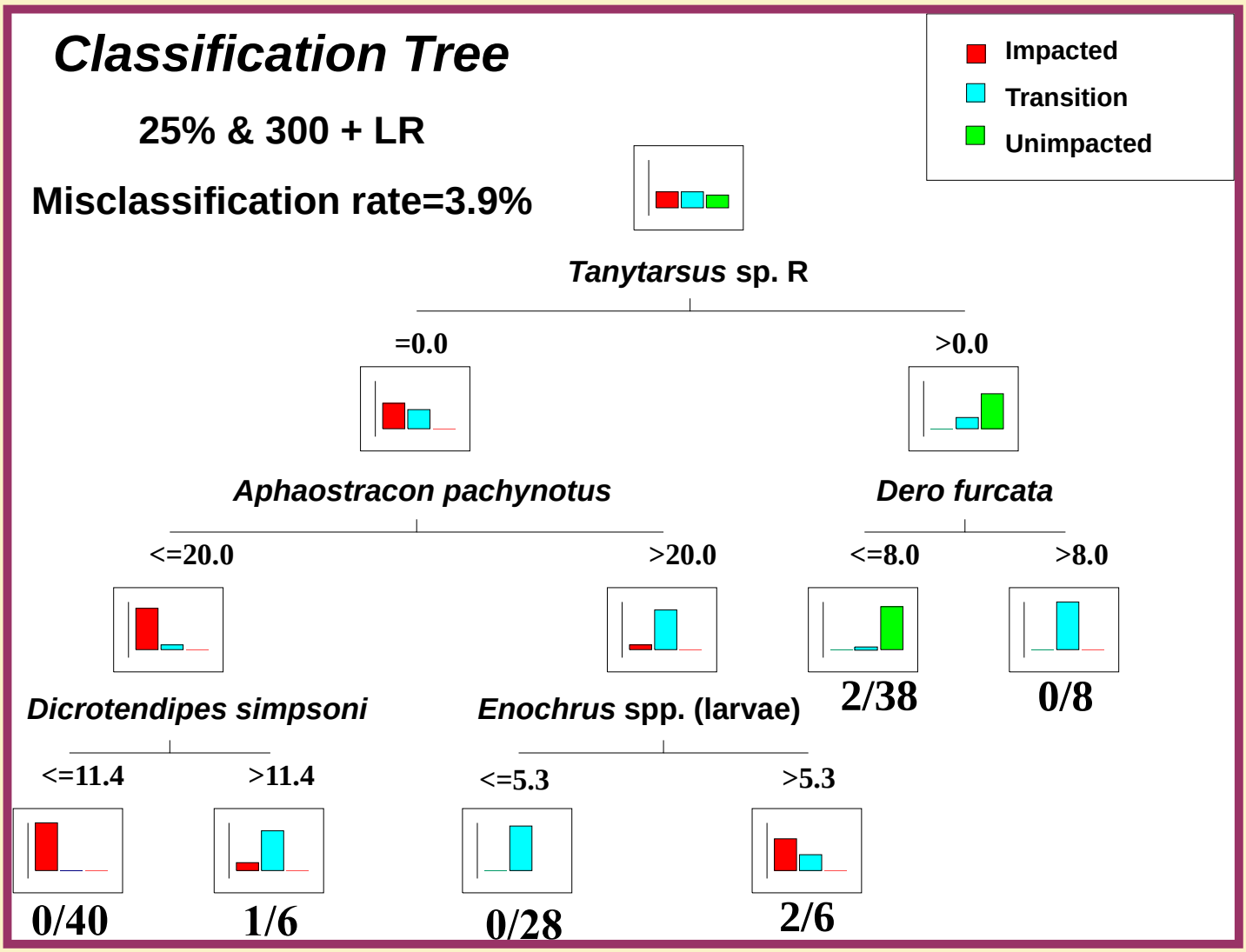

Figure 5. Example of a classification tree using species-level data. Names show the taxa that best classified plots into impact strata. Numbers below the terminal nodes (leaves) show the number of misclassified plots/total number of plots in each node.

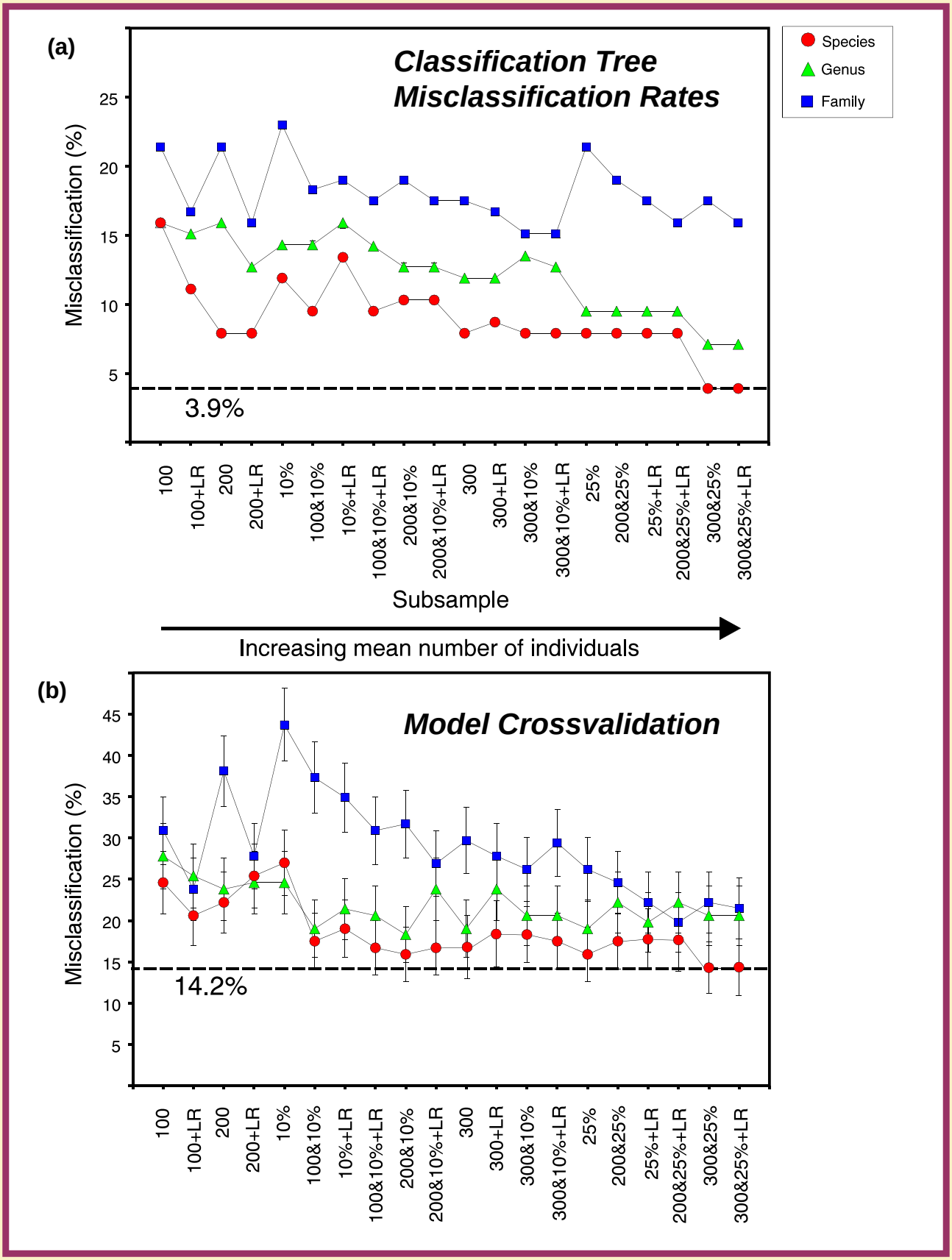

Figure 6. Results from classification tree analysis. Misclass-ification rates (%) indicate the accuracy of data in correctly classifying plots into impact strata (a) as well as the robustness of each tree model, as estimated using crossvalidation (b).

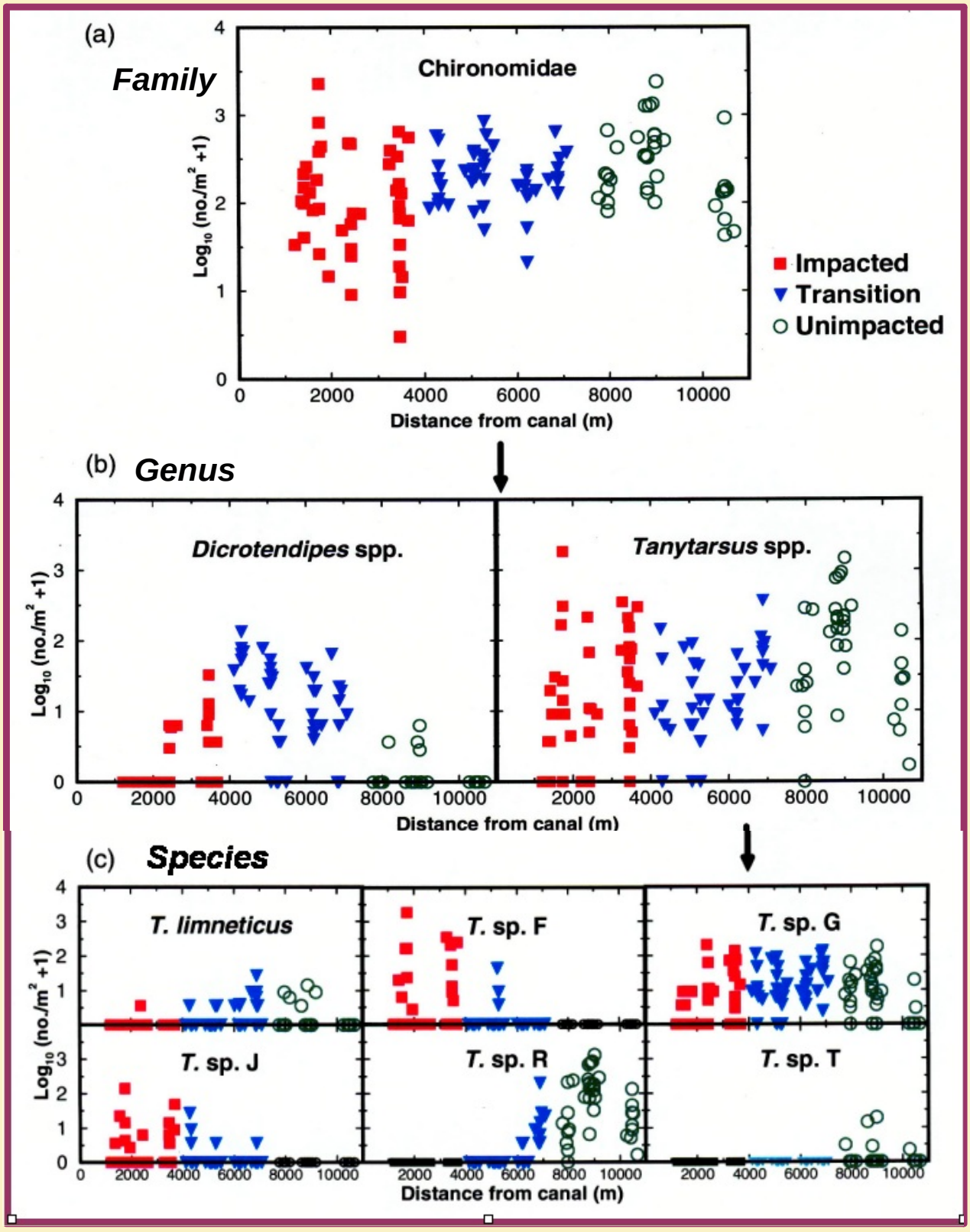

Figure 7. Illustration of the effect of differing levels of taxonomic resolution on the ecological signal provided by the family Chironomidae. Comparisons among family (all 51 species), two representative genera, *Dicrotendipes* and *Tanytarsus* (4 and 6 species, respectively), and each species of *Tanytarsus* show that much information is lost using family or genus level data.

CONCLUSIONS AND RECOMMENDATIONS

- **100-count subsamples performed significantly worse than all other approaches. This has serious implications for state and federal protocols that currently advocate this technique.**
- **The greatest marginal improvement was detected between 100 and 200-count subsamples. Beyond 200, marginal improvements diminished rapidly.**
- **Fixed-area subsampling results indicate that obtaining a minimum number, or "threshold", of organisms is more important than standardizing for a specific area. This also has implications for programs that use small "quantitative" samples and pick samples exhaustively.**
- **Integrating fixed counts and fixed areas helped make subsamples more accurate and robust, particularly for 10% fixed-area subsamples that had many small counts.**
- **Large-rare (LR) searches always resulted in slight improvements in assemblage-environment relationships.**
- **Family-level data performed much worse than genus or species; species was consistently the best.**
- **Profound adaptive radiation in many wetland macroinvertebrate groups (e.g., Chironomidae) will limit the utility of family-level data.**
- **We recommend the following for wetland bioassessment:**
  - **Fixed-count subsamples ^{3} 200 *or* integrated fixed count/fixed areas that consistently obtain at least 200 individuals.**
  - **Large-rare searches as a rapid, cost-effective way to improve accuracy in bioassessment.**
  - **Species-level taxonomy; genus-level at a minimum. We highly caution the use of family-level data for wetlands.**

REFERENCES

- Barbour, M. T. and J. Gerritsen. 1996. *J. N. Am. Benthol. Soc.* 15:386-391
- Barbour, M. T. et al. 1999. EPA 841-B-99-002.
- Courtemanch, D. L. 1996. *J. N. Am. Benthol. Soc.* 15:381-385
- Florida DEP. 1996. Standard operating procedures manual–benthic macroinvertebrates.
- USEPA. 1997. Field and laboratory methods for macroinvertebrate habitat assessment of low gradient, nontidal streams.
- Vinson, M. R., and C. P. Hawkins. 1996. *J. N. Am. Benthol. Soc.* 15:392-399
- Walsh, C. J. 1997. *Mar. Fresh. Res*. 48:241-248

ACKNOWLEDGMENTS

We thank J. Johnson and L. Karppi for assistance in field, D. Bressler, H. Jacobs, and L. Munger for help in subsampling, P. Heine, J. Rice, and W. Willis for soil and water analyses, M. Ho and P.V. Sundareshwar for technical assistance, and D. Urban for study design suggestions. Species identifications were verified by J. Epler, M. Milligan, F. Thompson, A. Rasmussen, J. Daigle, and M. Larsen. Funding was provided by a grant from the EAA Environmental Protection District to the Duke University Wetland Center.