

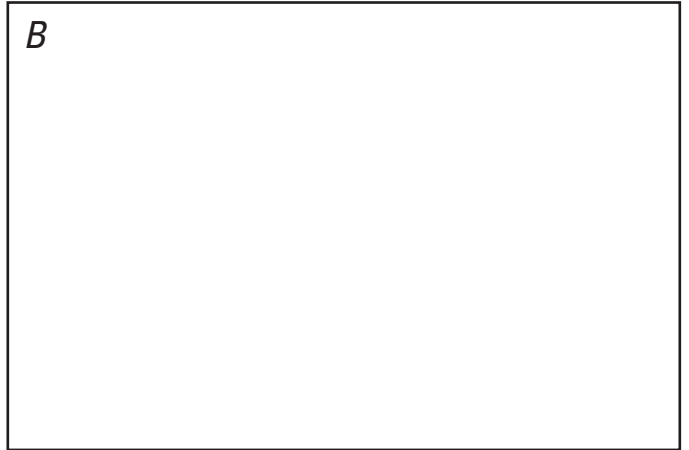
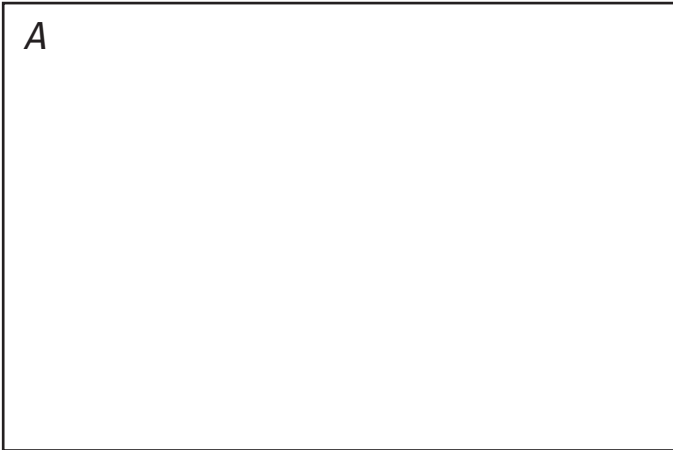
**National Water-Quality Assessment Program**

Prepared in cooperation with the Wyoming Department of Environmental Quality

**Comparison of Macroinvertebrate Community Structure  
between Two Riffle-Based Sampling Protocols in Wyoming,  
Colorado, and Montana, 2000-2001**



Scientific Investigations Report 2006-5117



**On the cover:** Photographs of streams in the Middle Rockies ecoregion (A and C), the Northwestern Great Plains ecoregion (B), and the Wyoming Basin ecoregion (D), and. Photographs A, B, and C by Ronald B. Zelt, U.S. Geological Survey, and photograph D by Gregory K. Boughton, U.S. Geological Survey, August and September 2000.

# **Comparison of Macroinvertebrate Community Structure between Two Riffle- Based Sampling Protocols in Wyoming, Colorado, and Montana, 2000-2001**

By David A. Peterson and Jeremy R. Zumberge

Prepared in cooperation with the Wyoming Department of Environmental Quality

Scientific Investigations Report 2006–5117

**U.S. Department of the Interior  
U.S. Geological Survey**

**U.S. Department of the Interior**  
P. Lynn Scarlett, Acting Secretary

**U.S. Geological Survey**  
P. Patrick Leahy, Acting Director

U.S. Geological Survey, Reston, Virginia: 2006

For product and ordering information:

World Wide Web: <http://www.usgs.gov/pubprod>

Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:

World Wide Web: <http://www.usgs.gov>

Telephone: 1-888-ASK-USGS

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

*Suggested citation:*

Peterson, D.A., and Zumberge, J.R., 2006, Comparison of macroinvertebrate community structure between two riffle-based sampling protocols in Wyoming, Colorado, and Montana, 2000-2001: U.S. Geological Survey Scientific Investigations Report 2006-5117, 12 p.

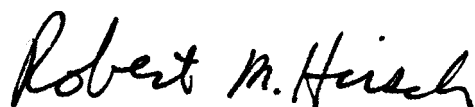
## FOREWORD

The U.S. Geological Survey (USGS) is committed to serve the Nation with accurate and timely scientific information that helps enhance and protect the overall quality of life, and facilitates effective management of water, biological, energy, and mineral resources. (<http://www.usgs.gov/>). Information on the quality of the Nation's water resources is of critical interest to the USGS because it is so integrally linked to the long-term availability of water that is clean and safe for drinking and recreation and that is suitable for industry, irrigation, and habitat for fish and wildlife. Escalating population growth and increasing demands for the multiple water uses make water availability, now measured in terms of quantity and quality, even more critical to the long-term sustainability of our communities and ecosystems.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program to support national, regional, and local information needs and decisions related to water-quality management and policy. (<http://water.usgs.gov/nawqa/>). Shaped by and coordinated with ongoing efforts of other Federal, State, and local agencies, the NAWQA Program is designed to answer: What is the condition of our Nation's streams and ground water? How are the conditions changing over time? How do natural features and human activities affect the quality of streams and ground water, and where are those effects most pronounced? By combining information on water chemistry, physical characteristics, stream habitat, and aquatic life, the NAWQA Program aims to provide science-based insights for current and emerging water issues and priorities. NAWQA results can contribute to informed decisions that result in practical and effective water-resource management and strategies that protect and restore water quality.

Since 1991, the NAWQA Program has implemented interdisciplinary assessments in more than 50 of the Nation's most important river basins and aquifers, referred to as Study Units. (<http://water.usgs.gov/nawqa/nawqamap.html>). Collectively, these Study Units account for more than 60 percent of the overall water use and population served by public water supply, and are representative of the Nation's major hydrologic landscapes, priority ecological resources, and agricultural, urban, and natural sources of contamination.

Each assessment is guided by a nationally consistent study design and methods of sampling and analysis. The assessments thereby build local knowledge about water-quality issues and trends in a particular stream or aquifer while providing an understanding of how and why water quality varies regionally and nationally. The consistent, multi-scale approach helps to determine if certain types of water-quality issues are isolated or pervasive, and allows direct comparisons of how human activities and natural processes affect water quality and ecological health in the Nation's diverse geographic and environmental settings. Comprehensive assessments on pesticides, nutrients, volatile organic compounds, trace metals, and aquatic ecology are developed at the national scale through comparative analysis of the Study-Unit findings. (<http://water.usgs.gov/nawqa/natsyn.html>).



Robert M. Hirsch  
Associate Director for Water

## Contents

Abstract.....	1
Introduction.....	1
Purpose and Scope .....	1
Description of the Study Area .....	2
Previous Investigations.....	2
Acknowledgments.....	3
Methods.....	3
Sample Collection.....	4
Taxonomic Analysis.....	4
Data Analysis.....	4
Comparison of Macroinvertebrate Community Structure .....	6
Initial Data Set.....	6
Taxa Reconciliation .....	6
Adjustment for Subsampling.....	8
Synopsis of Comparisons .....	9
Summary and Conclusions.....	9
References Cited.....	11

## Figures

1. Map showing locations of sampling sites, NAWQA-EMAP comparative study, Wyoming, Colorado, and Montana, 2000-2001.....	2
2-4. Graphs showing:	
2. Selected metrics of macroinvertebrate community structure and scores from the Wyoming Stream Integrity Index for 12 paired NAWQA and EMAP samples, comparative study, Wyoming, Colorado, and Montana, 2000-2001. ....	7
3. Total taxa richness in NAWQA and EMAP data sets after taxa reconciliation and subsampling, comparative study, Wyoming, Colorado, and Montana, 2000-2001.....	9
4. Non-metric multi-dimensional plots of macroinvertebrate community structure, comparative study, Wyoming, Colorado, and Montana, 2000-2001. ....	10

## Tables

1. Macroinvertebrate sampling sites in comparative study, Wyoming, Colorado, and Montana, 2000-2001.....	3
2. Metrics used in the Wyoming Stream Integrity Index to compute scores for mountain and non-mountain streams, comparative study, Wyoming, Colorado, and Montana, 2000-2001.....	6

## Conversion Factors

<b>Multiply</b>	<b>By</b>	<b>To obtain</b>
centimeter (cm)	0.3937	inch (in.)
micron	0.00003937	inch (in.)
millimeter (mm)	0.0394	inch (in.)
meter (m)	3.281	foot (ft)
square kilometer (km <sup>2</sup> )	0.3861	square mile (mi <sup>2</sup> )
square meter (m <sup>2</sup> )	10.76	square foot (ft <sup>2</sup> )

## Acronyms used in this report

EMAP	Environmental Monitoring and Assessment Program
EPT	Ephemeroptera, Plecoptera, and Trichoptera
IDAS	Invertebrate Data Analysis System
NAWQA	National Water-Quality Assessment Program
NMDS	Non-Metric Multi-dimensional Scaling
NWQL	National Water Quality Laboratory
RTH	Richest Targeted Habitat
USGS	U.S. Geological Survey
USEPA	U.S. Environmental Protection Agency
WDEQ	Wyoming Department of Environmental Quality
WSII	Wyoming Stream Integrity Index





# Comparison of Macroinvertebrate Community Structure between Two Riffle-Based Sampling Protocols in Wyoming, Colorado, and Montana, 2000-2001

By David A. Peterson<sup>1</sup> and Jeremy R. Zumberge<sup>2</sup>

## Abstract

Samples of benthic macroinvertebrates were collected side-by-side from riffles at 12 stream sites in Wyoming, Colorado, and Montana during 2000-2001, following protocols established by the U.S. Geological Survey National Water-Quality Assessment (NAWQA) Program and the U.S. Environmental Protection Agency Environmental Monitoring and Assessment Program (EMAP). Samples from riffles were collected following NAWQA protocols, using a sampler with 425-micron net mesh-opening size from a total area of 1.25 m<sup>2</sup> per sample in multiple riffles. Samples also were collected following EMAP protocols, using a sampler with 500-micron net mesh-opening size from a total area of 0.72 m<sup>2</sup> per sample in multiple riffles. The taxonomic identification and enumeration of the samples followed procedures established for each program. Benthic macroinvertebrate community structure was compared between the data sets using individual metrics, a multimetric index, and multivariate analysis.

Comparisons between the macroinvertebrate community structures were made after sequentially adjusting both data sets for: (1) ambiguous taxa, (2) taxonomic inconsistencies, and (3) differences in laboratory subsampling. After removal of ambiguous taxa, pair-wise differences in total taxa richness and Ephemeroptera taxa richness were statistically significant ( $p < 0.05$ ). Differences between the data sets generally were not significant for richness of other taxa, tolerant taxa, semi-voltine taxa, functional feeding groups, diversity, and dominance. Sample scores calculated using the Wyoming Stream Integrity Index were not significantly different between the two data sets. After reconciling both data sets for taxonomic inconsistencies, total taxa richness and Ephemeroptera taxa richness remained significantly different between the data sets. After adjusting the data for differences in laboratory subsampling, the differences in taxa richness were no longer significant. Bray-Curtis similarity coefficients and non-metric multi-dimensional scaling were used to examine macroinvertebrate community structure. Similarity in community structure

between sites was affected to a greater extent by taxa reconciliation than by adjustment for subsampling.

## Introduction

Biological assessment has become a common component of many water-quality programs across the Nation. The growing interest in biological monitoring has led to interest in sharing data among agencies collecting biological monitoring data. The U.S. Geological Survey (USGS) NAWQA Program and the U.S. Environmental Protection Agency (USEPA) EMAP are examples of two national programs that collect biological monitoring data. Aquatic ecological measures are one of a variety of tools that NAWQA uses to assess the quality of the Nation's streams and rivers (<http://water.usgs.gov/nawqa/>), whereas EMAP is designed to provide status and analysis of the Nation's aquatic ecological resources (<http://www.epa.gov/emap/>). The ability to show compatibility of NAWQA and EMAP aquatic ecological data sets would strengthen the basis for data integration and comparability of the programs. A study was conducted by the USGS in cooperation with the Wyoming Department of Environmental Quality (WDEQ) to compare benthic macroinvertebrate community structure between the riffle-based sampling protocols of NAWQA and EMAP. For this study, macroinvertebrate data were collected side-by-side using both protocols at 12 stream sites in Wyoming, Colorado, and Montana during 2000-2001.

## Purpose and Scope

The purpose of this report is to present a comparison of macroinvertebrate community structure among paired samples collected using the riffle-based sampling protocols of NAWQA and EMAP. Data from macroinvertebrate samples collected side-by-side at 12 sites in Wyoming, Colorado, and Montana during 2000-2001 are compared by examining the initial data set using individual metrics and a multimetric index, followed by resolution of taxonomic differences between the laboratories, and then adjusting for differences in

<sup>1</sup>U.S. Geological Survey, Cheyenne, Wyoming

<sup>2</sup>Wyoming Department of Environmental Quality, Sheridan, Wyoming

## 2 Comparison of Macroinvertebrate Community Structure, Wyoming, Colorado, and Montana, 2000-2001

subsampling procedures. The effects of differences in sampling technique cannot be completely separated from laboratory effects because interlaboratory splits were beyond the scope of study.

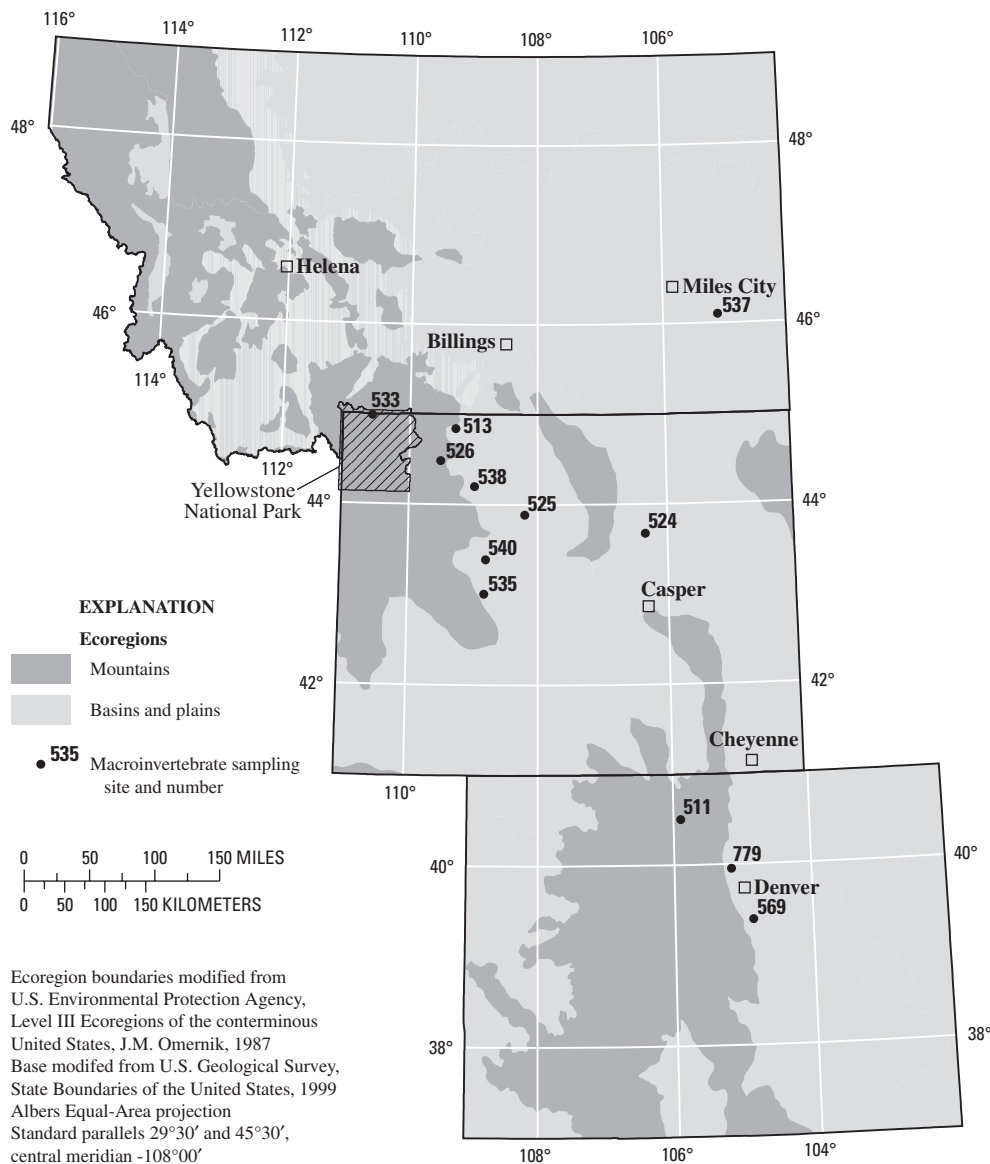
### Description of the Study Area

The environmental setting of the study area includes forested mountains and alpine zones, as well as shrub- and grasslands in the basins and plains of Wyoming, Colorado, and Montana (fig. 1). Mean annual precipitation ranges from about 150 mm in the driest parts of the basins and plains to more than 1,500 mm in the mountains near Yellowstone National Park (Zelt and others, 1999).

Streams and creeks that were sampled tended to be small, wadeable streams with riffle, run, and pool habitats. The mean wetted width of the streams at the time of sampling ranged from 1.7 m to 33.8 m (table 1); the median value was 5.0 m. The area of the drainage basins upstream from the sampling sites ranged from 13 to about 31,000 km<sup>2</sup>.

### Previous Investigations

Previous investigations have compared either NAWQA or EMAP benthic macroinvertebrate protocols to those of State or other agencies, but not to each other. For example, Justus and others (2001) described substantial differences in macroinvertebrate sampling methods and levels of taxonomic



**Figure 1.** Locations of sampling sites, NAWQA-EMAP comparative study, Wyoming, Colorado, and Montana, 2000-2001.

identification between NAWQA and State agencies from Missouri, Mississippi, and Arkansas. In spite of the procedural differences, five of nine metrics calculated from the respective data sets indicated comparable results. Differences among the other four metrics could be attributed, at least partially, to differences in net mesh size, number of organisms identified, and collection effort. Lenz and Miller (1996) noted water-quality ratings based on three metrics of environmental tolerance—the Hilsenhoff Biotic Index, the Family Biotic Index, and mean tolerance value—were similar among macroinvertebrate samples collected following NAWQA, State of Wisconsin, and U.S. Department of Agriculture-Forest Service protocols, but community structure varied significantly between the agencies' samples.

In a study using split samples from eight NAWQA sites in Montana and Wyoming, Vicki Watson (University of Montana, written commun., 2000) noted interlaboratory differences between the USGS laboratory and a contract laboratory had a greater effect on results than differences in field methods. Maret and others (2001) noted differences in total taxa richness and richness of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa were substantially greater in two interlaboratory (between the USGS laboratory and a contract laboratory) split samples than in two intralaboratory (within a laboratory) split samples.

## Acknowledgments

The efforts of Ron Zelt in the design and implementation of the early phases of this study are gratefully acknowledged. Cathy Tate graciously provided data from sampling efforts in Colorado. Daren Carlisle, Bob Zuellig, and Steve Moulton provided advice on taxa reconciliation. Colleague reviews by Daren Carlisle and Terry Maret led to many improvements in the report, and we also thank Karl Hermann from USEPA Region 8 for his review. The assistance of USGS personnel from Wyoming and Colorado, including Peter Wright and Greg Boughton, and Bill Schroeder from USEPA Region 8, in collection of the samples also is greatly appreciated.

## Methods

NAWQA protocols specify collection of two types of macroinvertebrate samples (Cuffney and others, 1993). The first type is a richest targeted habitat (RTH) sample, which is the type considered in this report, and the second type is a qualitative multi-habitat sample that is not discussed further in this report. Similarly, the EMAP protocols (Peck and others, 2000) specify collection of two types of macroinvertebrate samples; the first type is the targeted riffle sample described in this report and the second type is a transect-based, reach-wide sample that is not discussed further in this report.

**Table 1.** Macroinvertebrate sampling sites in comparative study, Wyoming, Colorado, and Montana, 2000-2001.

USGS site identification number	EMAP site identification number <sup>1</sup>	Stream name	Mean wetted width (meters)	Watershed area (square kilometers)
403012105533301	WCOP99-0511	Michigan River, CO	4.1	13
445032109164701	WWYP99-0513	Clarks Fork Yellowstone River, WY	33.8	1,997
434037106203001	WWYP99-0524	Salt Creek, WY	9.6	2,084
435313108115001	WWYP99-0525	Cottonwood Creek, WY	6.6	1,064
442859109302901	WWYP99-0526	Big Creek, WY	5.1	65
445854110353201	WWYP99-0533	Blacktail Deer Creek, WY	5.6	83
430029108491801	WWYP99-0535	Trout Creek, WY	3.2	120
460600105073801	WMTP99-0537	Powder River, MT	25.2	30,548
441202108584401	WWYP99-0538	Meeteetse Creek, WY	4.8	109
432323108474901	WWYP99-0540	Fivemile Creek, WY	1.7	114
392320104520001	WCOP99-0569	East Plum Creek, CO	3.3	272
395707105100401	WCOP01-0779	Coal Creek, CO	3.4	475

<sup>1</sup>Last 3 digits of EMAP number correspond to site numbers in figure 1.

## Sample Collection

Invertebrates were collected side-by-side following the NAWQA RTH protocol and the EMAP targeted riffle protocol at a total of 12 Western Pilot EMAP sites in Wyoming, Montana, and Colorado. The sites sampled for this study were chosen more or less randomly from a larger set of EMAP sites selected on a probabilistic design for sampling in the western United States (<http://www.epa.gov/emap/west/index.html>). At each site in this study, the samples were collected by USGS personnel that had been trained in use of both protocols.

Both the NAWQA and EMAP protocols specify collection of macroinvertebrates from multiple riffles where available within the sample reach. Each protocol specifies the collection of multiple samples that are then composited into one sample per site. The approach for both methods was to place the sampler frame firmly on the substrate, scrub individual rocks and other large items either by hand or with a vegetable brush, and then kick or stir the substrate vigorously for 30 seconds, so that the macroinvertebrates were washed downstream into the net. Where practical, the NAWQA and EMAP macroinvertebrate sample-collection nets were placed side-by-side to maximize comparability of the samples.

Macroinvertebrates were collected from five points at each site following the NAWQA protocol, using a Slack sampler with a 425-micron net mesh-opening size. The area sampled at each of the five points was 0.25 m<sup>2</sup>, which gives a total area of 1.25 m<sup>2</sup> per sample after compositing. The sample was elutriated to remove rocks and other large debris (Cuffney and others, 1993) and preserved with 10-percent buffered formalin.

Macroinvertebrates were collected from eight points at each site following the EMAP protocol, using a D-frame sampler that was 30.5-cm wide and with a 500-micron net mesh-opening size. The area sampled at each of the eight points was 0.09 m<sup>2</sup>, which gives a total area of 0.72 m<sup>2</sup> after compositing. The EMAP macroinvertebrate samples were preserved with 95-percent ethanol. Sample points in riffles were chosen at random where possible, but the small size of most of the streams and the large number of sample points often led to sampling nearly the entire riffle habitat in the reaches.

Quality-control samples collected for the study consisted of two splits and two replicate samples. At one site, each of the NAWQA and EMAP samples were split in the field and sent to the respective laboratories (intralaboratory splits). Ideally, all of the samples would have been split between the two taxonomic laboratories as interlaboratory splits but cost limitations prohibited that approach. Replicate samples were collected at one site and sent to the respective laboratories.

## Taxonomic Analysis

Macroinvertebrate samples collected following NAWQA protocols were sent to the Biological Group of the USGS National Water Quality Laboratory (NWQL) for taxonomic

identification and enumeration following methods described by Moulton and others (2000). The target count was a minimum of 300 organisms per sample; where organisms were more numerous, a subsampling frame was used to select macroinvertebrates. The remainder of each sample was then searched for an additional 15 minutes to find large or rare organisms that might have been missed in the initial sorting. The samples were sorted under a dissecting microscope with 10X magnification for a maximum of 8 hours. Insects usually were identified to genus or species, whereas higher taxonomic levels sometimes were reported for other organisms such as Acari (mites, class Arachnida). Levels of macroinvertebrate taxonomic identifications reported by the NWQL can be found in Moulton and others (2000, table 11). Quality-control procedures include having a second taxonomist check the sorting for at least 10 percent of the time that the sample was originally sorted, and taxonomic verification of a random selection of 10 percent of the taxa identified by laboratory taxonomists on a weekly basis (Moulton and others, 2000, p. 20-29).

Macroinvertebrate samples collected following EMAP protocols were sent to a contractor (EcoAnalysts, Inc., Moscow, Idaho) for taxonomic identification. The target count was 500 organisms (minimum) per sample; subsampling was conducted with a caton and tray when organisms were more numerous (Gary Lester, EcoAnalysts, Inc., written commun., 2005). After the initial processing, each sample was scanned for 5 minutes to locate large or rare organisms. Quality-control procedures included having a second person check the sorting and taxonomic verification of 10 percent of the samples. Insects usually were identified to genus or species, whereas higher taxonomic levels sometimes were reported for other organisms such as Oligochaeta.

## Data Analysis

Data were tested for normality prior to statistical comparisons. The macroinvertebrate abundance data tended to be log-normally distributed and therefore, the abundance data were transformed to either logarithms or presence/absence before making comparisons. The macroinvertebrate metric data typically were not normally or uniformly distributed, so pair-wise statistical tests were computed using the Wilcoxon signed-ranks test. The Wilcoxon signed-ranks test is a non-parametric counterpart to the t-test for paired samples (Helsel and Hirsch, 2002, p. 137-155). The signed-ranks tests were run in the statistical software S-Plus (Insightful Corp. 2002), at a two-tailed significance level of 0.05 of a Type I error (Helsel and Hirsch, 2002, p. 97-113). The null hypothesis is that the data sets are not different from each other. Probability values (p) greater than 0.05 indicate the null hypothesis likely is true and should not be rejected, whereas p values less than 0.05 indicate the null hypothesis likely is false and the data sets are different from each other.

The NAWQA and EMAP data sets each contained a number of ambiguous taxa. For example, taxa could be reported to

the family, with genus and species level identifications from within the same family reported from a given sample. The question is whether these ambiguous taxa are unique taxa, or are the same taxa that were undistinguishable due to problems such as immature or damaged specimens. The software routines in the Invertebrate Data and Analysis System (IDAS) (Cuffney, 2003) were used to provide a uniform, unbiased method of resolving ambiguous taxa for both data sets. Two options were tested for resolving ambiguous taxa, based on relative abundance of the parent (higher level) taxa and the children (Cuffney, 2003, p. 44-46). Under the first option, either the children's abundances are added to the parent's abundance if the ambiguous parent's abundance is greater than the sum of the children's abundances, or, the children are retained and the parent's abundances are deleted if the ambiguous parent's abundance is less than the sum of the children's abundances. Under the second option, ambiguous parents are distributed among the children in accordance with the relative abundance of the children but no abundances are deleted. Both options were computed and tested separately for the NAWQA data set and the EMAP data set. The total number of taxa was not significantly different (Wilcoxon signed-ranks test,  $p > 0.05$ ) between the two options for either data set, but the density was significantly different ( $p < 0.05$ ) between the options for both data sets. For this report, ambiguous taxa in both data sets were resolved by distributing the ambiguous parents to the children in accordance with relative abundance of the children to avoid unnecessary loss of information.

Macroinvertebrate metrics were calculated using the IDAS software (Cuffney, 2003). Selected metrics from the IDAS output were chosen to compare the data sets, based on their ecological relevance described in reports such as Barbour and others (1999) and Jessup and Stribling (2002). For example, total taxa richness is a measure of the diversity of invertebrates in the sample (Barbour and others, 1999), while the percent of the five dominant taxa (Jessup and Stribling, 2002) and the Shannon diversity index (Cuffney, 2003) represent different aspects of the sample diversity. Diversity is important because uneven distribution of taxa could indicate stress to the system. Taxa richness and relative abundance of EPT are of interest because those orders typically are associated with clean water, whereas non-insects, such as snails, tend to be more tolerant of degraded water quality. Tolerance values calculated in IDAS were derived in part from Barbour and others (1999) and in part from regional values for the northwestern U.S. (Cuffney, 2003) and unpublished data from the WDEQ. Those tolerance values are based on studies showing invertebrate response, by taxon, to organic pollution; higher tolerance values indicate more tolerant organisms. The semi-voltine taxa metric measures taxa that need multiple years to complete their life cycle, and generally are indicative of a stable environment (Jessup and Stribling, 2002). Variables affecting environmental stability include streamflow, bed-material size and mobility, and physicochemical condition. Metrics of functional feeding groups of macroinvertebrates (scrapers, filterer-collectors, and collector-gatherers) were tested because they can be

an important indicator of water quality where changes in water quality can affect the types of organisms present. For example, the abundance of scrapers is predicted to decrease in response to increasing perturbation (Barbour and others, 1999, p. 7-15) such as sediment deposition or water-quality degradation, both of which can affect periphyton production.

Scores from a multi-metric index, the Wyoming Stream Integrity Index (WSII) (Jessup and Stribling, 2002), were compared between the NAWQA and EMAP data sets. The WSII incorporates measures of taxa richness, composition, feeding group, tolerance, and voltinism into two different metric combinations—one for mountain streams and one for non-mountain streams (table 2). Metric scoring criteria were determined from reference stream data at the level III ecoregion-scale (Jessup and Stribling, 2002).

Similarity coefficients between the samples were computed pair-wise for all combinations of samples following methods described by Bray and Curtis (1957). Samples with more species in common received higher scores than those with fewer species in common. The Bray-Curtis similarity coefficients were used as input to non-metric multi-dimensional scaling (NMDS) ordinations (Clarke and Warwick, 2005) to show the relations of the macroinvertebrate community structure between the samples. The abundance data were log-transformed prior to calculation of the similarity coefficients because tests for normality indicated the macroinvertebrate abundance data generally were right-skewed. The abundance data also were transformed to presence/absence of taxa and re-analyzed with Bray-Curtis coefficients and NMDS ordinations. A stress level, or measure of the interpretive power, was calculated for each ordination. Lower levels of stress are desirable, as indicated by values of 0.2 or less, for example (Clark and Warwick, 2005, p. 5-6).

The taxonomic data from the NAWQA samples can be retrieved from the data warehouse: [http://infotrek.er.usgs.gov/doc/nawqa\\_www/bio/bio\\_communitysamples.htm](http://infotrek.er.usgs.gov/doc/nawqa_www/bio/bio_communitysamples.htm), using the site identification numbers listed in table 1. The taxonomic data from the EMAP samples currently (2005) are not publicly available, but the USEPA has announced plans to make the data publicly available in STORET (U.S. Environmental Protection Agency, 2004).

## 6 Comparison of Macroinvertebrate Community Structure, Wyoming, Colorado, and Montana, 2000-2001

**Table 2.** Metrics used in the Wyoming Stream Integrity Index (Jessup and Stribling, 2002) to compute scores for mountain and non-mountain streams, comparative study, Wyoming, Colorado, and Montana, 2000-2001.

Metric	Metric category	Mountain	Non-mountain
Total taxa	Richness		X
Ephemeroptera taxa	Richness	X	X
Plecoptera taxa	Richness	X	X
Trichoptera taxa	Richness	X	X
Percent Ephemeroptera (not including Baetidae)	Composition	X	
Percent Plecoptera	Composition		X
Percent Trichoptera (not including Hydropsychidae)	Composition	X	X
Percent non-insects	Composition		X
Percent five dominant taxa	Composition	X	
Percent scrapers	Feeding group	X	X
BCI CTQa	Tolerance	X	X
Hilsenhoff biotic index	Tolerance	X	
Semi-voltine taxa	Voltinism		X

### Comparison of Macroinvertebrate Community Structure

After resolving ambiguous taxa, the NAWQA data set contained 165 unique taxa, and the EMAP data set contained 193 unique taxa. These were the primary data sets used in subsequent comparisons unless specified otherwise.

#### Initial Data Set

The initial NAWQA and EMAP data sets, after removal of ambiguous taxa, were compared using common metrics, such as taxa richness of EPT and tolerance values. Less common metrics such as taxa richness of Chironomids (Diptera: Chironomidae) and Oligochaeta, and density of macroinvertebrates also were compared on the premise that the difference in mesh size of the sample collection devices used in the NAWQA and EMAP protocols could affect the data sets.

Total taxa richness and Ephemeroptera taxa richness were significantly different ( $p < 0.05$ ) between the NAWQA and EMAP data sets, and tended to be higher in the EMAP samples than in the NAWQA samples (fig. 2A and 2B). Trichoptera and Chironomid taxa richness also tended to be higher in the EMAP data set than in the NAWQA data set (fig. 2B and 2C), but were not significantly different ( $p > 0.05$ ). Taxa richness of Plecoptera was not significantly different between the data sets ( $p > 0.05$ ). Oligochaeta taxa richness (not shown) in the data sets was small compared to total taxa richness, but was significantly different between the data sets. Median

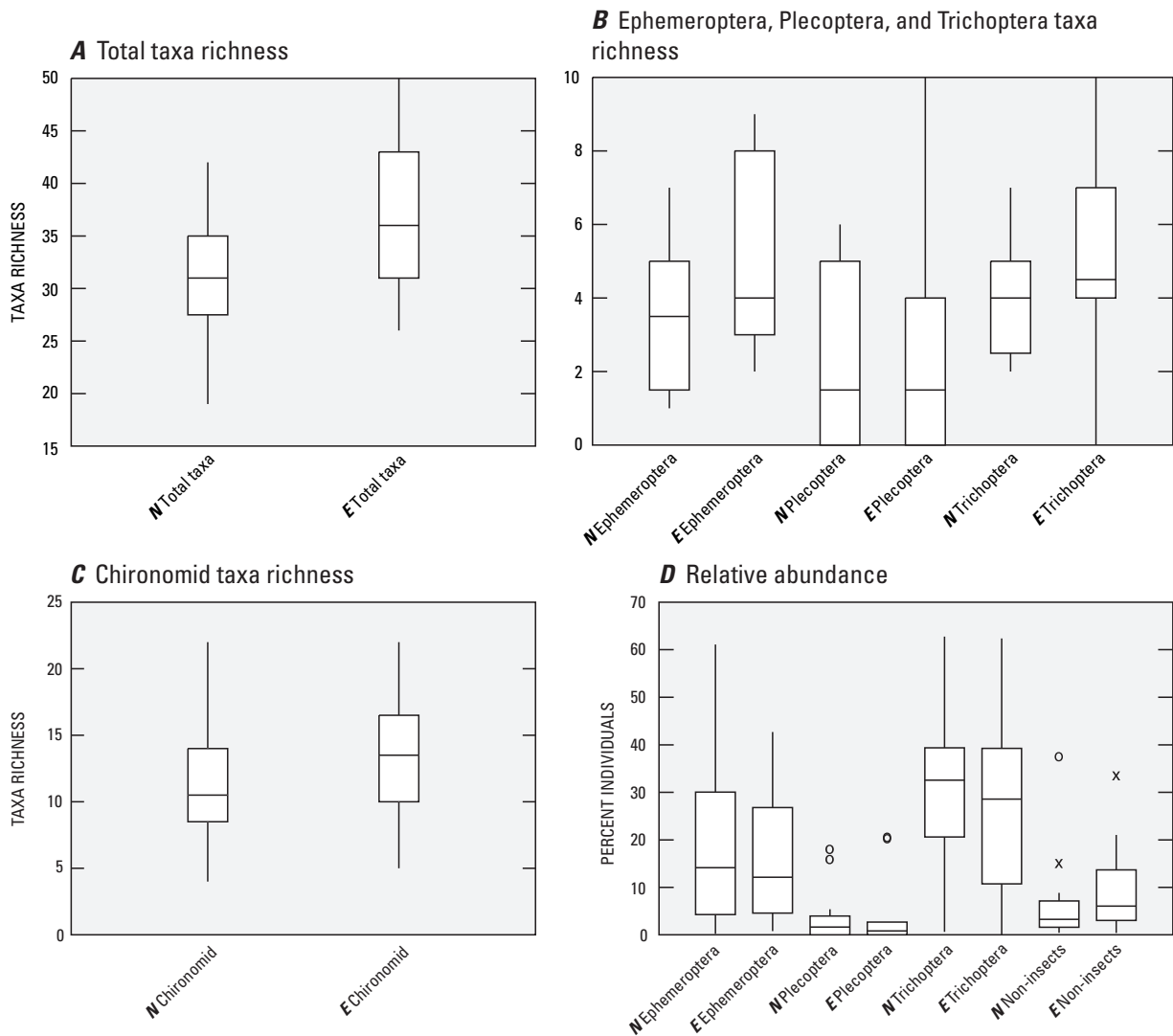
values of Oligochaeta taxa richness were 2.0 in the NAWQA samples and 1.0 in the EMAP samples.

Relative abundance of EPT and non-insects (fig. 2D), Chironomids, and Oligochaeta were not significantly different between the data sets. Other metrics tested that were not significantly different between the data sets ( $p > 0.05$ ) include tolerance values, number of semi-voltine taxa, abundance of the five dominant taxa, Shannon diversity, and relative abundances of the scraper, filter-collector, and collector-gatherer functional feeding groups. The density of macroinvertebrates per square meter was not significantly different between the two data sets (fig. 2E), in spite of differences in total area sampled and net-mesh size between the sampling protocols.

Invertebrate data in the initial data set were analyzed further through comparison of values calculated for a multi-metric index, the WSII. The WSII uses two combinations of metrics: one for mountain streams and one for non-mountain streams (table 2), in calculation of a score for each sample. The mountain stream metrics were computed for 4 of the 12 sites sampled in this study (sites 511, 513, 526, and 533), and the non-mountain metrics were computed for the other 8 sites. The computed scores for the WSII (fig. 2F) were not significantly different ( $p > 0.05$ ) between the NAWQA and EMAP data sets.

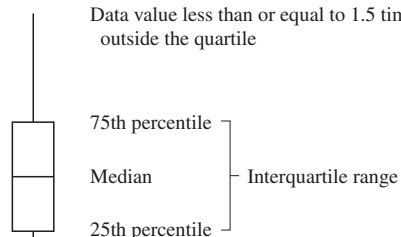
#### Taxa Reconciliation

The NAWQA and EMAP taxonomic laboratory procedures both specify identification of macroinvertebrates to species level where practical. A total of 29 taxa were identified to species in the NAWQA samples, and 35 taxa were identi-



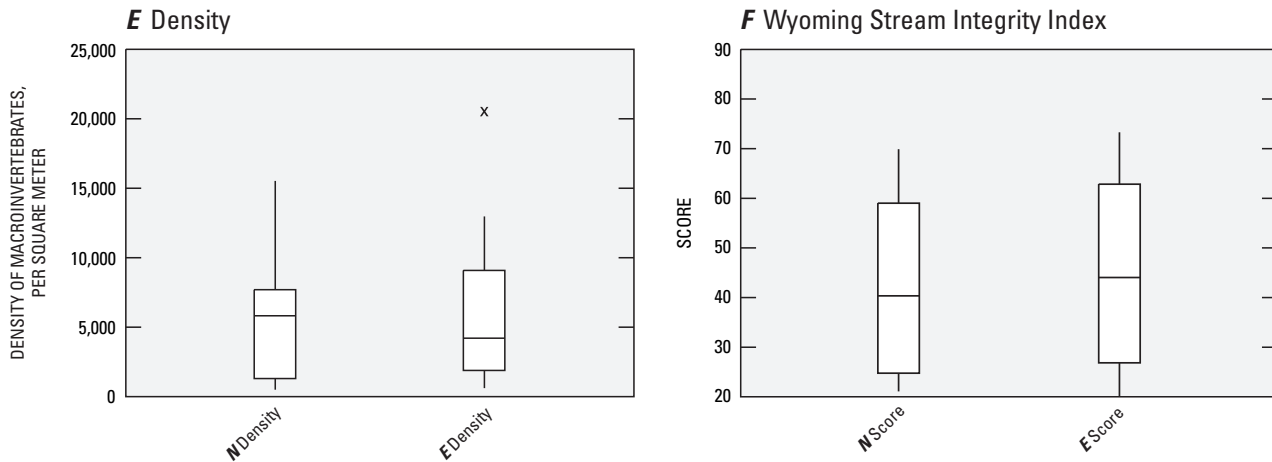
**EXPLANATION**

- o Outlier data value greater than 3 times the interquartile range outside the quartile
- x Outlier data value less than or equal to 3 and more than 1.5 times the interquartile range outside the quartile
- Data value less than or equal to 1.5 times the interquartile range outside the quartile



- N** National Water-Quality Assessment (NAWQA) Program samples
- E** Environmental Monitoring and Assessment Program (EMAP) samples

**Figure 2.** Selected metrics of macroinvertebrate community structure and scores from the Wyoming Stream Integrity Index for 12 paired NAWQA and EMAP samples, comparative study, Wyoming, Colorado, and Montana, 2000-2001.



**Figure 2.** Selected metrics of macroinvertebrate community structure and scores from the Wyoming Stream Integrity Index for 12 paired NAWQA and EMAP samples, comparative study, Wyoming, Colorado, and Montana, 2000-2001.—Continued

fied to species in the EMAP samples. Non-insects received different levels of taxonomic resolution between the laboratories. The NWQL laboratory identified mites to Acari (class Arachnida), Oligochaeta to family, and discarded Ostracoda, whereas the EMAP laboratory identified mites to genus, Oligochaeta to class, and counted Ostracoda. Some differences in taxonomic professional judgment also appeared, for example, where one laboratory identified *Gammarus* at a site, the other laboratory identified *Crangonyx* at the same site. Another example is where one laboratory identified caddisflies at a site as belonging to the *Rhyacophila brunnea* group, and the other laboratory reported *R. brunnea* group, *R. coloradensis* group, *R. hyalinata* group, and *R. valuma/pellisa* at the same site. As a test of the effect of the laboratory procedures on the results, both data sets were reconciled to set genus as the minimum level of identification, all Oligochaeta and mites were lumped to one respective taxon each, and Ostracoda were deleted. The total number of taxa in each reconciled data set was reduced from those of the initial data set, but the difference in total taxa richness (fig. 3) and Ephemeroptera taxa richness (not shown) between the reconciled data sets was still significant ( $p < 0.05$ ).

### Adjustment for Subsampling

The NAWQA laboratory uses a fixed-count target of 300 organisms whereas the EMAP laboratory uses a fixed-count target of 500 organisms. In order to test the effects of the different-sized fixed counts used by the laboratories, a software routine (Daren Carlisle, U.S. Geological Survey, written commun., 2005) was used to estimate taxa present if the EMAP samples were identified to 300 organisms instead of 500. Using a 300-organism, species-level count for EMAP and NAWQA data sets, the total number of taxa (fig. 3) and

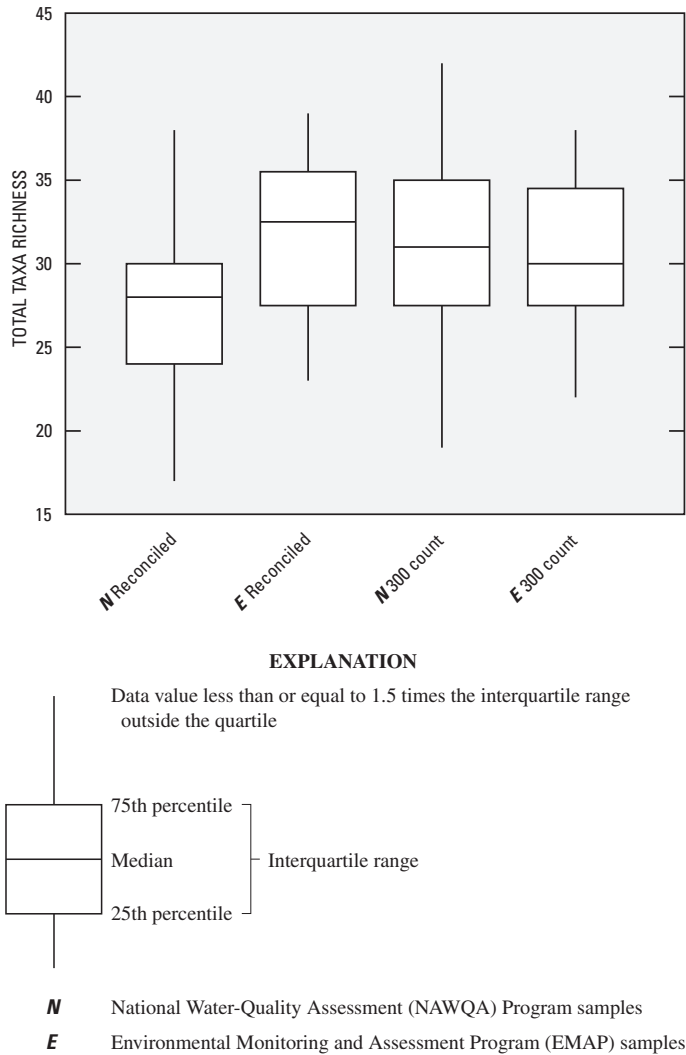
the Ephemeroptera taxa richness (not shown) were not significantly different between the data sets ( $p > 0.05$ ).

The effects of taxa reconciliation and subsampling on community structure also were tested using NMDS ordinations of Bray-Curtis similarity coefficients between the samples (Clarke and Warwick, 2005). In a NMDS ordination (fig. 4A), samples that are similar to each other will plot closer to each other than to other samples. For example, in the upper right portion of figure 4A, the community structure of the EMAP primary sample from sampling site 533 (533E) and the EMAP intralaboratory split sample (S) from site 533 (533ES) are more like each other than they are like other samples. Other intralaboratory quality-control samples—the NAWQA split sample from site 533 (533NS), and the NAWQA and EMAP replicate (R) samples from site 524 (524NR and 524ER)—also plotted relatively near to their respective primary sample.

Two patterns of community similarity in the species-level identifications of macroinvertebrate samples from NMDS ordinations of log-transformed macroinvertebrate abundance data are evident in figure 4A. First, all of the EMAP samples are plotted in the top half of figure 4A, and all of the NAWQA samples are plotted in the lower half. This pattern indicates the samples show some affinity due to their respective program, perhaps because of the differences in taxonomic identifications and subsampling between the programs. The second pattern is a gradient from left to right, corresponding to a gradient from plains to mountains streams, evident in both NAWQA and EMAP samples.

NMDS ordinations of log-transformed macroinvertebrate abundance data were tested for combinations of species-level data, before and after subsampling, and taxa reconciled data, before and after subsampling. All of the NMDS ordinations of log-transformed data showed a mirror image by program, similar to that shown in figure 4A, indicating greater affinity by program than by site.





**Figure 3.** Total taxa richness in NAWQA and EMAP data sets after taxa reconciliation and subsampling, comparative study, Wyoming, Colorado, and Montana, 2000-2001.

Additional testing was conducted using presence/absence of macroinvertebrate taxa as a more severe method of data transformation (Clarke and Warwick, 2005, p. 9-3). NMDS plots of presence/absence of taxa at the species level data (not shown), both before and after subsampling, did not change the pattern substantially from the log-transformed data shown in figure 4A. Use of presence/absence data changed the pattern substantially, however, for data after taxa reconciliation and subsampling.

The NMDS ordinations using presence/absence data after reconciliation of taxa showed a stronger affinity by site than by program (fig. 4B and 4C). The stress level of figure 4B (0.12, before subsampling) was the same as that of figure 4C (0.12, after subsampling), indicating no improvement in the interpretive power of the ordination. Of the various ordinations performed, taxa reconciliation of presence/absence data

(fig. 4B) provided the best indicator of similar macroinvertebrate community structure between the NAWQA and EMAP data sets.

## Synopsis of Comparisons

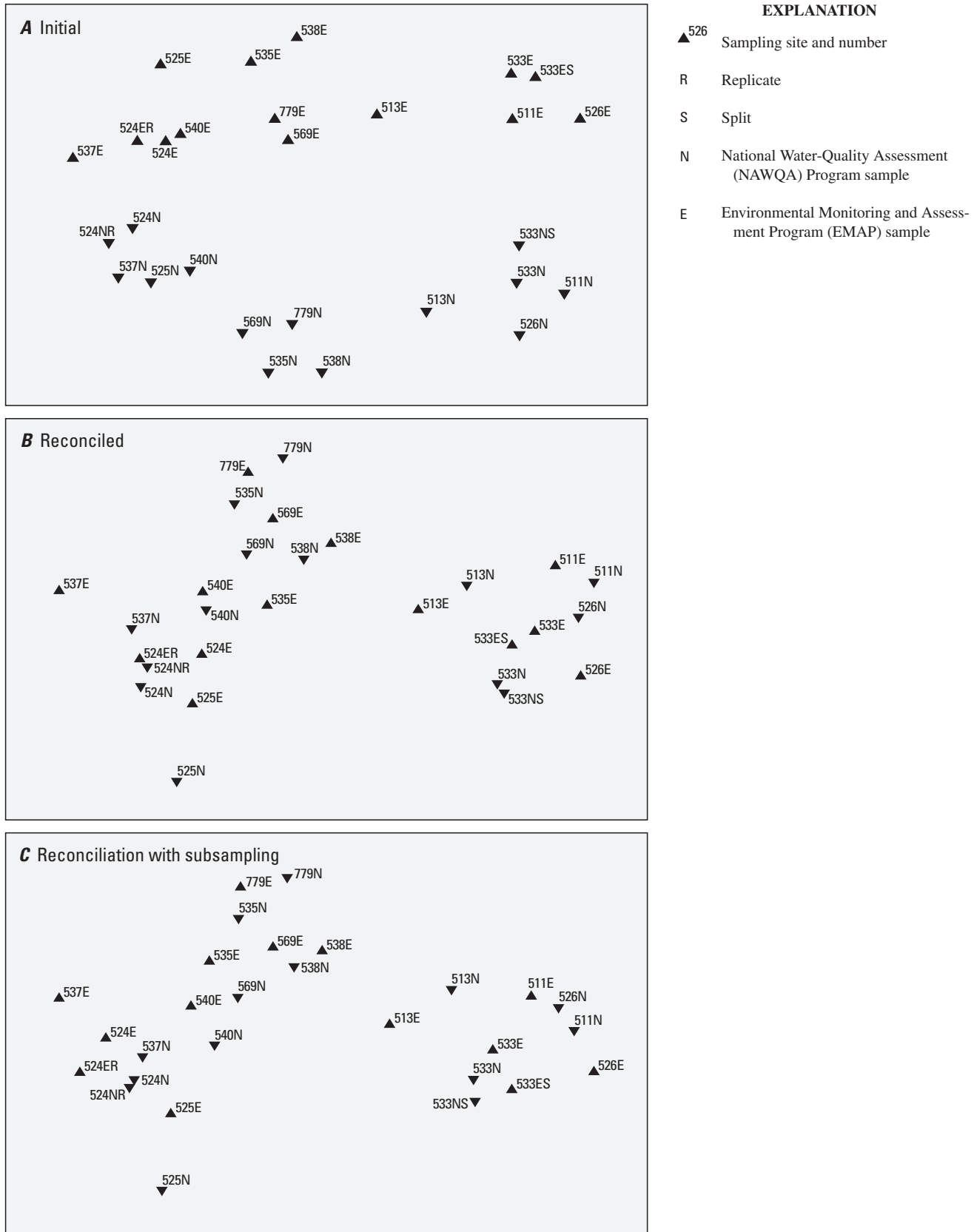
The ultimate goal of most comparative studies is to test the feasibility of combining data sets from different sources or protocols. This report presents methods that might be used to resolve differences between NAWQA and EMAP macroinvertebrate data prior to combining the data sets, but the small sample size and geographic area of this study limit the ability to extrapolate the results to other areas without additional study and further testing.

Reconciliation of taxa between the NAWQA and EMAP data sets resulted in substantial improvement of NMDS ordinations, but some metrics remained significantly different between the data sets. The reconciliation included lumping species-level identifications at the genus level. Elimination of species-level identifications also results in loss of information because water-quality tolerance values can vary greatly from genus to species level (Lenat and Resh, 2001).

Subsampling of the EMAP samples to a 300-organism count eliminated statistically significant differences in metric values, but did not improve the NMDS ordinations between the data sets. This finding is somewhat contradicted by previous studies, such as by Vinson and Hawkins (1996), where it was shown that most of the taxa that are present in a sample will be identified in the first 300 organisms or less, and new species identified per unit effort diminishes at more than 300 organisms. The difference between 300- and 500-count samples also reflects a difference in the amount of effort expended in the laboratory, and has cost and turn-around time implications to consider in designing aquatic ecology monitoring programs. Cao and others (2002) noted multivariate analyses, such as NMDS, can be unstable and are affected by sampling effort. A data transformation to presence/absence format was needed in the various NMDS ordinations to show greater similarity in community structure between samples than between the two programs. The data treatments used to minimize differences between the data sets include a loss of information—a loss of species-level data resulting from taxa reconciliation, and a loss of abundance data from the NMDS ordinations.

## Summary and Conclusions

A study was conducted to compare the benthic macroinvertebrate community structure in samples collected according to the riffle-based sampling protocols of the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program and the U.S. Environmental Protection Agency (USEPA) Environmental Monitoring and Assessment Program (EMAP). This study was conducted by the USGS in



**Figure 4.** Non-metric multi-dimensional plots of macroinvertebrate community structure, comparative study, Wyoming, Colorado, and Montana, 2000-2001. A, initial data set; B, after taxonomic reconciliation; and C, after taxonomic reconciliation and subsampling.

cooperation with the Wyoming Department of Environmental Quality (WDEQ). For the study, paired samples of benthic macroinvertebrates were collected from 12 stream sites in Wyoming, Colorado, and Montana. The taxonomic identification and enumeration of the samples followed procedures established for each program. The effects of differences in sampling technique cannot be separated completely from laboratory effects because interlaboratory splits were beyond the scope of this study. Benthic macroinvertebrate community structure was compared between the data sets using individual metrics, a multimetric index, and multivariate analysis.

Metrics of macroinvertebrate community structure were compared after adjusting both data sets for ambiguous taxa. Differences in total taxa richness and Ephemeroptera taxa richness between the data sets were statistically significant ( $p < 0.05$ ). Differences between the data sets generally were not significant for richness of other taxa, tolerant taxa, semi-voltine taxa, functional feeding groups, diversity, and dominance. Although the net mesh-opening size used in the NAWQA protocol (425 micron) was smaller than the net mesh-opening size used in the EMAP protocol (500 micron), no significant difference ( $p > 0.05$ ) was found in either Chironomid taxa richness or relative abundance of Chironomids between the data sets. NAWQA samples were collected from a larger area (1.25 m<sup>2</sup>) at each site than the EMAP samples (0.72 m<sup>2</sup>), but the density of macroinvertebrates was not significantly different ( $p > 0.05$ ) between the data sets. Sample scores calculated using a multi-metric index, the Wyoming Stream Integrity Index, were not significantly different between the two data sets ( $p > 0.05$ ).

Differences in taxonomic identifications between the data sets were partially reconciled by setting genus as the minimum level of identification, lumping Oligochaeta and mites to one respective taxon each, and deleting Ostracoda. After the taxa reconciliation, total taxa richness and Ephemeroptera taxa richness remained significantly different ( $p < 0.05$ ) between the data sets. Reconciliation did, however, result in substantial improvement in non-metric multi-dimensional scaling (NMDS) ordinations of presence/absence data. Macroinvertebrate community structure was similar between paired NAWQA and EMAP samples in NMDS ordinations of reconciled presence/absence data.

Subsampling of EMAP samples from a fixed-count target of 500 organisms to a 300-count equivalent of NAWQA samples eliminated any significant differences ( $p > 0.05$ ) between the data sets for the metrics tested. The subsampling did not improve the NMDS ordination, however, beyond that achieved with taxa reconciliation.

The similarity in results presented in this report is encouraging in terms of demonstrating the potential to aggregate data collected by two large-scale national programs using similar protocols. The potential benefits of larger, combined data sets must be weighed by the user against the loss of information associated with taxa reconciliation and subsampling. The size of the data set from this study and the geographic area are too small to extrapolate the results to the national

scale; however, the data do indicate that additional study and further testing are warranted.

## References Cited

- Barbour, M.T., Gerritsen, J., Snyder, B.D., and Stribling, J.B., 1999, Rapid bioassessment protocols for use in streams and wadeable rivers—Periphyton, benthic macroinvertebrates and fish (2d ed.): U.S. Environmental Protection Agency, Office of Water, EPA 841-B-99-002, variable pagination.
- Bray, J.R., and Curtis, J.T., 1957, An ordination of the upland forest communities of southern Wisconsin: *Ecological Monographs*, v. 27, p. 325-349.
- Cao, Y., Larsen, D.P., Hughes, R.M., Angermeier, P.L., and Patton, T.M., 2002, Sampling effort affects multivariate comparisons of stream assemblages: *Journal of the North American Benthological Society*, v. 21, no. 4, p. 701-714.
- Clarke, K.R., and Warwick, R.M., 2005, Change in marine communities—an approach to statistical analysis and interpretation (2d ed.): Plymouth, United Kingdom, Primer-E Ltd., variable pagination.
- Cuffney, T.F., 2003, User's manual for the National Water-Quality Assessment Program Invertebrate Data Analysis System (IDAS) software, version 3: U.S. Geological Survey Open-File Report 03-172, 103 p.
- Cuffney, T.F., Gurtz, M.E., and Meador, M.R., 1993, Methods for collecting benthic macroinvertebrate samples as part of the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 93-406, 66 p.
- Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, Book 4, Hydrologic analysis and interpretation, 510 p.
- Jessup, B.K., and Stribling, J.B., 2002, Further evaluation of the Wyoming Stream Integrity Index, considering quantitative and qualitative reference site criteria: Owing Mills, Md., Tetra Tech, Inc., variable pagination.
- Justus, B.G., Bray, D.G., Dossett, A., Hicks, M., Sarver, R., and Rogers, M., 2001, Comparison of multi-habitat aquatic macroinvertebrate sampling methods in streams of the Mississippi Alluvial Plain ecoregion: U.S. Geological Survey Water-Resources Investigations Report 00-4216, 16 p.
- Insightful Corp., 2002, S-Plus version 6.1, professional edition: Seattle, WA.

## 12 Comparison of Macroinvertebrate Community Structure, Wyoming, Colorado, and Montana, 2000-2001

Lenat, D.R., and Resh, V.H., 2001, Taxonomy and stream ecology—the benefits of genus- and species-level identifications: *Journal of the North American Benthological Society*, v. 20, no. 2, p. 287-298.

Lenz, B.N., and Miller, M.A., 1996, Comparison of aquatic macroinvertebrate samples collected using different field methods: U.S. Geological Survey Fact Sheet FS-216-96, 4 p.

Maret, T.R., MacCoy, D.E., Skinner, K.D., Moore, S.E., and O'Dell, I., 2001, Evaluation of macroinvertebrate assemblages in Idaho rivers using multimetric and multivariate techniques, 1996-1998: U.S. Geological Survey Water-Resources Investigations Report 01-4145, 69 p.

Moulton, S.R., Carter, J.L., Grotheer, S.A., Cuffney, T.F., and Short, T.M., 2000, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—processing, taxonomy, and quality control of benthic macroinvertebrate samples: U.S. Geological Survey Open-File Report 00-212, 49 p.

Omernik, J.M., 1987, Ecoregions of the conterminous United States: *Annals of the Association of American Geographers*, v. 77, no.1, p. 118-125, 1 pl., scale 1:7,500,000.

Peck, D.V., Lazorchak, J.M., and Klemm, D.V.J., eds., 2000, Environmental Monitoring and Assessment Program—surface waters—Western Pilot Study field operations manual for Wadeable streams: Washington D.C., U.S. Environmental Protection Agency, 230 p., available at <http://www.epa.gov/emap/html/pubs/docs/groupdocs/surfwatr/field/ewwsm01.html>.

U.S. Environmental Protection Agency, 2004, The Wadeable streams assessment: EPA-841-F-04-002, 2 p.

Vinson, M.R., and Hawkins, C.P., 1996, Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams: *Journal of the North American Benthological Society*, v. 15, no. 3, p. 392-399.

Zelt, R.B., Boughton, G.K., Miller, K.A., Mason, J.P., and Gianakos, L.M., 1999, Environmental setting of the Yellowstone River Basin, Montana, North Dakota, and Wyoming: U.S. Geological Survey Water-Resources Investigations Report 98-4269, 112 p.

For more information concerning the research in this report, contact:

Director  
Wyoming Water Science Center  
U.S. Geological Survey  
2617 E. Lincolnway, Suite B  
Cheyenne, Wyoming 82001-5662  
(307) 778-2931  
<http://wy.water.usgs.gov/>



