

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/228608811>

# After Site Selection and before Data Analysis: Sampling, Sorting, and Laboratory Procedures Used in Stream Benthic Macroinvertebrate Monitoring Programs by USA State ....

Article in *Journal of the North American Benthological Society* · December 2001

DOI: 10.2307/1468095

CITATIONS

195

READS

1,042

2 authors, including:



James L Carter

United States Geological Survey

46 PUBLICATIONS 1,250 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Metal transport in the Sacramento River, California [View project](#)



Trace elements in aquatic biota [View project](#)

## BRIDGES

*BRIDGES is a recurring feature of J-NABS intended to provide a forum for the interchange of ideas and information between basic and applied researchers in benthic science. Articles in this series focus on topical research areas and linkages between basic and applied aspects of research, monitoring, policy, and education. Readers with ideas for topics should contact Associate Editors, Nick Aumen and Marty Gurtz.*

*Biomonitoring streams using macroinvertebrate community composition is an integral part of water-quality programs throughout the USA. Carter and Resh use a questionnaire-based survey to evaluate biomonitoring methods used by state programs. They conclude that most programs use a similar suite of techniques both in the field and in the laboratory; however, significant differences among the programs in the specifics of individual steps (e.g., mesh size used for collecting, number of organisms sorted) affect comparability among programs. Carter and Resh suggest a need for increased research to address methods issues, and provide a list of questions as a starting point for determining the influences of differences in methods on biomonitoring data.*

### **After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies**

JAMES L. CARTER<sup>1</sup>

*US Geological Survey, 345 Middlefield Road, MS 465, Menlo Park, California 94025 USA*

VINCENT H. RESH

*Environmental Science, Policy, and Management, 201 Wellman Hall, University of California, Berkeley, California 94702 USA*

**Abstract.** A survey of methods used by US state agencies for collecting and processing benthic macroinvertebrate samples from streams was conducted by questionnaire; 90 responses were received and used to describe trends in methods. The responses represented an estimated 13,000–15,000 samples collected and processed per year. Kicknet devices were used in 64.5% of the methods; other sampling devices included fixed-area samplers (Surber and Hess), artificial substrates (Hester–Dendy and rock baskets), grabs, and dipnets. Regional differences existed, e.g., the 1-m kicknet was used more often in the eastern US than in the western US. Mesh sizes varied among programs but 80.2% of the methods used a mesh size between 500 and 600  $\mu\text{m}$ . Mesh size variations within US Environmental Protection Agency regions were large, with size differences ranging from 100 to 700  $\mu\text{m}$ . Most samples collected were composites; the mean area sampled was 1.7  $\text{m}^2$ . Samples rarely were collected using a random method (4.7%); most samples (70.6%) were collected using “expert opinion”, which may make data obtained operator-specific. Only 26.3% of the methods sorted all the organisms from a sample; the remainder subsampled in the laboratory. The most common method of subsampling was to remove 100 organisms (range = 100–550). The magnification used for sorting ranged from 1 (sorting by eye) to 30 $\times$ , which results in inconsistent separation of macroinvertebrates from detritus. In addition to subsampling, 53% of the methods sorted large/rare organisms from a sample. The taxonomic level used for identifying organisms varied among taxa; Ephemeroptera, Plecoptera, and Trichoptera were generally identified to a finer taxonomic resolution (genus and species) than other taxa. Because there currently exists a large range of field and laboratory methods used by state programs, calibration among all programs to increase data comparability would be exceptionally challenging. However, because many techniques are shared among methods, limited testing could be designed to evaluate whether procedural differences affect the ability to determine levels of environmental impairment using benthic macroinvertebrate communities.

**Key words:** bioassessment, biomonitoring, macroinvertebrate, sampling, processing, methods, streams.

<sup>1</sup> E-mail address: jlcarter@usgs.gov

Benthic biologists choose study sites and plan data analyses based on study objectives. However, the steps in between—the collection of samples, the separation of organisms from the substrate, the level of identifications used—are often a product of tradition or convenience. Yet, decisions concerning the choice of sampling device, where to take samples, whether to subsample, and how to sort samples, may greatly influence study conclusions and subsequent management considerations.

Studies comparing various sampling devices or proposing new ones, along with descriptions of sorting techniques, were a staple of benthological publications and meeting presentations in the 1960s and 1970s. As the science has developed, more emphasis has been placed on experimental approaches and less on descriptive approaches. However, although methodological topics have become of less interest, they are certainly not of less importance.

We present the results of a questionnaire survey conducted among benthic biologists connected with US state agencies that use macroinvertebrates for stream biomonitoring. The purpose of the survey was to examine what procedures are used in both the field and the laboratory for collecting and processing benthic macroinvertebrate samples from streams. We discuss the responses of these agencies in terms of potential sources of bias and cost, and speculate on how the approaches used in this essential aspect of experimental design can affect the interpretation of results obtained and the conclusions reached.

### Methods

The questionnaire consisted of 4 parts (see the Appendix for the questions asked). The 1st part addressed procedures associated with sample collection. Questions ranged from what device is used for collecting macroinvertebrates, to specification of the length of stream that defines the sampled reach. The 2nd part addressed the processing of samples in the field. Questions ranged from whether the sample is sorted only by eye in the field, to the sample-preservation techniques used. The 3rd part addressed laboratory procedures. Questions asked whether and how samples are subsampled, and included an assessment of the level of taxonomic identification used. The final part of the questionnaire

assessed aspects related to an evaluation of data quality. Questions ranged from whether Quality Assurance/Quality Control procedures are used, to whether and how reference collections are maintained.

The data used for the analyses represent the responses from 48 states and the District of Columbia. Analyses were limited to US state programs because we suspected that their development was influenced by similar factors (i.e., both within-state and federal water-quality information needs). State programs were contacted based on the list of participants in Davis et al. (1996). Programs were requested to submit one questionnaire for each unique combination of field collecting method and field and/or laboratory processing method. Results are based on the responses from 90 questionnaires. Percentages were calculated based on the number of responses to each specific question. Because some questions eliminated the need to answer other questions, not all questions were answered on each questionnaire. In addition, some respondents did not answer all the possible questions. Therefore, the total number of responses varied among questions. Regional differences in methods were evaluated with ANOVA by partitioning states based on location. First, states were separated into groups east or west of the Mississippi River. Second, states were assigned to the 10 USEPA regions. Statistical analyses and percentages were calculated using STATISTICA (1999, STATISTICA for Windows, StatSoft Inc., Tulsa, Oklahoma).

Several previous studies have analyzed trends in methods used in benthic research (Winterbourn 1985, Voshell et al. 1989, Resh and McElravy 1993). The results of our survey are not directly comparable because the agency personnel we solicited are mandated to conduct bioassessments or do biomonitoring rather than research, *per se*. However, we recognize that it often is difficult to distinguish between specific aspects of the methods used in research studies addressing environmental influences on benthic communities and those used in biomonitoring.

### Results and Discussion

#### *Sample collection*

*Type of sampler used.*—It is often said that the number of different samplers used in studying

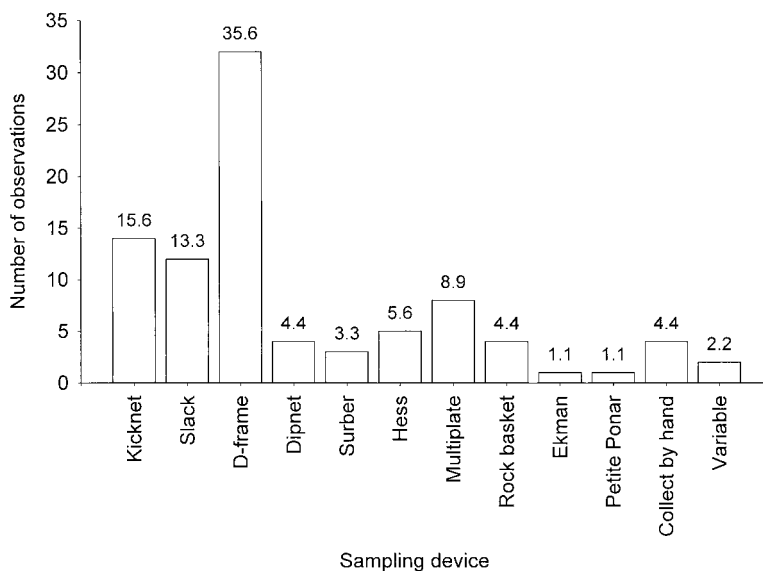


FIG. 1. Percentage (above each bar) of methods using each sampling device ( $n = 90$ ).

benthic macroinvertebrates in lotic systems is proportional to the number of investigators (Cummins 1962, Resh 1979, Merritt and Cummins 1996). However, this statement does not apply to sampling devices used in the state programs surveyed (Fig. 1). Kick-type devices, such as a D-frame sampler (0.3 m wide), Slack sampler (0.5 m wide), and/or kicknet (1 m wide), in which a predetermined effort (e.g., area of substrate, collection time, or number of sweeps or jabs) is used for sample collection comprised 64.5% of the devices. In contrast, the delineated, fixed-quadrat samplers such as the Surber and Hess comprised only 8.9% of the responses. Artificial substrates such as multiplates and rock baskets comprised 13.3% of the devices used.

Kick-type samplers have been the most commonly used devices in rapid bioassessment approaches (Resh and Jackson 1993), but they are far less used than fixed-quadrat samplers in the assessment of pollution, based on the results of earlier surveys (Winterbourn 1985, Voshell et al. 1989, Resh and McElravy 1993). Advantages of kick-type samplers include low cost, ease of transport, and usefulness in sampling a variety of habitats (including deep-water habitats) more easily than most commonly used fixed-quadrat samplers. The latter often are considered to have an advantage because numbers of organisms

collected can be related to a more precisely defined sample area, thereby providing an absolute (rather than relative) measure of density per taxon. Furthermore, fixed-quadrat replicates are considered to have higher precision as represented by lower coefficients of variation (but see Hornig and Pollard 1978).

*Mesh size.*—The size of the openings in the sampler net or in the sieves used for cleaning a sample determines the lower size limit of the organisms collected. However, the size of the mesh also affects the efficiency of the sampler (e.g., the backwash created if the mesh is too fine can sweep organisms out of the net). Our survey indicated that the most common mesh size used by US state agencies was 500  $\mu\text{m}$  (38.3% or 31/81 of respondents) followed by 600  $\mu\text{m}$  (21.0% or 17/81). However, the 500 to 600  $\mu\text{m}$  range covered 80.2% (65/81) of all programs (Fig. 2). Earlier surveys of benthic macroinvertebrate research studies in streams reported the use of smaller mesh sizes (201–300  $\mu\text{m}$ : Winterbourn 1985, 300–500  $\mu\text{m}$ : Resh and McElravy 1993). However, earlier surveys of lotic studies (Downing 1984) indicated mesh sizes of 450 to 600  $\mu\text{m}$  being used in lake investigations, which are close to the ranges reported in our survey.

More early instars of aquatic insects and

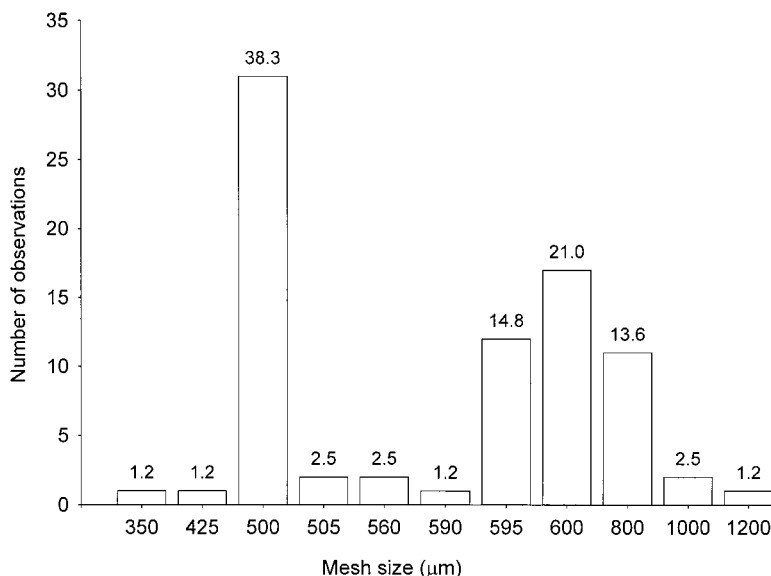


FIG. 2. Percentage (above each bar) of methods using each mesh size ( $n = 81$ ).

smaller organisms such as microcrustaceans are collected when netted collecting devices have finer meshes (Slack et al. 1991). However, these smaller organisms are often difficult for most benthic biologists to identify, and are simply included in taxa lists at higher taxonomic levels (family or even order). The inclusion (or exclusion) of early instars may greatly influence the results of some quantitative descriptions of the benthos (e.g., secondary production) (Huryn and Wallace 2000). In biomonitoring studies, metrics dependent on estimates of density or % composition also will be affected.

There is also an effective mesh size that varies over the time in which a sample is being collected. For example, small organisms that would not normally be retained given a certain mesh size, are retained as the net becomes filled with debris or clogged with silt. Different rates of clogging will lead to inconsistencies in the collection of organisms smaller than the mesh size.

*Compositing of samples.*—The formation of single sample units by compositing >1 collection was not done typically in earlier benthic studies (e.g., in only 10% of the studies reviewed by Resh and McElravy 1993). However, our survey indicated that compositing is now a common practice (74.4% or 64/86). The number of collections composited varied with different samplers

(Table 1). For example, survey results indicated that 3 collections were composited to form a single sample unit with Hess samplers, 3 to 5 with multiplates, 3 to 8 with Surber samplers, and 2 to 20 with D-frame samplers.

Most bioassessments described the benthos at a site using a single sample formed by collecting over a relatively large, contiguous area ( $\geq 1 \text{ m}^2$ ) or by combining several smaller noncontiguous collections. However, there was large variation in the area covered (Fig. 3). There are numerous advantages to large samples. They combine the assemblages representing individual microhabitat patches and reduce the high intersample (= intersite when a single sample represents a site) variation in benthic composition attributed to the often-described, high variance of small, individual, fixed-quadrat samples (e.g., Needham and Usinger 1956, Resh 1979, Merritt and Cummins 1996). In addition, samples composited from all available habitats within a reach may detect generalized habitat impairment.

The disadvantage of compositing is that a measure of within-site variance is lost, which limits the traditional use of inferential statistics for comparing among sites. However, individual replicates collected from within a single habitat (e.g., one riffle) are not true replicates but represent pseudoreplicates (Hurlbert 1984). Com-

TABLE 1. Number of samples composited per sampling device. SD = standard deviation. (No respondents reported compositing rock baskets.)

Sampling device	Number of responses	Mean number of samples	SD	Range
Kicknet	9	2.3	0.50	2–3
Slack	8	10.0	7.01	2–20
D-frame	18	12.6	8.37	2–20
Dipnet	1	18.0	0	18–18
Surber	3	4.7	2.89	3–8
Hess	2	3.0	0	3–3
Multiplate	2	4.0	1.41	3–5
Ekman	1	12.0	0	12–12

positing also tends to create large-volume samples that need to be subsampled.

*Length of stream section sampled.*—The length of stream (and the diversity of habitats) over which a sample is composited influences both the number of species encountered and the species-abundance distribution of the sample (Vinson and Hawkins 1996, Larsen and Herlihy 1998). Our data indicated that samples were collected over 3 principal lengths of stream: 1) the shortest length of stream sampled was when a single kick or Surber-type sample was collected from within a single habitat; 2) the next was when a composite was formed from  $\geq 2$  collec-

tions from  $\geq 1$  habitats within a single riffle-pool sequence; and 3) the largest was when a composite was formed from  $\geq 2$  collections from  $\geq 1$  habitats over distances of multiple riffle-pool sequences. We suspect most bioassessment programs expect that their sample represents lengths of stream that are at least the reach scale if not the segment scale (*sensu* Frissell et al. 1986).

The length of the stream section sampled was most often based on channel morphology; however, many programs collected their sample from a predetermined fixed length of stream. The most commonly sampled stream section

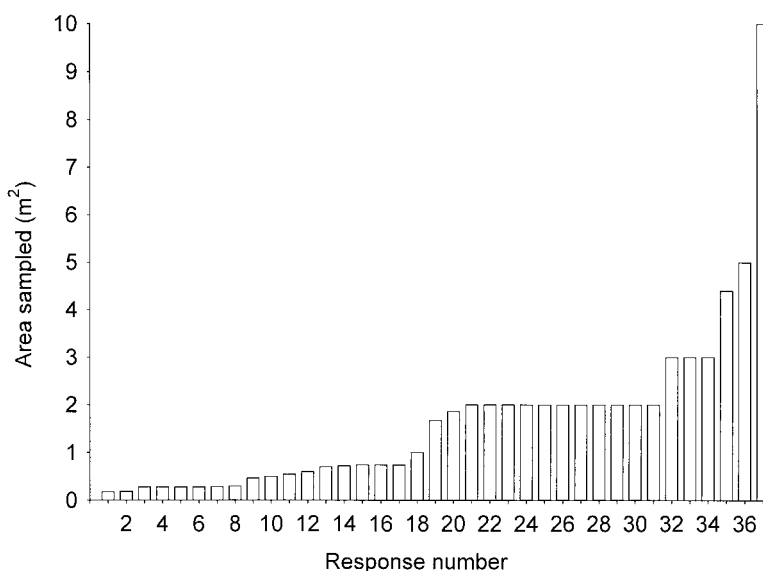


FIG. 3. Sequential frequency distribution of the total area sampled per site (from small to large area). Mean area sampled = 1.7 m<sup>2</sup>.

TABLE 2. Number and % (in parentheses) of composited samples collected from single compared to multiple habitat types and the method used to determine the length of stream over which the composite was collected.

Length of stream sampled	Types of habitats sampled		Totals
	Single	Multiple	
1 riffle-pool sequence	15 (26.8)	5 (8.9)	20 (35.7)
>1 riffle-pool sequence	2 (3.6)	3 (5.4)	5 (8.9)
Multiple channel widths	3 (5.4)	5 (8.9)	8 (14.3)
Fixed length of stream	9 (16.1)	9 (16.1)	18 (32.1)
Other	0 (0)	5 (8.9)	5 (8.9)
Totals	29 (51.8)	27 (48.2)	56 (100.0)

was a single riffle-pool sequence (35.7% or 20/56, Table 2); samples were composited over multiple riffle-pool sequences by 8.9% (5/56) of the programs. The 2nd most common method was to collect the sample over a fixed length of stream (32.1% or 18/56), such as 100 m. Only 14.3% (8/56) of the programs defined the section of stream to be sampled using multiples of the channel's width. When the section of stream over which the sample was collected exceeded 1 riffle-pool sequence, 2.8 riffle-pool sequences, 30 channel widths, or 126 m in fixed-channel length were the average lengths sampled.

The number of programs that based their sample on a composite from a single habitat type collected within a single riffle-pool sequence (15/20) was almost equal to those that formed a composite from a single habitat type over a section of stream greater than a single riffle-pool sequence (14/20). The distances over which samples were collected that exceeded a single riffle-pool sequence were longer than we expected and likely reflected attempts to avoid the atypical patterns possible if only a single habitat was examined from a single riffle-pool sequence, a topic that has been discussed as part of the pseudoreplication issue (Hurlbert 1984). Programs that sample macroinvertebrates over the same length of stream from which they sample fish and/or assess habitat also may select larger distances.

*Habitats sampled.*—Most sampling designs confined sampling to a single habitat type (63.4% or 52/82). Sampling only riffles (25.6%

or 21/82), followed by sampling “fast water habitats” (riffles and runs, 24.4% or 20/82), were the most common individual habitats sampled. However, 23% (19/82) of the programs attempted to sample all available individual habitats, usually in proportion to their occurrence (14.6% or 12/82). Riffle/run areas offer some level of standardization in terms of habitat stratification, which facilitates comparisons among sites. Also, these habitats are usually considered areas of highest species richness in riffle-pool dominated streams. Such stratification may reduce variation for intersite comparisons (Resh and Jackson 1993), but also may exaggerate differences if the strata are actually different among sites.

Collecting macroinvertebrates by sampling habitats in proportion to their occurrence forms a sample that is more representative of the organisms (and habitats) present in the sampled reach than collecting from a single habitat. However, replicability and, consequently, comparability among sites depends on the method used to define the % of individual habitats present in a reach, and is most likely operator dependent. Collecting from all available habitats produces a sample that contains more species and more life-history stages than collecting a sample from only a single habitat type; however, sampling multiple habitats in proportion to their % cover may be more difficult to standardize than collecting from a single habitat type.

*Placement of sampling devices.*—In what is cer-



tainly a departure from traditional random sampling, most state programs chose the spot to place the sampling device by using "expert opinion" (70.6% or 60/85) and rarely (4.7% or 4/85) by using random techniques (e.g., with a random numbers table and a mapped grid). *Expert* choice of a collecting location may lead to lower replicability among collectors (unless criteria are detailed and those choosing locations have sufficient training or experience). Errors in judgment about *best spots* also could rank 2 sites as being different or similar when they are not, simply because judgments about what are the *richest* sampling locations can be different. Although using expert opinion to choose sampling locations may maximize the number of species collected, it also may compromise the ability to obtain unbiased estimates of the relative composition of individual taxa or functional feeding groups among sites.

Nevertheless, the bias resulting from purposely choosing a sampling location is intentional: the programs using this approach are designed to collect in microhabitats with the highest species richness. Personnel thoroughly trained in the habitat requirements of benthic macroinvertebrates will most likely collect a sample that better represents the taxa present at a site by collecting expertly than by collecting randomly given the fixed, limited effort (fixed area sampled, fixed time for sampling, etc.) per site that is characteristic of most bioassessment sampling protocols.

*Replicates.*—Over ½ of the programs (56.1% or 46/82) collected replicates, with a mean of 2.3 replicates collected per site. However, most programs (67.4% or 31/46) collected only a single sample per site. The lack of sample replication in almost ½ of the programs is a major departure from sampling approaches in traditional benthic studies to assess pollution effects (Resh and McElravy 1993).

The question of replication has been hotly debated since the earliest day of benthic macroinvertebrate studies (Needham and Usinger 1956, Chutter and Noble 1966, Resh 1979) and continues to the present (Norris and Georges 1993, Norris et al. 1995, Merritt and Cummins 1996). In the studies reviewed by Resh and McElravy (1993), 3 to 5 replicates typically were taken. The survey of Winterbourn (1985) indicated that  $n \leq 5$  was the most common choice, and that all studies had at least some replication. Our sur-

vey indicated that if replicates were not collected at all sites, they often were collected at a % of the sites (2.5–55% of sites examined). Moreover, 45.7% (21/46) of the programs reported collecting replicates at 10% of the sites, as recommended by Plafkin et al. (1989) and Barbour et al. (1999). Although many programs collected replicates, they were often only collected for Quality Assurance/Quality Control, and generally were not processed and used in analyses.

### *Field Processing*

*Field sorting.*—Sorting (i.e., removing organisms from the sample matrix) was restricted to the field in 22.9% (19/83) of the programs and, of these, 26.3% (5/19) sorted a fixed number of organisms. The number of organisms ranged from 100 to 300, with a mean of 150. Seventy-five percent (12/16) of the programs that relied entirely on field sorting selected specimens to maximize an estimate of faunal richness. Two of these programs only sorted Ephemeroptera, Plecoptera, and Trichoptera (EPTs). Three programs with intensive field collections followed by laboratory processing also reported having reconnaissance programs where macroinvertebrates were only sorted in the field (some of these procedures were similar to the RBP-I protocol of Plafkin et al. 1989 or the BioRecon of Barbour et al. 1999).

Field sorting for assessing richness is highly dependent on an ability to distinguish among closely related taxa, and larger and smaller individuals of the same taxon, and is affected by ambient environmental factors (e.g., available light, weather conditions). Inconsistency among sorters is likely quite high, and quality control is also difficult. Clearly, field sorting to maximize faunal diversity is easier done than field sorting to produce proportionally representative counts of specific groups. The bias will be on large and/or rapidly moving taxa when sorting live material.

*Preliminary field processing.*—Of the groups that relied on sample processing in the laboratory, many removed large organic debris and stones (49.2% or 32/65) in the field, and some also elutriated (separated less dense organic material from more dense inorganic material) samples (20.0% or 13/65). In contrast, some programs did no field processing (27.7% or 18/65), and took the entire sample back to the labora-



tory. However, in  $\frac{1}{2}$  of these cases, artificial substrates were used, and the samples probably contained less extraneous detritus than kick-type samples.

The advantages of doing at least some processing in the field are to: 1) reduce the amount of material brought back (a particularly important problem if samples must be carried long distances); 2) enhance preservation of samples; 3) reduce damage to organisms; and 4) lessen laboratory sorting time. Undoubtedly, specimens are lost during field processing and at least some under-representation results from specimens adhering to discarded debris or from specimens remaining in the inorganic substrate following elutriation (e.g., cased caddisflies and mollusks).

*Field sieving.*—Most (56.7% or 38/67) respondents used sieves in the field to process samples. All programs reported using a sieve that was less than or equal to the size of the mesh used for collecting. Most programs (89.5% or 34/38) only used the specimens retained on the sieves in their analyses.

*Subsampling in the field.*—Field subsampling typically was not done by most programs. Of the 6 programs that subsampled in the field, 1 took  $\frac{1}{4}$  of the sample, 2 took random subsamples from a gridded tray, and 3 took a certain % of the specimens collected. Given the large area ( $\bar{x} = 1.7 \text{ m}^2$ ) over which specimens were collected by many programs, it was surprising that more programs did not subsample in the field.

If intersite comparisons are based on a fixed sample area, comparisons may be compromised if a program reduces sample volume by a variable amount among sites because richness is a function of area. However, because most programs only identified macroinvertebrates from a very small subsample of the original sample, the loss of some portion of the sample by field subsampling may have little effect.

*Preservatives.*—Most programs (64.1% or 50/78) preserved samples with ethyl alcohol, in a concentration range of 70 to 100% ( $\bar{x} = 85\%$ ). Some programs (17.9% or 14/78) used formalin. The mean concentration of formalin used was 10% and ranged from 5 to 30%.

It has long been known that alcohol preservatives cause a weight loss in specimens collected (e.g., Leonard 1939), but ethyl alcohol has benefits that make its use popular: it is relatively

inexpensive, widely available, and less unpleasant to use than formalin. It does, however, need to be periodically changed. The weight-loss issue is probably of minor importance because few programs use standing stock or biomass (or calculate secondary production estimates) as descriptors of macroinvertebrate communities. Conversely, not all organisms are best preserved initially in ethanol. For example, buffered formalin is the preferred initial preservative for oligochaetes because they tend to degrade if they are first preserved in alcohol (Klemm 1985).

The safe disposal of preservatives is a concern of industrial hygiene managers. Most of the respondents (64.5% or 49/76) disposed of preservatives by pouring them down the drain. This method of disposal is possible for some preservatives (e.g., ethyl alcohol with sufficient dilution), but not with others (e.g., formalin). Only 21.2% (16/76) of the programs used a hazardous waste processor to dispose of preservatives.

#### *Laboratory processing*

*Subsampling.*—Only 26.3% (19/72) of the respondents sorted all the organisms from a sample. Most (73.6% or 53/72) sorted a subsample of the entire sample. This result differed substantially from an earlier survey by Resh et al. (1985) in which  $\sim\frac{1}{2}$  of the respondents sorted all of the invertebrates from a sample. However, the mean area sampled by respondents in the Resh et al. (1985) survey was  $\sim 0.1 \text{ m}^2$ , over an order of magnitude smaller than the mean area sampled by the respondents to the present study (Fig. 3). Large-volume samples necessitate subsampling to make sample processing more economical. Little consensus has been reached on the effects of subsampling and the most effective methods for subsampling (Barbour and Gerritsen 1996, Courtemanch 1996, Vinson and Hawkins 1996, Grown et al. 1997, Cao et al. 1998, Larsen and Herlihy 1998).

Most programs subsampled by removing a fixed number of organisms, usually from 100 to 550 organisms (Fig. 4). A 100-organism subsample was the most common number (53.5% or 23/43 of the programs). The fixed number frequency distribution was somewhat bimodal because 300 organisms were subsampled in 25.6% of the protocols (Fig. 4). Eight programs subsampled by sorting a fixed proportion (most often  $\frac{1}{4}$ ) of the original sample. The device most

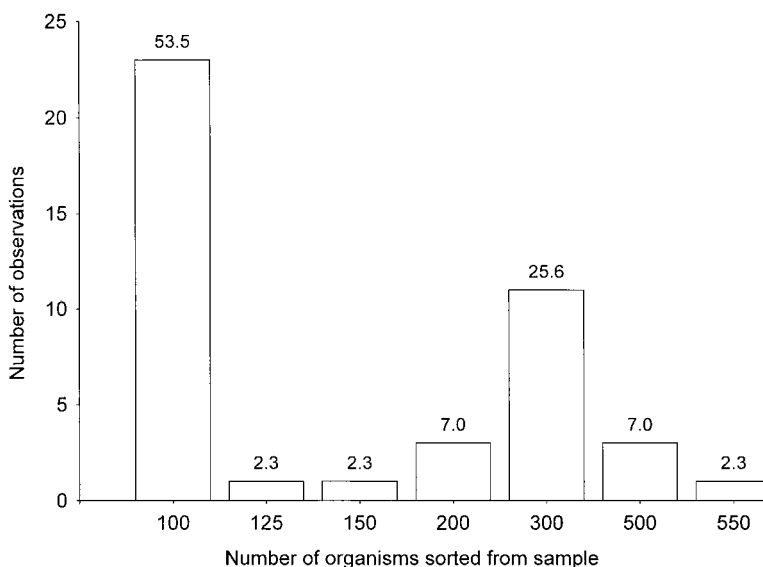


FIG. 4. Percentage (above each bar) of programs subsampling a fixed number of organisms from a sample ( $n = 43$ ).

commonly used for subsampling was a gridded tray or frame (Caton 1991, Moulton et al. 2000). Only 2 programs reported using an Imhoff cone (Wrona et al. 1982) for subsampling. Of the 43 programs that reported subsampling by acquiring a fixed number of organisms, 95% reported that they selected subsamples randomly.

Some programs (35.7% or 25/70) subsampled further when extremely abundant taxa were present. The taxonomic groups most often chosen for this procedure were chironomids and oligochaetes, but several respondents stated that they further subsampled any taxon that was abundant enough to limit economical processing.

*Sorting large/rare organisms.*—Vinson and Hawkins (1996) reported that sorting large/rare organisms increased estimated richness by ~28%. In our study, 53% (26/49) of respondents sorted large/rare organisms (e.g., see techniques presented in Vinson and Hawkins 1996, Moulton et al. 2000) from the sample. Of the 26 programs that sorted large/rare organisms, 35% sorted them before subsampling the original sample and 65% sorted them from the remainder of the sample after the subsample(s) was removed. Twenty-three programs did not sort large/rare organisms from the sample during any part of the processing. Although sub-

sampling randomly is less biased than sorting large/rare organisms, it seems reasonable to sort large/rare organisms in the laboratory, particularly when samples are collected using "expert opinion" to maximize estimated richness.

*Magnifications used for sorting.*—The magnification used for sorting macroinvertebrates in the laboratory ranged from 1 (sorting by eye) to 30 $\times$  (Fig. 5). Although a modal magnification of 10 $\times$  was used by 32.8% (22/67) of the programs, ~21% only sorted by eye. Furthermore, >50% of the programs used  $\leq 5\times$  magnification. Two respondents stated that they only sorted organisms that were greater than a given size (0.5 mm for one program and 1.2 mm for the other). If the objective is to sort *all* organisms collected, sorting by eye will result in fewer organisms being removed than when sorting under higher magnification.

*Taxonomic levels generally used in studies.*—The level of identifications for studies using the benthic community as an indicator of impairment is contentious. One view maintains that the lowest possible levels are necessary (i.e., species or the lowest possible taxon, usually a combination of genus and species levels) and the other view maintains that higher levels are sufficient (e.g., family) either because of similar information content at low and high levels or because of cost

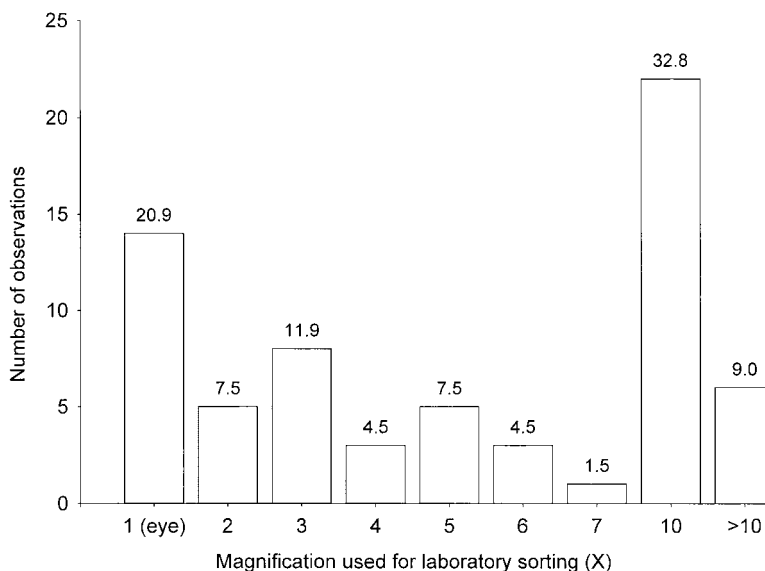


FIG. 5. Percentage (above each bar) of magnifications used in the laboratory for sorting macroinvertebrates ( $n = 67$ ).

considerations. The answer to this issue depends on several factors, some of which are intrinsic to the study area and objective. For example, is the taxa richness in an area low (so that families may only be represented by a single species) or high (having many confamilials and congeners)? What is the geographic extent of the study? Is the level of impairment expected to be large or small?

The type of analysis planned also influences this decision. For example, multivariate analyses probably show similar patterns between species and higher taxonomic levels because rare taxa are eliminated from an analysis (Lenat and Resh 2001). In reality, the laboratory procedure of counting fixed numbers of organisms (e.g., 100) has the same effect as deliberately removing or reducing rare taxa from an analysis, particularly if large/rare organisms are not sorted. Bailey et al. (2001) and Lenat and Resh (2001) discuss the pros and cons of various taxonomic resolution issues.

Even though there is much discussion on levels of identification to be used, it is clear that most US benthic macroinvertebrate biomonitoring programs use finer rather than coarser resolution in this regard. Typically, macroinvertebrates are identified to genus (41% or 30/73), genus and species (24.7% or 18/73), or species

(17.8% or 13/73). Only 16.4% (12/73) listed family level as typical, and 2 of these described this level as appropriate for reconnaissance studies.

*Taxonomic levels and specific taxa.*—Although genus and species were the typical taxonomic levels used in programs, there was variation in levels used with different taxonomic groups (Figs 6, 7). For example, microcrustaceans, mites, oligochaetes, and mollusks generally were identified at higher taxonomic levels (if at all, Fig. 6A–D), whereas Ephemeroptera, Plecoptera, Trichoptera, and Chironomidae were usually identified to lower taxonomic levels (Fig. 7A–D). Inclusion of some groups (but not others) and identification of groups to different taxonomic levels will affect calculation of various community measurements and, possibly, direct comparisons of metrics among programs.

*Organisms not identified.*—In general, aquatic vertebrates such as fish and amphibians (36.1% or 22/61), terrestrial invertebrates (50.8% or 31/61), adult stages of aquatic insects (19.6% or 12/61), and aquatic insect pupae (16.4% or 10/61) were not identified from benthic samples. Omission of the latter 2 types of organisms is unfortunate because they can be used to give specific names to the immature taxa collected. Some programs did not identify nematodes (18.0% or

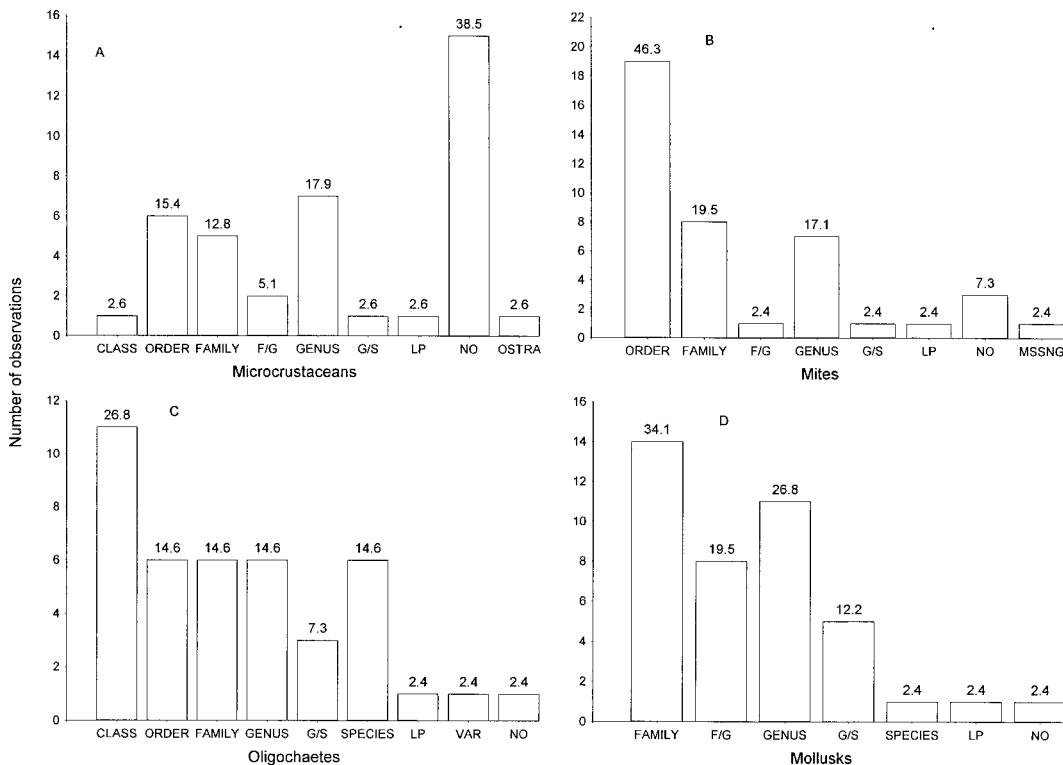


FIG. 6. Percentage (above each bar) of taxonomic levels used for identifying specific taxa ( $n = 41$ , except Microcrustaceans where  $n = 39$ ). A.—Microcrustaceans. B.—Mites. C.—Oligochaetes. D.—Mollusks. F/G = family to genus, G/S = genus to species, LP = lowest practicable level, NO = taxon not identified, MSSNG = taxonomic level not indicated, VAR = variable taxonomic level used, OSTRA = Ostracoda.

11/61) and microcrustaceans (16.4% or 10/61), perhaps because these taxa are inconsistently collected by the 500- to 600- $\mu$ m meshes used. However, early instars of most aquatic insects also are inconsistently collected using these mesh sizes; therefore, not identifying these non-insect taxa likely reflects uncertainty in their identification because of a lack of keys, a lack of program interest, and/or a lack of available expertise.

*Enumeration of early instars and damaged organisms.*—Early instars of aquatic insects were generally identified at higher taxonomic levels and counted as part of the sample by 88.9% (64/72) of the programs. Damaged organisms were identified and counted by 81.7% (58/71) of the programs if the specimens had sufficient diagnostic body parts remaining. Many (72.9% or 43/59) programs identified and counted an organism even if only the head was present.

#### *Questions related to confidence in data quality*

Most of the programs had developed Standard Operating Procedures for sample collection and processing (97.2% or 70/72), had quality control procedures (97.3% or 73/75), and had numeric criteria for at least some portion of their quality assurance plan (70% or 49/70). Sorting efficiency and macroinvertebrate identifications were checked by 88% (22/25) of the programs. Moreover, 71.4% (50/70) regularly consulted taxonomic specialists.

Most samples were collected by program staff (90.8% or 69/76) or program staff with the aid of contractors (98.7% or 75/76). Few programs had contractors solely responsible for sample collection. However, even though most samples were collected by program staff, only 64.3% (45/70) of the programs processed their own samples. Contractors were solely responsible for

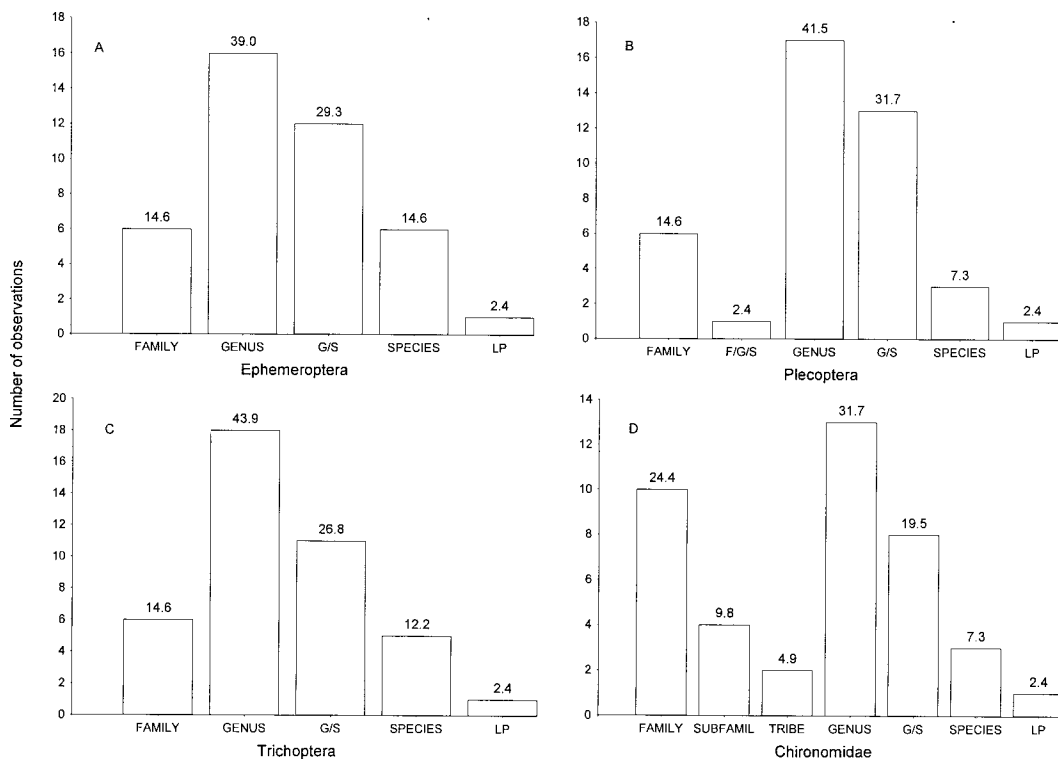


FIG. 7. Percentage (above each bar) of taxonomic levels used for identifying specific taxa ( $n = 41$ ). A.—Ephemeroptera. B.—Plecoptera. C.—Trichoptera. D.—Chironomidae. SUBFAMIL = subfamily, F/G = family to genus, F/G/S = family to species, G/S = genus to species, LP = lowest practicable level, NO = taxon not identified.

processing samples for 21.3% (16/75) of the programs, and program staff with the aid of contractors processed samples for 18.7% (14/75) of the programs.

Few samples were discarded initially or retained for <1 y (8.2% or 6/73); most programs kept specimens indefinitely (47.9% or 35/73). Most programs maintained reference or voucher collections. However, most (87.9% or 58/66) programs kept their collections in-house, and only 2% of the collections were kept by contractors. Moreover, <10% (6/66) of the programs kept their collections in university museums.

#### *Extent of monitoring programs*

One important result of conducting this survey was to document the remarkably large number of benthic macroinvertebrate samples that are collected each year by the respondents from these state programs. We asked respondents to estimate both the minimum and max-

imum number of samples collected per year. The average number of samples/program/year was 178. The minimum number of samples collected per year when summed across all the programs was 13,180 and the maximum was 15,236. Regardless, it is evident that far more samples need to be collected, both spatially and temporally, to adequately represent the distribution of benthic macroinvertebrates in potentially impaired streams in the US.

#### *Regional differences*

Numerous regional evaluations of the variables contained in the questionnaire could be made, but we highlight a few that we thought could most influence biomonitoring data and comparability among programs. We examined these patterns by grouping states east and west of the Mississippi River, or by US Environmental Protection Agency (USEPA) jurisdictional regions.

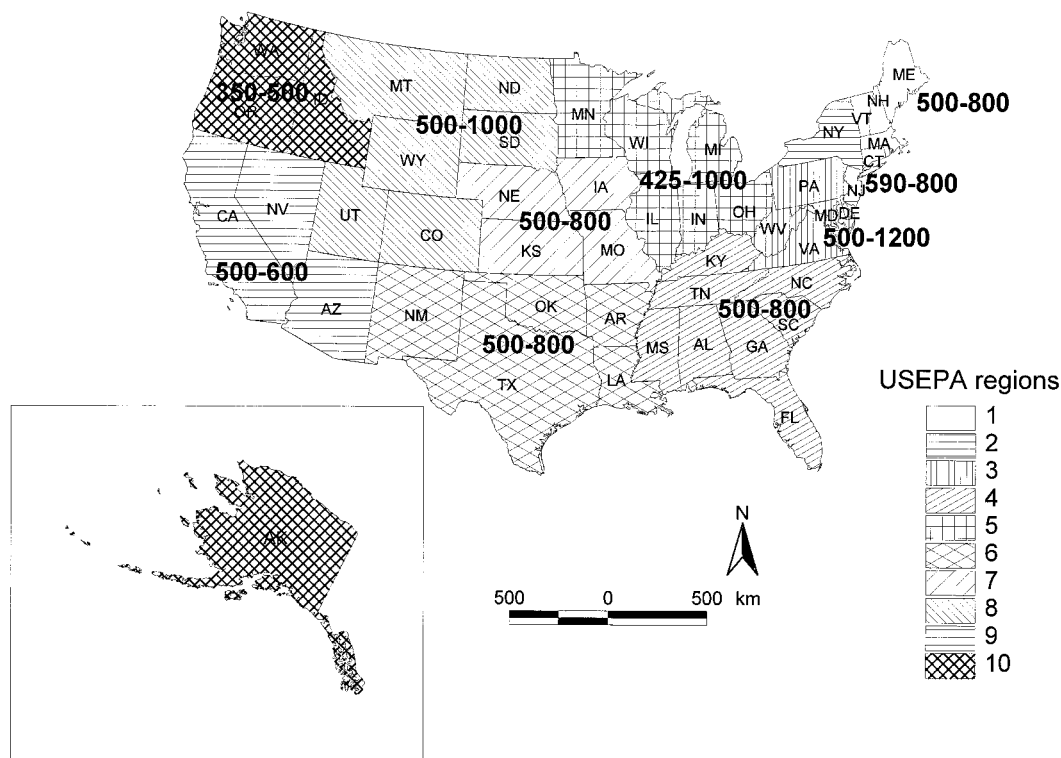


FIG. 8. Range of mesh sizes ( $\mu\text{m}$ ) used by  $\geq 1$  methods for collecting benthic macroinvertebrates within each US Environmental Protection Agency region.

Mesh size differences, even within a region, varied greatly (Fig. 8). From an east-west perspective, mesh sizes used for collecting were significantly larger in the east ( $\bar{x} = 628 \pm 143.3$  SD,  $n = 49$ ) than in the west ( $\bar{x} = 548 \pm 118.5$  SD,  $n = 32$ ).

Sampling devices varied greatly from region to region (Fig. 9); the D-frame net was the only device used within all 10 USEPA regions. The 1 m kicknet was used far more often by programs in the east than in the west, and fixed-quadrat samplers were used in the west but not in the east (Fig. 10).

Differences in the number of organisms sorted per sample ranged from 0 (all programs within a region sorted the same number of organisms) to 250 (Fig. 11). In general, most (72.7% or 16/22) programs in the east sorted 100 organisms, whereas in the west most (62% or 9/21) programs sorted  $\geq 300$  organisms (Fig. 12).

Differences in methods between the eastern and western states probably resulted because

bioassessment protocols were established earlier in the east than in the west, and some western states developed protocols independently of eastern programs.

### Conclusion

Although methodological research in freshwater benthic science is not often a popular research topic, the need for rigorous, well-designed additional studies of sampling and sorting procedures is evident from the results of this survey. A few recent studies have addressed some of these topics: subsampling (Barbour and Gerritsen 1996, Courtemanch 1996, Vinson and Hawkins 1996); field sampling and subsampling (Gowns et al. 1997, Larsen and Herlihy 1998, Rabeni et al. 1999); and taxonomic resolution (Hewlett 2000, Bailey et al. 2001, Lenat and Resh 2001). However, specific comparisons of currently used techniques to determine whether and/or how differences in methods affect our ability to detect spatial and temporal

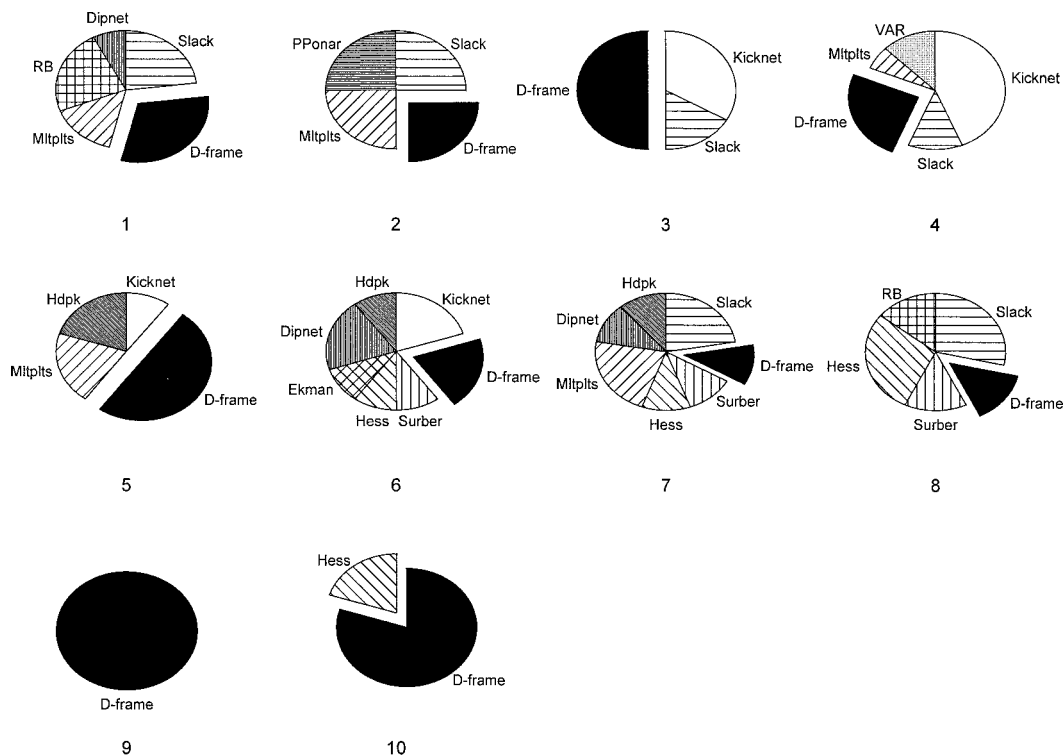


FIG. 9. Percentage of each device used within each US Environmental Protection Agency region (number below each pie diagram). Mltplts = multiplate artificial substrates, RB = rock basket, Hdpc = hand picking, Pponar = petit ponar grab, VAR = various methods used in combination.

changes in benthic communities is a research question that must be further addressed.

Many questions are particularly appropriate for undergraduate theses and honors projects, but other topics require considerable effort to determine the specific effects of sampling and sorting on our ability to detect environmental change. These types of studies should be rigorously designed, evaluated for the introduction of method-specific errors, funded by monitoring agencies and industry, and carried out by multiple, well-established, public and private monitoring teams and processing laboratories.

Some research questions emerge as being obvious and useful:

- 1) Do different mesh sizes (e.g., 500, 600, 1000  $\mu\text{m}$ ) affect our ability to detect spatial and/or temporal change in benthic communities? If early instars of aquatic insects are more susceptible to impairment, a mesh-size effect would certainly be true; however, early

instars also are probably more temporally variable.

- 2) What are the changes in functional mesh size of netted samplers as the collection of a sample proceeds from initial flow of water through the net to gradual clogging with debris?
- 3) How do different mesh sizes compare in terms of indicating impact with either multimeric or multivariate approaches? Are 1 to 2 mm mesh sizes ever appropriate for monitoring (e.g., by volunteer groups)?
- 4) Is there a difference in richness values obtained when a few large collections are compared to many smaller collections are composited? Is there an optimal composite sample size for general biomonitoring?
- 5) Is there an increased benefit to cost ratio in sampling multiple riffle-pool sequences compared to a single riffle-pool sequence to represent a site?



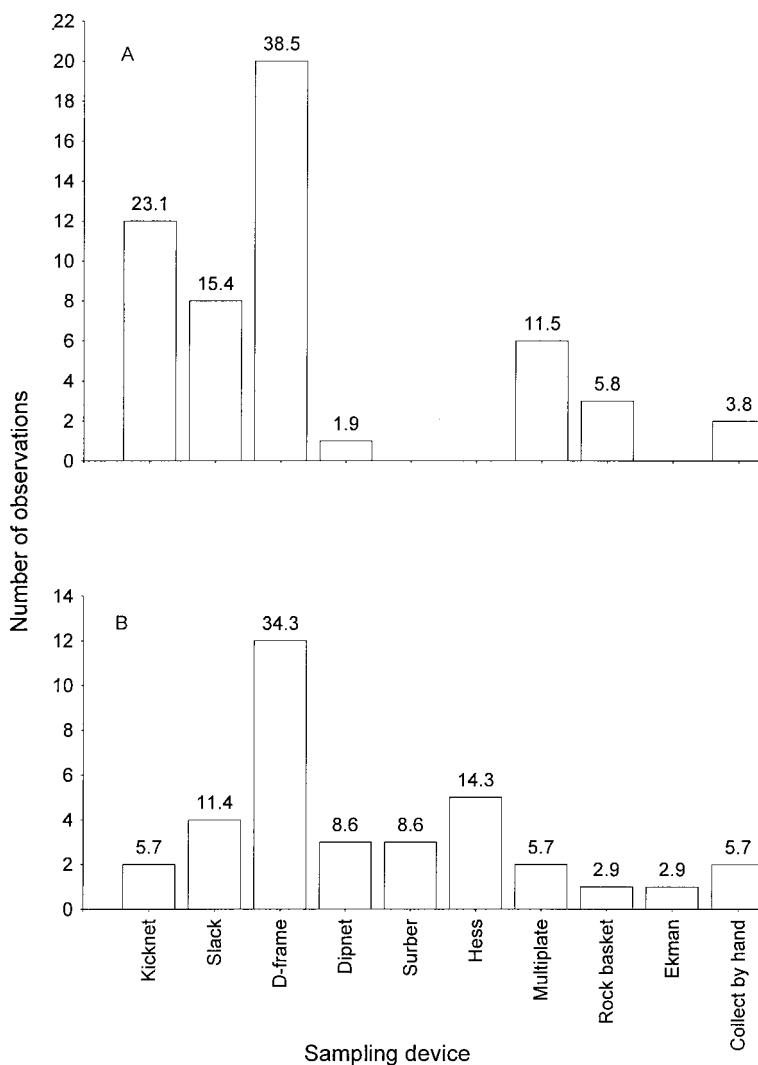


FIG. 10. Differences in percentage (above each bar) of methods used for sampling between the states east (A,  $n = 52$ ) and west (B,  $n = 35$ ) of the Mississippi River.

- 6) Are there significant differences in the information content of samples from single compared to multiple habitats?
- 7) Can we develop standard protocols for consistently (and from operator-to-operator) sampling habitats "in proportion to their occurrence?"
- 8) How do we standardize "expert opinion" in terms of sampler placement?
- 9) How many macroinvertebrates should be sorted from a sample (100, 200, 300, etc.) and is this value related to regional differ-

ences in species richness and/or level of impairment?

- 10) Is sorting large/rare organisms advantageous? If so, what is the best method to standardize sorting such organisms?
- 11) What are the effects of sorting at various magnifications on the values of commonly used metrics and how many (what %) of organisms can be *missed* without affecting metric values? What amount of time should be allotted to check for additional organisms?

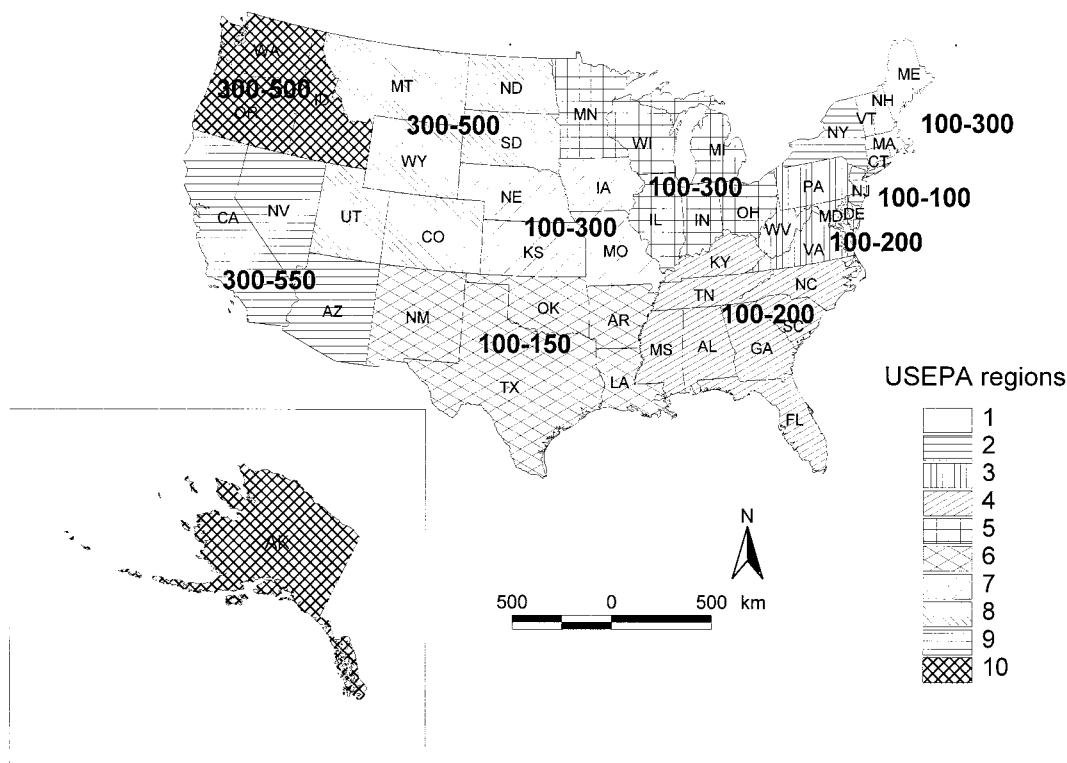


FIG. 11. Ranges of fixed numbers of organisms sorted from a sample in the laboratory within each US Environmental Protection Agency region.

- 12) What level of taxonomic resolution is necessary for general biomonitoring in different geographic/biotic regions? Should only taxa identified to the same level throughout the year and/or at all sites be used in analyses?

There are wonderful opportunities for gad-geteers, hobby taxonomists, and armchair statisticians to pursue. Are there portable magnifiers that can be used in the field to overcome problems of lighting and differences in eyesight among sorters? Are there more effective methods for field elutriation? What is the optimal subsampling device and the best method to use for subsampling benthic macroinvertebrates in complex substrate matrices (e.g., sediment, filamentous algae, etc.)?

Identification keys that are produced by local agencies need to be made more easily available and be more widely distributed. The internet and websites offer marvelous opportunities for

the interactive sharing of taxonomic expertise without the time and hassle of journal publication. In addition, we highly recommend the establishment of a national website where identification of individual taxa can be discussed. Such discussion would significantly improve the precision of taxonomic identifications, development of pollution tolerance values, etc.

In conclusion, the results of our survey clearly indicate that the collection and processing of benthic samples provides a major source of employment for benthic biologists in the US. From 13,000 to 15,000 samples per year are processed by the 48 states responding to our survey. The large number of intra- and inter-regional differences in methods used both in the field and laboratory indicate that calibration on a national scale would be extremely challenging, although a few programs in neighboring states have standardized and/or calibrated certain aspects of their methods. The sharing of sites among agen-

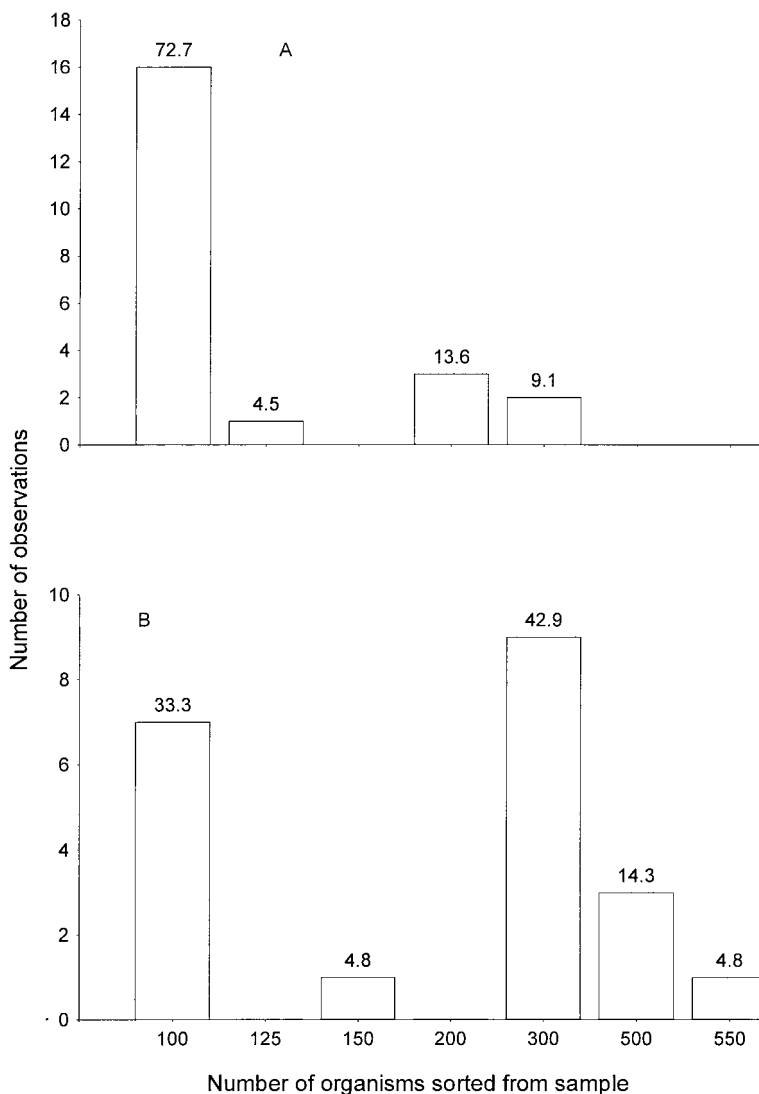


FIG. 12. Differences in percentage (above each bar) of fixed numbers of organisms removed during sample sorting in the laboratory between the states east (A,  $n = 22$ ) and west (B,  $n = 21$ ) of the Mississippi River.

cies, and maintaining regional/national voucher collections and identification keys can make bioassessments more accurate and more efficient.

Well-designed tests of the most commonly used sampling and processing methods to determine whether the effects introduced by differences in methods significantly influence the interpretation of benthic macroinvertebrate data are needed. As a starting point, answers to the 12 questions listed above would allow more in-

formed decisions to be made by biomonitoring researchers on the most appropriate method to use for monitoring using macroinvertebrates, and increase national comparability among programs.

#### Acknowledgements

We thank all the state biologists and managers for the time and effort they put into completing our questionnaire. We also thank Mi-

chael Barbour, Steve Fend, Mike Miller, Steve Moulton, Terry Short, and James Stribling for their helpful comments and suggestions on an earlier version of this manuscript.

### Literature Cited

- BAILEY, R. C., R. H. NORRIS, AND T. B. REYNOLDS. 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, 2nd edition. EPA 841-B-99-002. Office of Water, US Environmental Protection Agency, Washington, DC.
- CAO, Y., D. D. WILLIAMS, AND N. E. WILLIAMS. 1998. How important are rare species in aquatic community ecology and bioassessment? *Limnology and Oceanography* 47:1403–1409.
- CATON, L. W. 1991. Improving subsampling methods for the EPA "Rapid Bioassessment" benthic protocols. *Bulletin of the North American Benthological Society* 8:317–319.
- CHUTTER, F. M., AND R. G. NOBLE. 1966. The reliability of a method of sampling stream invertebrates. *Archiv für Hydrobiologie* 62:95–103.
- COURTEMANCH, D. L. 1996. Commentary on the subsampling procedures used for rapid bioassessments. *Journal of the North American Benthological Society* 15:381–385.
- CUMMINS, K. W. 1962. An evaluation of some techniques for the collection and analysis of benthic samples with special emphasis on lotic waters. *American Midland Naturalist* 67:477–504.
- DAVIS, W. S., B. D. SNYDER, J. B. STRIBLING, AND C. STOUGHTON. 1996. Summary of state biological assessment programs for streams and wadeable rivers. EPA 230-R-96-007. Office of Policy, Planning, and Evaluation, US Environmental Protection Agency, Washington, DC.
- DOWNING, J. A. 1984. Sampling the benthos of standing waters. Pages 87–130 in J. A. Downing and F. H. Rigler (editors). *A manual on methods for the assessment of secondary productivity in fresh waters*. 2nd edition. IBP Handbook 17. Blackwell Scientific Publications, Oxford, UK.
- FRISSELL, C. A., W. J. LISS, C. E. WARREN, AND M. D. HURLEY. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management* 10:199–414.
- 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.
- 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.
- HORNIG, C. E., AND J. E. POLLARD. 1978. Macroinvertebrate sampling techniques for streams in semi-arid regions. EPA-600/4-78-040. Office of Research and Development, Environmental Monitoring and Support Laboratory, US Environmental Protection Agency, Las Vegas, Nevada.
- HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* 54:187–211.
- HURYIN, A. D., AND J. B. WALLACE. 2000. Life history and production of stream insects. *Annual Review of Entomology* 45:83–110.
- KLEMM, D. J. (EDITOR). 1985. *A guide to the freshwater Annelida (Polychaeta, naidid and tubificid Oligochaeta, and Hirudinea) of North America*. Kendall/Hunt Publishing Company, Dubuque, Iowa.
- 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.
- 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.
- LEONARD, J. W. 1939. Comments on the adequacy of accepted stream bottom sampling techniques. *Transactions of the North American Wildlife and Natural Resources Conference* 4:288–295.
- MERRITT, R. W., AND K. W. CUMMINS (EDITORS). 1996. *An introduction to the aquatic insects of North America*. Kendall/Hunt Publishing Company, Dubuque, Iowa.
- MOULTON, S. R., J. L. CARTER, S. A. GROTHEER, T. F. CUFFNEY, AND T. M. SHORT. 2000. Methods of analysis by the US Geological Survey National Water Quality Laboratory—processing, taxonomy, and quality control of benthic macroinvertebrate samples. US Geological Survey Open-File Report 00-212. US Geological Survey, Denver, Colorado.
- NEEDHAM, P. R., AND R. L. USINGER. 1956. Variability in the macrofauna of a single riffle in Prosser Creek, California, as indicated by the Surber sampler. *Hilgardia* 24:383–409.
- NORRIS, R. H., AND A. GEORGES. 1993. Analysis and interpretation of benthic macroinvertebrate sur-

- veys. Pages 234–286 in D. M. Rosenberg and V. H. Resh (editors). *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York.
- NORRIS, R. H., E. P. MCELRAVY, AND V. H. RESH. 1995. The sampling problem. Pages 282–306 in P. Calow and G. E. Petts (editors). *The rivers handbook*. Blackwell Scientific Publications, Oxford, UK.
- 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/444/4-89/001. Office of Water, US Environmental Protection Agency, Washington, DC.
- RABENI, C. F., N. WANG, AND R. J. SARVER. 1999. Evaluating adequacy of the representative stream reach used in invertebrate monitoring programs. *Journal of the North American Benthological Society* 18:284–291.
- RESH, V. H. 1979. Sampling variability and life history features: basic considerations in the design of aquatic insect studies. *Journal of the Fisheries Research Board of Canada* 36:290–311.
- RESH, V. H., AND J. K. JACKSON. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. Pages 195–233 in D. M. Rosenberg and V. H. Resh (editors). *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York.
- RESH, V. H., AND E. P. MCELRAVY. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. Pages 159–194 in D. M. Rosenberg and V. H. Resh (editors). *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York.
- RESH, V. H., D. M. ROSENBERG, AND J. W. FEMINELLA. 1985. The processing of benthic samples: responses to the 1983 NABS questionnaire. *Bulletin of the North American Benthological Society* 2:5–11.
- SLACK, K. V., L. J. TILLEY, AND S. S. KENNELLY. 1991. Mesh-size effects on drift sample composition as determined with a triple net sampler. *Hydrobiologia* 209:215–226.
- 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.
- VOSHELL, J. R., R. J. LAYTON, AND S. W. HINER. 1989. Field techniques for determining the effects of toxic substances on benthic macroinvertebrates in rocky-bottom streams. Pages 134–155 in U. M. Cowgill and L. R. Williams (editors). *Aquatic toxicology and hazard assessment*. Volume 12. ASTM Special Technical Publication 1027. American Society for Testing and Materials, Philadelphia, Pennsylvania.
- WINTERBOURN, M. J. 1985. Sampling stream invertebrates. Pages 241–258 in R. D. Pridmore and A. B. Cooper (editors). *Biological monitoring in freshwaters: proceedings of a seminar*, Hamilton, November 21–23, 1984, Part 2. Water and Soil Miscellaneous Publication No. 83. National Water and Soil Conservation Authority, Wellington, New Zealand.
- WRONA, F. J., J. M. CULP, AND R. W. DAVIES. 1982. Macroinvertebrate subsampling: a simplified apparatus and approach. *Canadian Journal of Fisheries and Aquatic Sciences* 39:1051–1054.

## ABOUT THE AUTHORS

*Jim Carter is an aquatic ecologist with the National Research Program, Water Resources Division of the US Geological Survey. He studies the influence of physical and chemical factors on community structure of benthic invertebrates in streams. Vince Resh is a Professor of Entomology at the University of California, Berkeley, and has taught there for 25 y. He has done extensive research on stream and river bioassessment using macroinvertebrates.*

Appendix. Questionnaire used for obtaining data on methods used by state agencies for collecting and processing benthic macroinvertebrate samples from streams.

## INTRODUCTION

We are interested in determining the current methods used to collect and process benthic invertebrate samples for bioassessments and other lotic studies. To acquire these data we are polling researchers actively involved in bioassessments and/or the development of bioassessment techniques. Because we are contacting a limited number of researchers for this survey, we greatly appreciate you taking the time to fill out this questionnaire.

Benthic sample collecting and processing varies greatly among programs and researchers; therefore, some

questions may not exactly represent your methods. However, please respond as best you can to as many questions as possible.

Thank you,

Jim Carter and Vince Resh

Jim Carter  
U.S. Geological Survey  
345 Middlefield Rd. MS 465  
Menlo Park, CA 94025  
USA

Tel: 650-329-4439

jlcarter@usgs.gov

Vince Resh  
UC Berkeley  
ESPM/IB  
201 Wellman Hall  
Berkeley, CA 94720  
USA

Tel: 510-642-3763

vresh@nature.berkeley.edu

Please fill out the following information.

Name: \_\_\_\_\_

Organization: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Email: \_\_\_\_\_

Tele: \_\_\_\_\_

General comments about your program:

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Please fill out one questionnaire for each sampling/processing protocol used by your program. A protocol for our purposes is defined as a unique combination of samplers, mesh size, habitats sampled, and field/laboratory processing that are consistently applied for collecting and processing samples within a study.

#### SECTION A: SAMPLE COLLECTION

\*\*\*\*\*

This section is designed to collect data on field sampling. We make a distinction between two commonly used approaches for collecting a benthic sample. We distinguish between a *discrete sample* and a *composite sample*. A *discrete sample* is defined as a sample collected from a contiguous area (e.g., a single Surber sample or a single kick net sample). A *composite sample* is defined as a sample collected from a non-contiguous area (e.g., collecting several Surber samples (=units) within a riffle or reach and combining all of them in the field).

\*\*\*\*\*

A-I. What type of sampling device is used (e.g., Kicknet (net between 2 poles), D-frame, Surber, Box, Hess, Grabs (Ekman, Peterson, Ponar, other), Artificial Substrates (Multiplates, Rock-filled basket, other), Corer, etc.)? Please describe. \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

A-II. What mesh size is used during sample collection or, for artificial substrates, sampler retrieval?

\_\_\_\_\_ (Please provide units)

A-III. Is the sample a composite sample as defined above?      **Yes**      **No**

**If yes,**

- A. How many "units" are collected per composite sample? \_\_\_\_\_  
 B. What is the **total** area of the composite sample?

\_\_\_\_\_ (**Please provide units**)

A-IV. What is the area of a single unit of the composite **or** single discrete sample if the sample is not a composite? If the sample is limited by time instead of area (e.g., 3 minutes), please describe in "Other". (**Please circle only one**)

- A. dipnet (no area)  
 B.  $<0.09\text{-m}^2$  (1-ft<sup>2</sup>)  
 C.  $\sim 0.1\text{-m}^2$   
 D.  $\sim 0.18\text{-m}^2$  (2-ft<sup>2</sup>)  
 E.  $\sim 0.25\text{-m}^2$   
 F.  $\sim 0.5\text{-m}^2$   
 G.  $\sim 1.0\text{-m}^2$   
 H.  $>1.0\text{-m}^2$   
 I. Other \_\_\_\_\_

(**Please provide units**)

A-V. Is the sample collected from a single habitat-type (e.g., riffle, pool, run, etc.)?      **Yes**      **No**

\*\*\*\*\*

The next question addresses the length of stream over which only a composite sample is collected. If your sample is not a composite, please go to question A-VII.

\*\*\*\*\*

A-VI. Please indicate which of the following lengths of stream best describe the length over which the composite is collected. **Indicate all that apply to the collection of the composite.**

- A. Is the composite collected within a single riffle-pool sequence?      **Yes**      **No**  
 B. Is the composite collected over a fixed number of riffle-pool sequences?      **Yes**      **No**

**If yes,**

1. How many? \_\_\_\_\_  
 C. Is the composite collected over stream lengths of multiple channel widths (e.g., 20× the channel width)?      **Yes**      **No**

**If yes,**

1. How many channel widths? \_\_\_\_\_  
 D. Is the composite collected over a fixed length of stream?      **Yes**      **No**

**If yes,**

1. What is the length? \_\_\_\_\_ (**Please provide units**)  
 E. Other, \_\_\_\_\_

\*\*\*\*\*

Please answer the following three questions (A-VII through A-IX) regardless of the type of sample collected.

\*\*\*\*\*

A-VII. Please circle the method that best describes the placement of the sampling device in the habitat or area of stream sampled.

- A. Using random numbers  
 B. Systematically (e.g., along the thalweg; at 1/3 intervals across the stream; etc.)  
 C. Expert opinion (e.g., the area of a riffle that appears most complex and may yield the greatest number of taxa)  
 D. Haphazardly (i.e., no system—not even random)  
 E. Other \_\_\_\_\_

Please describe: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

A-VIII. What habitat(s) are sampled? (**Please circle only one**)

- A. Riffle(s) only  
 B. Pool(s) only  
 C. Run/glide(s) only



- D. Fast water habitats (i.e., riffles and runs)
- E. A composite of fast and slow water habitats
- F. A composite of all habitats present in proportion to their occurrence
- G. A composite of all habitats present with the same effort (equal time, equal number of sweeps, etc.) per habitat
- H. Other \_\_\_\_\_

A-IX. Are replicates collected (replicates are defined as additional discrete samples or additional composite samples collected at a "site" and processed separately—a replicate is **not** defined as one part of a composite sample)? **Yes** **No**

If yes,

- A. How many replicates are collected per site? \_\_\_\_\_
- B. What is the percentage of sites from which replicates are collected? \_\_\_\_\_ %

## SECTION B: FIELD PROCESSING

\*\*\*\*\*  
Please answer all relevant questions regardless of the type of sample collected.  
\*\*\*\*\*

B-I. Is the sample **only** sorted by eye in the field (i.e., without magnification)? **Yes** **No**

If yes,

- A. Are a fixed count of organisms chosen? **Yes** **No**

If yes,

- 1. How many? \_\_\_\_\_

- B. Are only certain taxa chosen (e.g., EPT)? **Yes** **No**

If yes,

- 1. Which taxa are chosen? \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

- C. Are organisms chosen to maximize the number of different taxa collected? **Yes** **No**

- D. Are actual counts made of the number of individuals per taxon? **Yes** **No**

- E. Are ordinal values assigned to the taxa sorted in the field (e.g., rare, common, abundant)? **Yes** **No**

\*\*\*\*\*

If you answered **Yes** to B-I, please go to question B-V

\*\*\*\*\*

B-II. Please circle **all** of the statements that best fit your field procedure.

- A. Absolutely no field processing is done—the entire sample is taken to the laboratory.
- B. The sample is elutriated in the field.
- C. Large organic debris and large stones are removed and discarded in the field.
- D. Other \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

B-III. Even though the sample was most likely collected with a net, is the sample further sieved in the field? **Yes** **No**

If yes,

- A. What mesh size of sieve is used? \_\_\_\_\_ (Please provide units)

- B. Are only individuals retained on the sieve used for analyses? **Yes** **No**

B-IV. Are samples subsampled in the field? **Yes** **No**

If yes, which of the following methods is used to subsample in the field?

- A. The subsample is a percentage of the total sample (e.g., 50%, 25%, etc.). **Yes** **No**

If yes,

- 1. Do you record the percentage taken from each sample? **Yes** **No**

If yes,

- a. Is the same percentage taken from each sample? **Yes** **No**

If yes,

- 1. What percentage is taken? \_\_\_\_\_ %

2. How is the percentage taken? \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

B. The subsample is a fixed count of organisms.      **Yes**      **No**

**If yes,**

1. Are fixed counts of organisms randomly acquired?      **Yes**      **No**

2. How many individuals are collected? \_\_\_\_\_

C. Other, please describe: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\* \* \* \* \*

B-V. What preservative is used in the field (**concentration**)?

A. No preservative used

B. Formalin ( \_\_\_\_\_ %)

C. Ethyl Alcohol ( \_\_\_\_\_ %)

D. Isopropyl Alcohol ( \_\_\_\_\_ %)

E. Kahle's

F. Other \_\_\_\_\_ ( \_\_\_\_\_ %)

B-VI. How do you dispose of the preservative used in the field when you are processing the samples in the laboratory? \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

#### SECTION C: LABORATORY PROCESSING

C-I. Are all organisms sorted from the sample?      **Yes**      **No**

**If yes,** please go to question C-VI.

**If no,** continue . . .

\* \* \* \* \*

Please answer questions C-II through C-V, if any portion of the sample is subsampled.

\* \* \* \* \*

C-II. Do you size fractionate your sample with sieves before sorting (or subsampling)?      **Yes**      **No**

**If yes,**

A. What mesh size(s) is used? \_\_\_\_\_  $\mu\text{m}$

B. Are analyses performed on only those organisms retained on the sieve?      **Yes**      **No**

**If no,**

1. Are the organisms that pass the sieve further subsampled?      **Yes**      **No**

**If yes,**

a. What device (method) is used? \_\_\_\_\_

\_\_\_\_\_

b. How many organisms are typically identified from this portion of the subsample? \_\_\_\_\_

c. Are the organisms that passed the sieve used in analysis?      **Yes**      **No**

C-III. Do you subsample the sample, regardless of whether or not it is size fractionated?      **Yes**      **No**

**If yes,** please circle and fill out A, B, or C.

A. The subsample is acquired by sorting a fixed count of organisms.

1. Are subsamples randomly chosen?      **Yes**      **No**

2. What is your target number of organisms (e.g., 50, 100, 300, etc.)? \_\_\_\_\_

\_\_\_\_\_

B. The subsample is a percentage of the sample, regardless of the number of organisms the percentage contains.

1. Is the same percentage of the sample always sorted (e.g., 50%, 25%, etc.)?      **Yes**      **No**

**If yes,**

a. What percentage is sorted? \_\_\_\_\_ %

C. Other \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

C-IV. What device(s) or method(s) is used for subsampling the sample in the laboratory (e.g., gridded tray, Folsom plankton splitter, Imhoff cone, etc.)? \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

C-V. In addition to acquiring organisms for your analyses by subsampling, do you also sort out large-rare organisms from your sample?     **Yes**     **No**

**If yes,**  
A. Do you sort out large-rare organisms from your sample before or after subsampling? **Please circle one . . . Before    After**  
B. Describe the method(s) used for acquiring these invertebrates from the sample. \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

\*\*\*\*\*  
End of subsampling section.  
\*\*\*\*\*

C-VI. Do you have an established maximum time limit for sorting a sample?     **Yes**     **No**  
**If yes,**  
A. What is your time limit? \_\_\_\_\_

C-VII. What magnification is used when sorting a sample or subsample? (**Please circle only one**)  
A. None (sort by eye)  
B. ~2×  
C. ~3×  
D. ~4×  
E. ~5×  
F. ~6×  
G. ~7×  
H. ~8×  
I. ~10×  
J. >10× **Please indicate:** \_\_\_\_\_

C-VIII. Do you subsample (or further subsample) certain extremely abundant taxa (e.g., Chironomidae) prior to identifying them to lower taxonomic levels?     **Yes**     **No**  
**If yes,**  
A. List the taxa that are identified using the above process.  
\_\_\_\_\_  
\_\_\_\_\_  
B. Describe the method used for subsampling these taxa.  
\_\_\_\_\_  
\_\_\_\_\_

C-IX. If you were asked to what taxonomic level you typically identified invertebrates—what would you respond (e.g., Order, Family, Genus, Species, etc.)? \_\_\_\_\_

C-X. Are certain commonly collected organisms not identified (i.e., not even recorded on the data sheet [e.g., fish, terrestrials, adult aquatics, pupae, mites, worms, Nematoda, etc.]?)?     **Yes**     **No**  
**If yes,**  
A. Please indicate which taxa and/or life-stages. \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

C-XI. To what level are the following taxa typically identified (**please indicate which taxa are normally slide mounted for identification**)?  
A. Oligochaeta \_\_\_\_\_  
B. Microcrustacea \_\_\_\_\_  
C. Ephemeroptera \_\_\_\_\_

- D. Odonata \_\_\_\_\_
- E. Plecoptera \_\_\_\_\_
- F. Hemiptera \_\_\_\_\_
- G. Trichoptera \_\_\_\_\_
- H. Coleoptera \_\_\_\_\_
- I. Tipulidae \_\_\_\_\_
- J. Simuliidae \_\_\_\_\_
- K. Chironomidae \_\_\_\_\_
- L. Higher Diptera \_\_\_\_\_
- M. Mites \_\_\_\_\_
- N. Mollusca \_\_\_\_\_
- O. Other \_\_\_\_\_

C-XII. Additional questions regarding enumeration.

- A. In general, are early instars and immatures categorized at higher taxonomic levels and counted?

**Yes No**

- B. If a fragment of an organism can be identified—is it counted? **Yes No**

**If yes,**

1. Are only heads counted? **Yes No**

#### SECTION D: GENERAL QUESTIONS

- D-I. Do you use specific procedures for maintaining confidence in your data quality? **Yes No**

**If yes,**

- A. Do you have a documented Standard Operating Procedure for sample collection and processing represented by this questionnaire? **Yes No**
- B. Is the laboratory processing subjected to Quality Control procedures and numeric criteria that are documented in a Quality Assurance plan? **Yes No**
- C. **If you answered yes to D-I, but you answered No to D-I.A. and D-I.B.,** do you regularly check the quality of sample sorting and invertebrate identification? **Yes No**

**If yes,** briefly describe your method: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

- D-II. Are taxonomic specialists regularly consulted? **Yes No**

- D-III. Are samples **collected** by you and your staff or by contractors? **Me Contractors**

- D-IV. Are samples **processed** by you and your staff or by contractors? **Me Contractors**

D-V. Sample retention and curation.

- A. How long are samples retained? \_\_\_\_\_

- B. Is a reference/voucher collection maintained? **Yes No**

**If yes,**

1. Where is the above maintained? \_\_\_\_\_  
 \_\_\_\_\_

2. How long is the above maintained? \_\_\_\_\_

D-VI. Number of samples.

- A. On average, how many samples do you (or your program) process per year using the protocol this questionnaire represents? \_\_\_\_\_

- B. How many years have you (or your program) been processing this many samples? \_\_\_\_\_

- C. How many more years do you (or your program) expect to process this number of samples? \_\_\_\_\_

#### CONCLUSION

Thank you very much for filling out our questionnaire. The data derived from this survey will allow all of us to better define, classify, and understand current sampling and processing techniques used in bioassessments and other lotic studies.