Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality?

Bernard W. Sweeney[1], Juliann M. Battle[2], AND John K. Jackson[3]

Stroud Water Research Center, 970 Spencer Road, Avondale, Pennsylvania 19311 USA

Tanya Dapkey[4]

University of Pennsylvania, Department of Biology, Philadelphia, Pennsylvania 19104 USA

Abstract. Four approaches to or levels of identifying macroinvertebrates (amateur/family, expert entomologist/genus, expert entomologist/genus–species, and DNA barcoding/species) were used to assess community structure and water quality in White Clay Creek, Pennsylvania, USA. Macroinvertebrates were collected in March 2008 from 2 riffle sites 3.9 km apart on the same stream. The downstream site was known to be degraded by land and water use. About 98% of the 1617 specimens used for analysis, including small, immature, and damaged specimens, were successfully barcoded (sequenced) for the mitochondrial cytochrome c oxidase subunit I gene. A criterion of 2 to 4% genetic divergence provided good separation of presumptive species. Barcodes increased the taxonomic inventory across the 2 sites by 475% (124 taxa) relative to the amateur level, and 125% (83 taxa) and 70% (62 taxa) relative to the expert genus and species levels, respectively. Barcoding revealed species not currently described in larval taxonomic keys, including multiple (2–11) coexisting congeneric species. That 150 species were revealed by barcoding samples collected on the same date and in the same habitat was unprecedented, as was the fact that 60 cm^2 of stream bottom supported an average of 248 to 347 individuals representing 55 to 68 species. Most barcode species were rare, with 42% represented by ≤2 individuals. Across all species, 43 of 89 barcode species were unique to upstream site 11 and 60 of 107 were unique to downstream site 12. In terms of water-quality assessment, most of the 17 metrics studied changed significantly (α = 0.05) when taxonomy changed from family to genus–species (79 and 93% for sites 11 and 12, respectively), and many also changed when taxonomy changed from genus to species (59 and 65% for sites 11 and 12, respectively). The proportion of metrics able to detect a difference (α = 0.05) between sites 11 and 12 increased with improved taxonomic resolution (36, 47, 65, and 76% for family, genus, genus–species, and barcode, respectively). The results revealed a *pollution-tolerance gap* because barcoding pushed larval taxonomy beyond the available pollution-tolerance data. Regardless, the combined morphological and molecular approach provides a finer resolution for evaluating environmental change associated with both natural and anthropogenic processes. The ability to distinguish larvae at the species level through barcoding finally puts biodiversity assessments for aquatic communities in terms comparable to those used for terrestrial ecosystems where estimates of biodiversity for plants and animals are never quantified at the level of genus or family. We conclude that DNA barcodes of stream macroinvertebrates will improve descriptions of community structure and water quality for both ecological and bioassessment purposes.

Key words: DNA barcoding, cytochrome c oxidase, COI gene, macroinvertebrates, freshwater biosurveillance, water-quality monitoring, community structure.

One of today's largest scientific challenges is to understand how natural and human-induced changes in watersheds affect the structure and function of

freshwater ecosystems and their ability to deliver services (sensu Daily and Ellison 2001). Formulating and testing hypotheses about population and community ecology of freshwater ecosystems, as well as the evolutionary processes associated with their biota, continue to be the primary means of addressing that challenge (Giller and O'Donovan 2002). As a diverse, productive component of most freshwater food webs

[1] E-mail addresses: sweeney@stroudcenter.org

[2] jbattle@stroudcenter.org

[3] jkjackson@stroudcenter.org

[4] tdapkey@sas.upenn.edu

(Allan and Castillo 2007) and a mainstay of water-quality monitoring programs (Barbour et al. 1999), fly fishermen (Knopp and Cormier 1997), and wildlife enthusiasts (Waters 2000) around the world, macro-invertebrates have assumed a central role in tests of such hypotheses, assessments of stream and river pollution (Lenat 1993, Resh et al. 1995), and development of ambient water-quality policy (Kenney et al. 2009).

A longstanding constraint on use of aquatic macroinvertebrates for environmental assessments has been the difficulty of identifying them to species, especially in the immature stages. This task is challenging for even the best taxonomists because species identification keys for the larval stages of many aquatic macroinvertebrates are incomplete, unreliable, and nonexistent in some cases (Gresens et al. 2007). In addition, immature and small macroinvertebrate specimens often cannot be identi-fied to species because of their life stage, size, or condition (i.e., damaged), which not only prevents full access to the ecological and evolutionary infor-mation they hold (e.g., physiological mechanisms of pollution tolerance; Buchwalter and Luoma 2005, Buchwalter et al. 2008), but also leads to errors and imprecision in assessments of habitat and water quality (Lenat and Resh 2001, Stribling et al. 2008). To complicate matters, cryptic species continue to be problematic for aquatic taxonomists (e.g., Funk et al. 1988, Sweeney and Funk 1991, Sharley et al. 2004, Stahls and Savolainen 2008, Alexander et al. 2009, Krosch et al. 2009, Zhou et al. 2010).

All of this leads to widespread use of generic and higher levels of taxonomic resolution for ecological assessments, despite the knowledge that congeneric and confamilial species can display a broad range of sensitivities to various environmental stresses. One consequence has been that authors of studies of stream pollution comparing species-, genus-, and family-level analyses often report similar findings with regard to degree of impact (see Lenat and Resh 2001) because of the need to assign to species pollution-tolerance information based on data from the generic or family level. We presently lack the taxonomic knowledge and tools to assess the com-plete aquatic biodiversity from local to continental scales at an ecologically meaningful taxonomic reso-lution (e.g., species). This inability, in turn, compro-mises our understanding of both the natural processes that influence local and regional patterns of stream biodiversity, and the consequences of past, present, and future human management of those systems. Thus, the current state of the art is that species identification is not implemented, nor necessarily required, for every bioassessment effort (Resh and Jackson 1993, Bowman and Bailey 1997) and, at this time, the ability of species-level data to aid biological assessments is not completely understood.

We propose that molecular genetic analysis can help resolve some of these issues. This approach has been used in the past to distinguish morphologically cryptic species of macroinvertebrates in ecological and evolutionary studies, but these methods were expensive and time consuming (e.g., Sweeney et al. 1987, Funk et al. 1988, Funk and Sweeney 1990, Sweeney and Funk 1991, Jackson and Resh 1992, 1998). Recent advances in direct sequencing of deoxyribonucleic acid (DNA) make molecular meth-ods more readily available to help resolve taxonomic challenges presented by fauna in general (Hebert et al. 2003) and freshwater macroinvertebrates in particular (Sharley et al. 2004, Pfenninger et al. 2007, Zhou et al. 2007, 2009, 2010, Sinclair and Gresens 2008, Stahls and Savolainen 2008, Krosch et al. 2009). Our purpose here is neither to criticize past taxonomic efforts nor to replace morphological taxonomy with DNA. Rather, it is to evaluate the efficacy of DNA barcoding for assessing aquatic macroinvertebrate biodiversity and water quality in White Clay Creek and, more generally, to show that DNA barcoding can improve taxonomy and eventually lead to a greater under-standing of community structure, pollution tolerance at the species level, and, ultimately, more effective bioassessment of stream and river ecosystems.

Methods

Study sites, collection, and identification

Our study was conducted on the East Branch of White Clay Creek near Avondale, Chester County, Pennsylvania, USA. Land use in the White Clay Creek watershed is mixed agriculture and intact forest. Macroinvertebrates were collected from 2 locations (sites 11 and 12). Site 11 is at the Stroud Water Research Center (lat 39°51.579′N, long −75°47.059′W) and is only ~3.9 km upstream of site 12 (lat 39°50.146′N, long −75°46.915′W). Site 11 has an intact riparian forest along the stream with light agriculture but minimal human development upstream. Several changes in land use occur between sites 11 and 12. These land uses include a golf course, a mushroom-growing operation, an orchard, and fairly intense row-crop agriculture and livestock farming. The sites were chosen because of their proximity on the same stream and because long-term water-quality monitor-ing data (spanning 1994–2008) demonstrates that site 12 is degraded relative to site 11 (JKJ, unpublished data). For example, Macroinvertebrate Aggregated
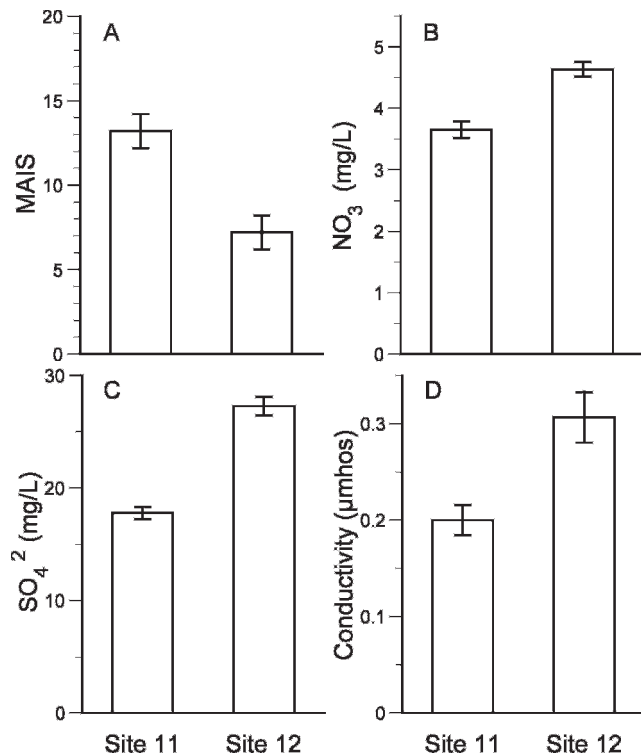
Fig. 1. Mean ($\pm$1 SE) values of the Macroinvertebrate Aggregated Index for Streams (MAIS) (A), NO_3^- (B), SO_4^{2-} (C), and conductivity (D) measured annually between 1994 and 2008 at sites 11 and 12 on White Clay Creek, Chester County, Pennsylvania (JKJ, unpublished data). All variables differed significantly among sites (analysis of variance, $p \leq 0.001$). The number of years sampled varied for MAIS ($n = 10$), NO_3^- ($n = 8$), SO_4^{2-} ($n = 4$), and conductivity ($n = 6$).

Index for Streams (MAIS; Smith and Voshell 1997) scores at site 12 (mean = 7.2; water-quality rating: fair) were consistently lower than scores at site 11 (mean = 13.2; rating: good; Fig. 1A), and levels of NO_3^- (Fig. 1B), SO_4^{2-} (Fig. 1C), and conductivity (Fig. 1D) were consistently elevated at site 12 vs site 11.

On 10 March 2008, a Surber sampler (250-μm mesh, 0.092 m^2) was used to sample macroinvertebrates from riffle habitat in >100-m stream reach at each site. Three composite samples were collected at each site and preserved in 95% nondenatured ethanol (ETOH). A composite sample consisted of 4 Surber samples that were combined in a large bucket. One quarter of the material was randomly removed with a quadrate splitting tool. Thus, 12 Surber samples covering ~1.1 m^2 of stream bottom were reduced to 3 composite samples. Sample compositing was used to increase the accuracy of a site description by increasing the area sampled without increasing the number of individuals or samples processed. Samples

were immediately taken back to the laboratory where ETOH was replaced, and samples were allowed to soak for $\geq$1 h in the fresh ETOH before processing. The composite sample was then further subsampled (into ½, ¼, ⅛, or 1/16 fractions) and processed to separate $\geq$200 macroinvertebrates for identification. A dissecting microscope (12$\times$ magnification) was used for processing, and all individuals in a subsample (including 1st-instar larvae) were removed. Three composite samples were collected per site, so this protocol resulted in >600 specimens per site, which is more than the 500 individuals commonly suggested as adequate to describe species richness at a site (e.g., Cao et al. 2002). Each specimen was photographed, placed in a separate voucher vial, identified first by an amateur (a college student with family-level taxonomy skills) and then by expert taxonomists with an average of 23 y experience.

DNA and sequence analyses

DNA was sequenced from a piece of each specimen (a leg for most insects, posterior parapods for chironomid Diptera, or a piece of body tissue for other dipteran insects, oligochaetes, crustaceans, and snails). Genomic mitochondrial DNA was extracted, and the 657 base pair (bp) barcoding region of the cytochrome c oxidase subunit 1 (COI) gene was amplified and sequenced using standard barcoding protocols (Ivanova et al. 2006) at the University of Guelph Canadian Centre for DNA Barcoding. DNA was extracted in a 96-well lysis plate that contained 50 μL of insect lysis buffer with proteinase K. Approximately ½ of the samples were extracted robotically using the Biomek FX-1 (Beckman Coulter, Inc., Fullerton, California), and DNA was eluted in 40 μL H_2O. The remaining samples were extracted manually, and the DNA was eluted in 25 μL H_2O. Full-length COI barcodes were amplified using the M13-tailed versions of the Folmer primers: LCO1490-t1 (5'-TGTAAAACGACGGCCAGTGGTCAACAAA-TCATAAAGATATTGG-3')/HCO2198-t1 (5'-CAGGA-AACAGCTATGACTAAACTTCAGGGTGACCAAA-AAATCA-3') (Folmer et al. 1994). PolyLCO and polyHCO primers were used when full-length polymerase chain reaction (PCR) amplification was not successful (Carr 2010). Each PCR had a total volume of 12.5 μL and contained 5% trehalose (D-(+)-trehalose dehydrate), 1.25 μL of 10$\times$ reaction buffer, 2.5 mM $MgCl_2$, 1.25 pmol each forward and reverse primer, 50 μM dNTP (Promega Corp., Madison, Wisconsin), 0.3 units of platinum *Taq* DNA polymerase (Invitrogen Corp., Carlsbad, California), and 2 μL of genomic DNA. PCR products were examined using 2%

agarose E-Gel® 96-well system (Invitrogen) and imaged using an AlphaImager workstation (Alpha Innotech, San Leandro, California) and E-Editor software (version 2.0.2, Invitrogen). Successful PCR products were sent for cycle sequencing. The remaining PCR product was diluted (~1:4) and prepared for bidirectional cycle sequencing via BigDye v3.1 using an ABI 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, California) as explained in Hajibabaei et al. (2005) and deWaard et al. (2008). Contiguous sequences (contigs) were generated and edited using Sequencer 4.5 software (Gene Codes Corp., Ann Arbor, Michigan). Edited sequences were aligned using MEGA 3.1 software (http://www.megasoftware.net/index.html) and uploaded onto the Barcode of Life Datasystems (BOLD) website (http://www.barcodinglife.com/). Sequences and detailed information about all specimens including photographs are stored on GenBank and are publicly accessible within the project 'Stroud Water Research Center - White Clay Creek 2008 Study'.

MEGA 4.0 (Tamura et al. 2007) was used to determine genetic divergence for specimens with DNA barcodes >350 bp except for Plecoptera (>600 bp) and to create profiles of neighbor-joining (NJ) trees of Kimura-2-parameter distances for each insect order. A total of 1786 specimens were available for barcoding. Turbellaria flatworms (4), Acari (31), and Oligochaete segmented worms (131) were removed from further analysis because of sequencing problems, as were 3 specimens suspected of handling contamination. Of the remaining 1617 specimens (mostly insects), 98% were successfully barcoded. Of those, 97% had 657-bp sequences.

Data analysis

Macroinvertebrate data were summarized as community-structure metrics (e.g., richness and diversity) commonly used in water-quality monitoring programs. Every effort was made by the amateur and experts to identify all individuals to the lowest taxonomic level possible. However, because of small size and poor condition, some specimens were left at genus, family, or order by taxonomists. The following protocol was used to determine richness (i.e., the number of species or taxa in a sample) for specimens identified morphologically. If a specimen of a given genus in a sample was identified to species (e.g., *Teloganopsis deficiens* (Morgan)), and another specimen was collected in that sample and identified to the same genus but not to the species level (e.g., left at *Teloganopsis*), then the 2nd specimen was not counted as representing a new and distinct taxon. Likewise, if

TABLE 1. Percentage of specimens identified by expert entomologists to species, genus, or higher taxonomic levels. Number in parentheses is the total number of individuals examined. All barcoded taxa refers to specimens that were successfully barcoded and were used in all analyses. All taxa sampled includes individuals that were omitted from analyses because of problems during barcoding (oligochaetes [131], Acari [31], EPT [23], other [22]).

Taxonomic group	% species	% genus	% higher than genus
Ephemeroptera (415)	70	15	15
Plecoptera (47)	2	0	98
Trichoptera (153)	75	24	1
Nonchironomid Diptera (114)	59	36	5
Chironomidae (702)	40	56	4
Coleoptera (146)	36	30	34
Other (2)	0	0	100
All barcoded taxa (1579)	51	36	12
All taxa sampled (1786)	46	33	21

a specimen in a given family (Ephemerellidae) was identified to genus (e.g., *Teloganopsis*) but not to species, then other specimens collected in that sample and identified to the same family (e.g., Ephemerellidae) but not to the level of genus were not counted as representing a new and distinct taxon. This approach is conservative and prevents overestimation of richness because it treats small or damaged specimens as redundant taxonomic units rather than as unique species. These small or damaged individuals may not have contributed to estimates of richness, but they were used in calculations of all other metrics (see Results for information about taxonomic richness based on DNA barcodes).

The fundamental levels of identification compared in our study were amateur (largely family level), expert (genus–species level), and barcode (species level), but a 4th level was created (*expert genus*) by changing all species-level identifications provided by our experts to the level of genus. This level was created to mimic the typical level of taxonomy (i.e., genus) associated with state and federal monitoring projects and most work done by environmental consulting firms. Expert data identified to genus–species level was called *expert species* even though only 51% of the specimens were identified to species and 36% to genus (Table 1).

All metrics were calculated in 3 ways: 1) combining the 3 samples and using them in their entirety, 2) taking the average of the 3 samples in their entirety, and 3) using a rarefaction method. In rarefaction methods, all samples are standardized so that measures of community structure based on richness are not biased by the number of individuals examined

in a sample (Ricklefs and Miller 2000, Gotelli and Colwell 2001). Our rarefaction process used a resampling (1000 random draws without replacement) routine to standardize samples to 150 individuals (SAS statistical package, version 8; SAS Institute, Cary, North Carolina).

In addition to the several measures of richness, other common water-quality monitoring metrics including Beck's Index (Beck 1955), Shannon Index (Weaver and Shannon 1949), Simpson's Index (Simpson 1949), and Hilsenhoff Biotic Index (HBI; Hilsenhoff 1977) were calculated. In an effort to quantify differences within the Chironomidae (which was abundant and speciose at both sites), total richness and the Shannon and Simpson indices also were calculated with only chironomid data (i.e., Chironomidae only) and a standardized sample of 60 individual chironomids in the rarefaction process.

Values derived from Smith and Voshell (1997) for family-level data, unpublished genus–species-level data obtained in 2005 from the US Environmental Protection Agency, and genus-level data for Pennsylvania (PA DEP 2009) and Maryland streams (Southerland et al. 2005) were used with metrics that required tolerance values (i.e., Intolerant and Tolerant richness, Beck's Index, and HBI). Differences in metrics among identification levels (amateur, expert genus, expert species, barcode) or sites were tested using analyses of variance (ANOVAs) followed by a Tukey's test or a Student's t-test (equal variances), respectively. To avoid a Type I error (indicating a difference when one is not present), a sequential Bonferroni correction was used to adjust the significance level for similar derived metrics that were grouped under richness or indices. To avoid a Type II error (i.e., failing to indicate a difference when one was present), significance was also judged at $\alpha = 0.05$ because of our small sample size ($n = 3$).

Results

Macroinvertebrate inventory

An analysis of the number of pairwise comparisons as a function of % genetic divergence for barcode data was used to establish the appropriate level of genetic divergence for distinguishing species in each major macroinvertebrate group (Fig. 2A–F). Our approach was to look for distinct regions of 0 divergence (a barcode gap) separating clusters of significant divergence in the data, which we think represent differences in levels of genetic divergence among species, genera, families, etc. Different levels of divergence were used to distinguish species, depending on the macroinvertebrate group. A 2% level of genetic

divergence was used to separate species for Ephemeroptera, Trichoptera, Coleoptera, and nonchironomid Diptera, 3% for Plecoptera, and 4% for Chironomidae. For Plecoptera, most specimens were too small to identify beyond the family level, so the sequence of our specimens was matched to known sequences to determine that 36 of the 48 specimens belonged to the genus *Perlesta*. Pairwise comparisons of *Perlesta* sequences showed no apparent gap (0 divergence), but the remaining Plecoptera had divergences of 2% (Fig. 2B). Neighbor-joining trees based on barcode with associated names based on morphological identifications by expert entomologists are in Figs 3 and 4.

Overall, across both study sites, the total number of taxa distinguished by amateur, expert genus, expert species, and barcode approaches were 26, 67, 88, and 150, respectively (Table 2). For some groups (e.g., Plecoptera and Chironomidae) where amateur identifications were limited to order or family, the number of taxa identified increased markedly with experts and doubled again with barcodes. These increases were apparent at both sites—experts identified ~3× the number of taxa at each site than did amateurs, and barcoding increased the number of taxa at each site another ~34 to 38%. This difference between expert species and barcode was greatest for the Chironomidae, the most abundant and diverse macroinvertebrate family in our study. For example, experts were unable to identify 60% of the chironomids to species (Table 1). For the other 40%, the experts recognized 24 different species, of which 7 were rare (≤ 2 individuals or $\leq 0.1\%$ of total individuals). Of the remaining 17 more-common species, 9 of the morphological designations by experts (19% of the total chironomids) agreed with barcodes (i.e., all individuals named a species formed 1 haplotype cluster). In 1 instance, experts split specimens into 2 species (viz., *Eukiefferiella*) but barcoding indicated only 1 species (haplotype cluster *Eukiefferiella* 1; Fig. 4). For the remaining 7 species (15% of the total chironomids), barcodes indicated that the specimens assigned to these species represented from 2 to 4 cryptic species for each of them [i.e., *Cricotopus tremulus* (Linnaeus), *Orthocladius dorenus* (Roback), *Orthocladius oliveri* Soponis, *Polypedilum convictum* (Walker), *Tanytarsus glabrescens*, *Tanytarsus guerlus* (Roback), and *Thienemanniella xena* (Roback); Fig. 4]. Almost every major group of macroinvertebrate in the study contained clusters of specimens that were morphologically indistinguishable and were lumped at the generic level by experts. In every case, barcodes could be used to separate them into 2 to 11 distinct genetic groups or species (e.g., 2 Ephemeroptera: *Maccaffertium* [Fig. 3A],
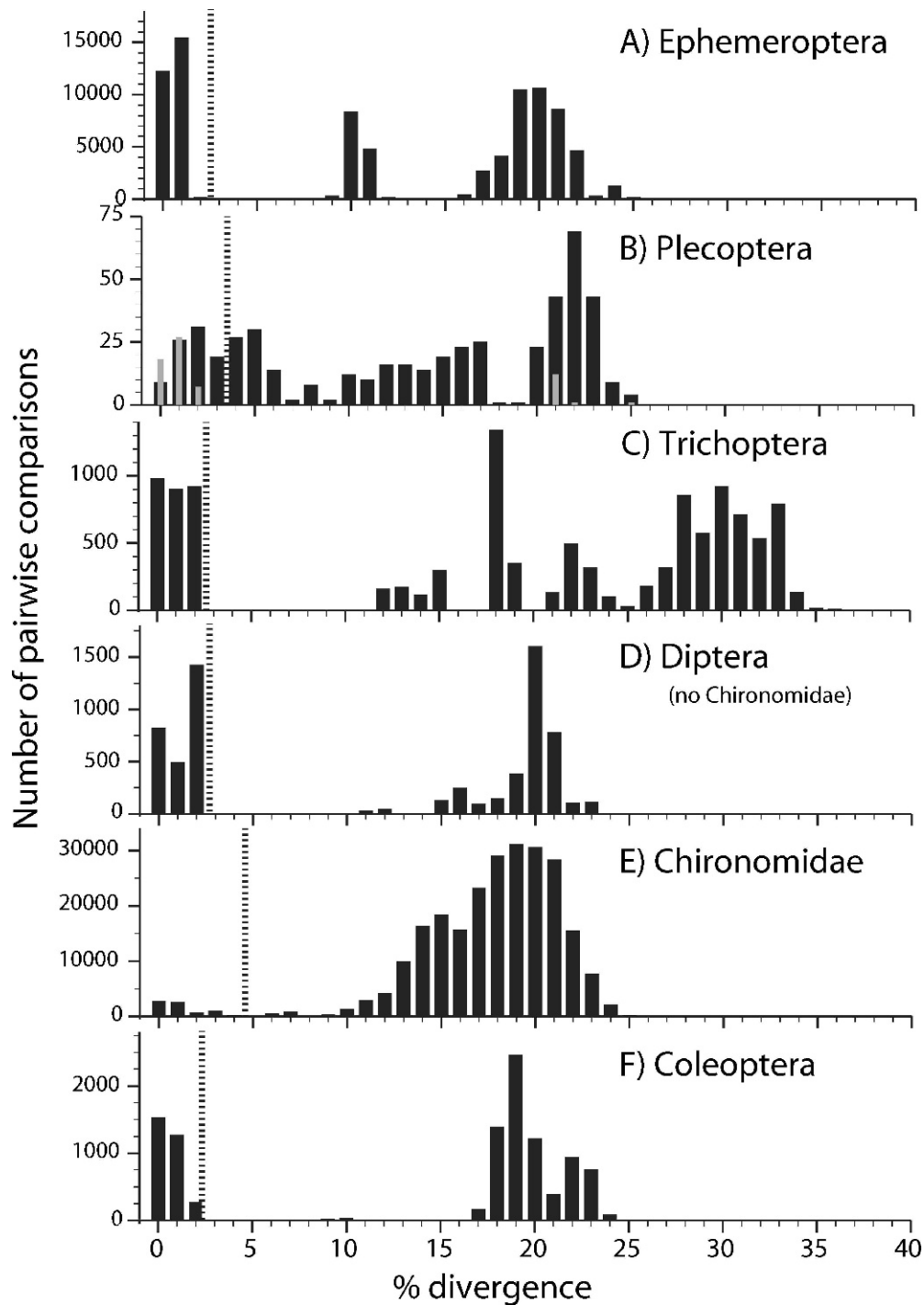
FIG. 2. Number of pairwise comparisons vs % genetic divergence for Ephemeroptera (A), Plecoptera (B), Trichoptera (C), Diptera without Chironomidae (D), Chironomidae (E), and Coleoptera (F) from White Clay Creek, Pennsylvania. The vertical dashed line indicates the value of genetic divergence used to separate species (2% for Ephemeroptera, Trichoptera, Coleoptera, and nonchironomid Diptera; 3% for Plecoptera; 4% for Chironomidae). Plecoptera are represented as *Perlesta* (black) and other stoneflies, mainly Leuctridae (gray).
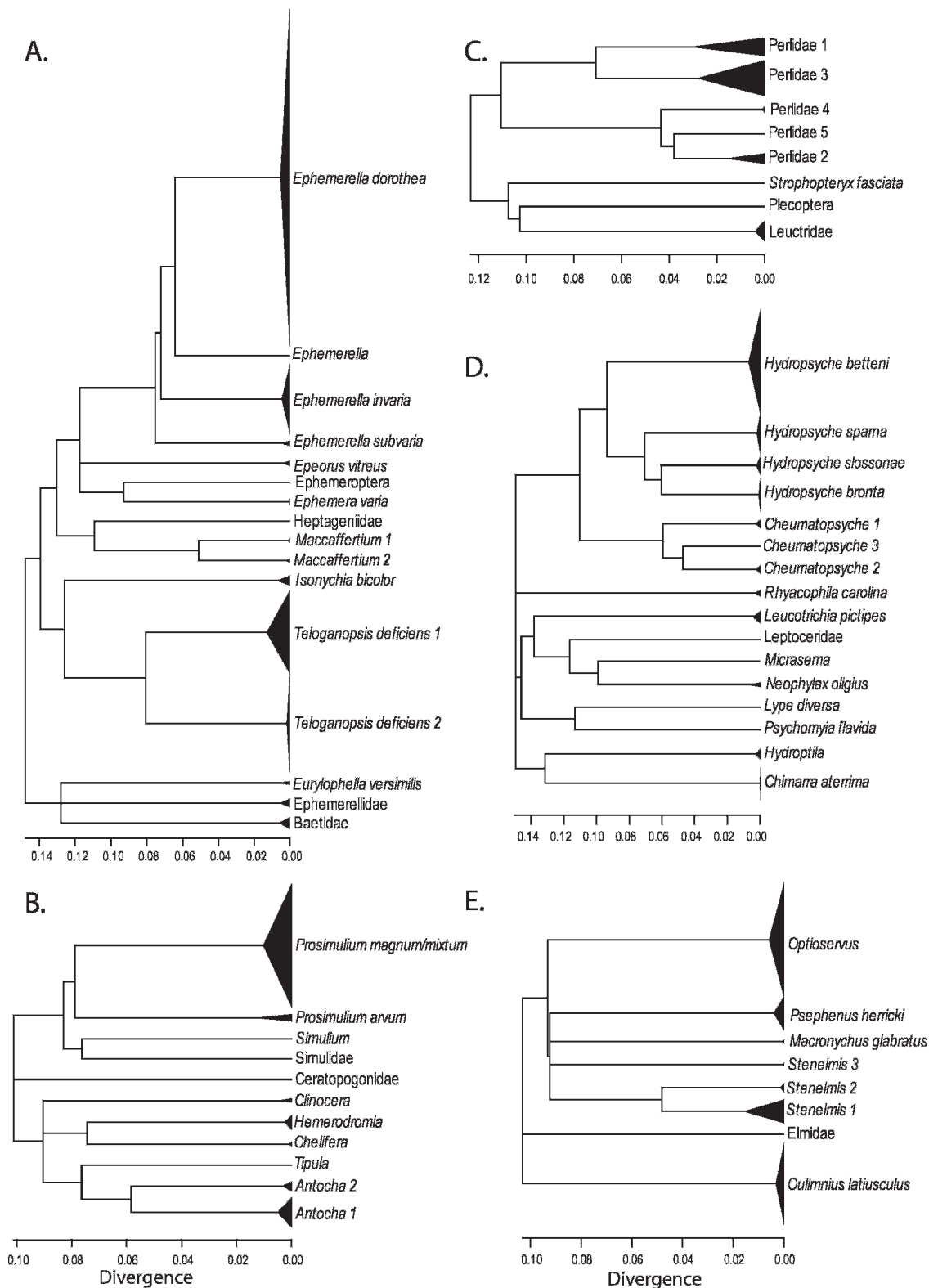
Fig. 3. Neighbor-joining groupings of barcode COI haplotype diversity with morphology-based names applied for Ephemeroptera (A), nonchironomid Diptera (B), Plecoptera (C), Trichoptera (D), and Coleoptera (E) collected from White Clay Creek, Pennsylvania. For each taxon, the vertical distance of triangle is proportional to number of specimens collected and horizontal length represents maximum genetic diversity within the branch. Numbers after names indicate the number of barcode taxa within the morphology-based taxon.
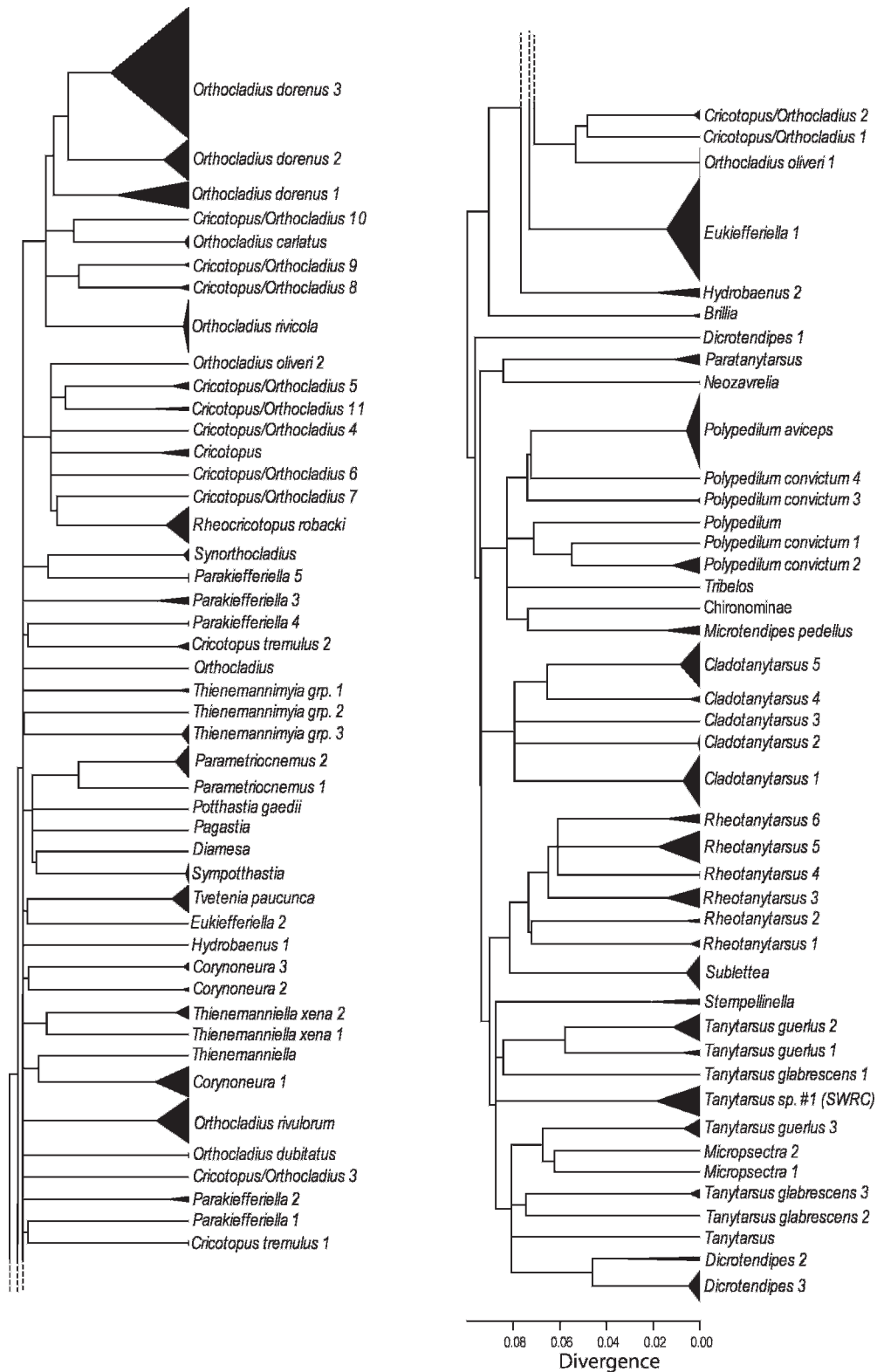
Fig. 4. Neighbor-joining groupings of barcode COI haplotype diversity with morphology-based names applied for Diptera:Chironomidae larvae collected from White Clay Creek, Pennsylvania. For each taxon, the vertical distance of triangle is proportional to number of specimens collected and horizontal length represents maximum genetic diversity within the branch. Numbers after names indicate the number of barcode taxa within the morphology-based taxon.

TABLE 2.   Richness determined by amateur (A), expert genus (EG), expert species (ES), and barcode (B) for samples collected at 2 sites from White Clay Creek. Sites were considered to have good (Site 11) and fair (Site 12) water quality.

Taxonomic group	Site 11 $n = 743$ individuals 28% Chironomidae				Site 12 $n = 836$ individuals 59% Chironomidae				Both sites $n = 1579$ individuals 45% Chironomidae			
	A	EG	ES	B	A	EG	ES	B	A	EG	ES	B
Plecoptera	1	2	2	7	1	2	2	3	1	3	3	8
Ephemeroptera	4	7	9	11	4	5	6	9	6	8	10	16
Trichoptera	8	7	10	12	6	7	8	8	10	11	14	16
Nonchironomid Diptera	3	5	7	7	4	6	7	10	4	8	10	11
Chironomidae	1	21	25	45	1	27	36	70	1	29	42	89
Coleoptera	2	5	5	6	2	5	6	6	2	6	7	8
Other	1	1	1	1	1	1	1	1	2	2	2	2
Total richness	20	48	59	89	19	53	66	107	26	67	88	150

3 Trichoptera: *Cheumatopsyche* [Fig. 3D], 5 Diptera: *Cladotanytarsus* [Fig. 4], 6 Diptera: *Rheotanytarsus* [Fig. 4], 11 Diptera: *Cricotopus/Orthocladius* [Fig. 4]). Barcoding even revealed cryptic species in the well-known Ephemeroptera—all *Teloganopsis* larvae were identified as *T. deficiens* by experts, but barcoding revealed 2 genetically distinct species (Fig. 3A).

Rare species (i.e., ≤2 individuals or 0.1%) at the 2 WCC sites made up 42% of the 150 total barcode species collected or 5% of total individuals. Eighty-six of the barcode species had ≤5 individuals (i.e., ≤0.3% of total individuals). Only 5 barcode species had ≥50 individuals collected (Fig. 5). Both sites exhibited the same abundance structure for species, but site 12 had higher densities of macroinvertebrates than site 11, and 93 more individuals (836 vs 743) were identified at site 12 than at site 11 (Table 2). The 2 sites also differed in taxonomic composition based on barcode and expert taxonomy (but not at the amateur level; Table 2). Intersite differences in taxonomic structure occurred among both rare and common species. For example, of the 15 most common barcode species, ⅔ occurred either uniquely

or almost uniquely (>85%) at one or the other site (Fig. 6). Across all species, 43 of 89 barcode species were unique to site 11 and 60 of 107 were unique to site 12, and some but not all were rare (Table 3). Fifty-five of the 88 species identified by experts were present at both sites, resulting in a Jaccard Similarity Index of 63%. By contrast, only 47 of the 150 barcode species occurred at both sites, resulting in a Jaccard Similarity Index of only 31%, <½ of either the expert or the amateur level.

No resampling vs rarefaction

Rarefaction and no resampling (3 samples averaged) reduced total taxon richness estimates relative to the entire sample for both expert species and barcode at both sites, but only for rarefied expert genus data at site 12 and not at all for expert genus data at site 11 and amateur data at either site (Fig. 7A, B). This result was expected a priori because most species in the expert species and barcode data sets were rare and, thus, would be lost when the number of specimens was reduced by rarefaction. This
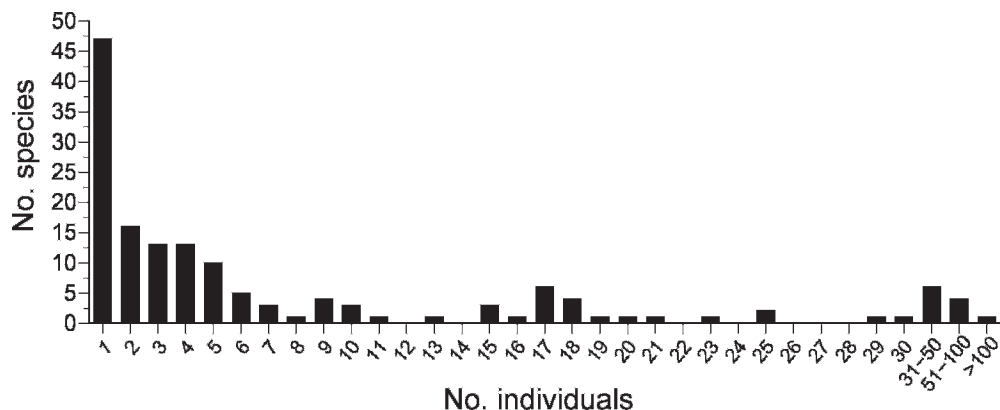

FIG. 5.   Relative abundance of macroinvertebrate barcode taxa collected in White Clay Creek.
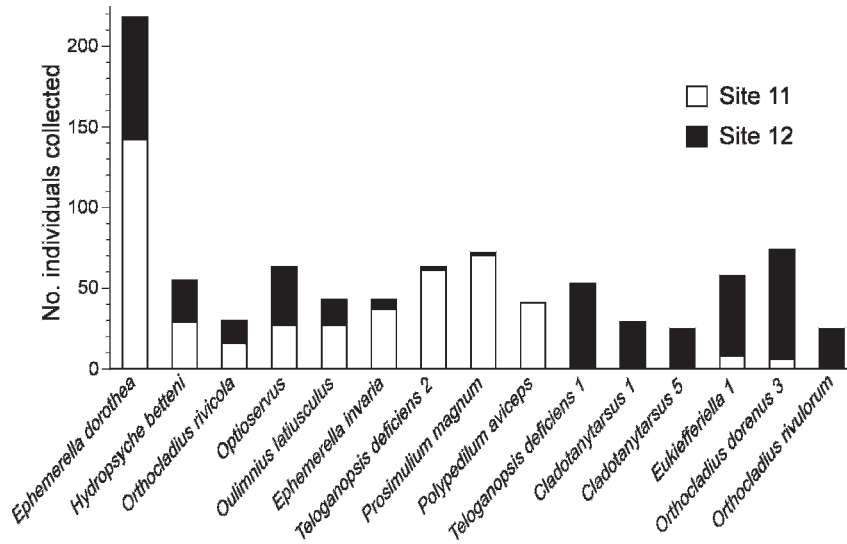
Fig. 6.   Relative abundance of the most common (≥ 25 individuals collected) barcode taxa in White Clay Creek.

difference did not occur for the expert genus and amateur data because of the coarser level of identification. Thus, the loss of richness caused by rarefaction was least for the amateur data and increased with increased levels of taxonomic discrimination. Despite the reduction in richness caused by rarefaction, all richness values based on amateur identifications remained lower than values based on expert genus, expert species, and barcode (Fig. 7A, B).

The accumulation curve (using rarefied data) showing the number of chironomid taxa based on barcode data as a function of number of chironomid individuals in a given sample differed from the curves based on expert genus and expert species identifications (e.g., site 12, Fig. 8). The slope of the barcode-species accumulation curve decreased, but it did not reach an asymptote even with a sample of 500 individuals. The expert genus and expert species accumulation curves reached an asymptote with 400 individuals at 27 and 36 taxa, respectively.

As identification method changed from expert genus to expert species to barcode, the number of

common chironomid taxa (≥ 3 individuals collected at each site) decreased whereas the number of unique abundant (≥ 3 individuals collected at only 1 site) and unique rare (≤ 2 individuals collected at only 1 site) taxa increased (Fig. 9). Overall, only 7 and 24% of the total chironomid density from barcoding were common to both sites 11 and 12, respectively.

Water-quality analysis: amateur vs expert species vs barcode

Using rarefied data, we calculated 14 metrics for amateur data and 17 for expert species and barcode data (Table 4). Given that our results were vulnerable to both Type I and II errors, we report and discuss the statistics with the Bonferroni correction applied and with $\alpha = 0.05$ (i.e., uncorrected). Sixty and 80% (80 and 90%; uncorrected) of the 10 richness metrics increased for sites 11 and 12, respectively, when the level of taxonomy changed from amateur to expert species. In contrast, 0 and 27% (36 and 55%; uncorrected) of the 11 richness metrics increased for sites 11 and 12, respectively, when the taxonomy changed from expert species to barcode.

Indices related to diversity and evenness.—Shannon and Simpson Indices both increased (corrected and uncorrected) for sites 11 and 12 when expert species data were compared with amateur data. In contrast, only the Simpson Index at site 11 increased (and only for uncorrected α) when expert species data were compared with barcode data. For chironomid-only data, no significant (corrected or uncorrected) change in either the Shannon or the Simpson Index was observed when calculated with expert species and barcode data.

TABLE 3.   Number of barcode species that were exclusive to Site 11 (43) and Site 12 (60). Numbers in parentheses indicate the number of rare taxa (≤ 2 individuals).

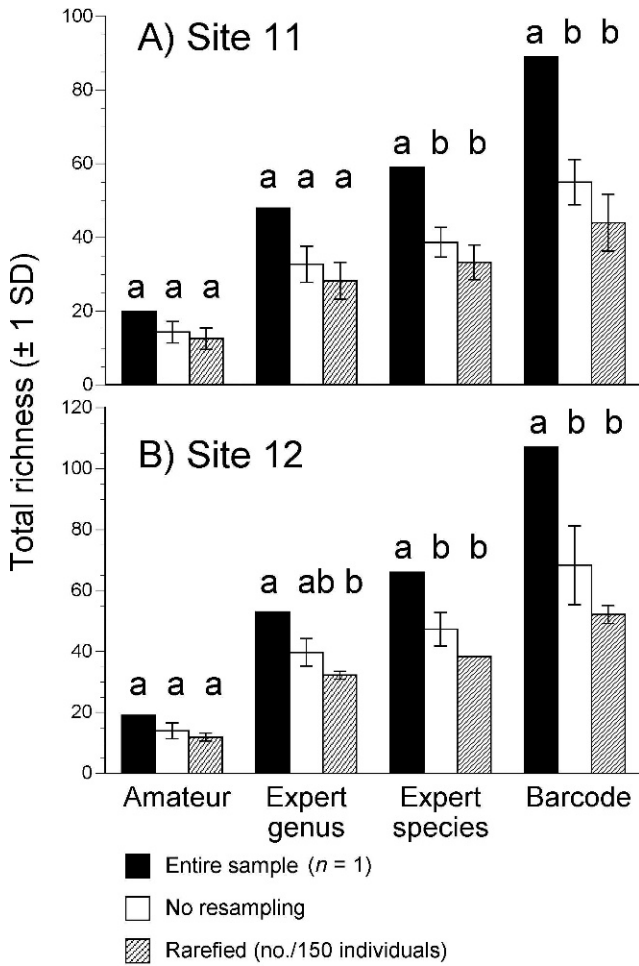| Taxonomic group	Site 11	Site 12
Plecoptera	5 (2)	1 (1)
Ephemeroptera	7 (1)	5 (4)
Trichoptera	8 (3)	4 (2)
Nonchironomid Diptera	1 (1)	4 (3)
Chironomidae	19 (15)	43 (22)
Coleoptera	2 (2)	2 (1)
Other (Ancylidae, Nemertea)	1 (1)	1 (1)

FIG. 7. Mean (± 1 SD) total richness of macroinvertebrates determined by 4 identification methods (amateur, expert genus, expert species, and barcode) at sites 11 (A) and 12 (B) on White Clay Creek, Pennsylvania. Richness values were determined from entire sample (3 samples combined) with no resampling (based on average of 3 samples) or with rarefaction (150 individuals resampled 1000 times). Within each identification method, bars with the same letters are not significantly different (analysis of variance, Tukey's test, $p > 0.05$).

Indices related to pollution.—Both HBI and Beck's Index changed at site 12 (corrected) and Beck's Index at site 11 (uncorrected) when expert species data were compared with amateur data. However, only Beck's Index changed (uncorrected) when barcode data were compared with expert species data.

Water-quality analysis: expert genus vs barcode

We compared metrics (based on rarefied data) calculated with expert genus and barcode data. Nine and 55% (73 and 64%; uncorrected) of the 11 richness metrics increased for sites 11 and 12, respectively.
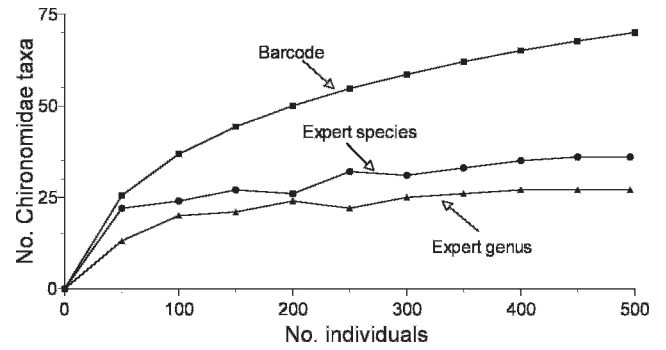

FIG. 8. Accumulation curves for Chironomidae taxa based on expert genus, expert species, and barcode identification for site 12 on White Clay Creek. Richness values were determined from 3 pooled samples with rarefaction (increments of 25 individuals were resampled 1000 times).

Shannon and Simpson Indices both increased (corrected) for site 12 but not 11. HBI did not change for either site and Beck's Index changed only at site 11 (uncorrected only) in response to the expert genus to barcode change in taxonomy. Comparison across all metrics showed that, at an uncorrected $\alpha = 0.05$, the percentages of metrics changed by improving taxonomy from expert genus to barcode were 59% and 65% for sites 11 and 12, respectively. These percentages were substantially higher than the percentage of metrics changed at the corrected $\alpha = 0.05$ level by improving taxonomy from expert species to barcode (i.e., 35 and 47% for sites 11 and 12, respectively).
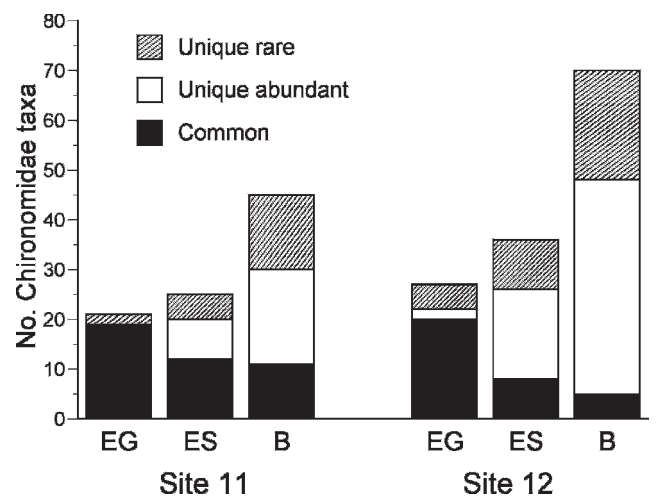

FIG. 9. Number of chironomid taxa determined by expert genus (EG), expert species (ES), and barcode (B) identifications. Taxa were designated as common (≥ 3 individuals collected at each site), unique abundant (≥ 3 individuals collected at only 1 site), or unique rare (≤ 2 individuals collected at only 1 site).

TABLE 4. Metrics calculated based on amateur (A), expert species (ES), expert genus (EG), and barcode (B) identifications (ID) for macroinvertebrates collected at 2 sites. Metric values for A vs ES, EG vs B, and ES vs B were compared with paired t-tests on rarefied samples (150 individuals resampled 1000 times for all metrics except Chironomidae-only metrics for which 60 individuals were used). The significance level was adjusted with sequential Bonferroni correction for richness and index metrics and significant p values are shown. Nearly significant results ($0.05 \leq p > \alpha$ based on Bonferroni correction) are indicated by asterisks. EPT = Ephemeroptera, Plecoptera, Trichoptera; TV = tolerance value; HBI = Hilsenhoff Biotic Index; n/a = not applicable.

| | Comparisons of ID methods by sites | | | | | | | | | | | |
| | Site 11 | | | | | | Site 12 | | | | | |
Metric	A vs ES	p	EG vs B	p	ES vs B	p	A vs ES	p	EG vs B	p	ES vs B	p
Richness												
Total	13 vs 33	0.003	28 vs 44	*	33 vs 44		12 vs 38	<0.001	32 vs 52	<0.001	38 vs 52	0.001
EPT	9 vs 15	0.006	11 vs 18	*	15 vs 18		6 vs 9	0.004	8 vs 10	0.006	9 vs 10	*
Ephemeroptera	3 vs 6	<0.001	5 vs 7	0.002	6 vs 7		2 vs 3	0.016	3 vs 4		3 vs 4	
Plecoptera	1 vs 2	*	2 vs 4	*	2 vs 4	*	1 vs 1		1 vs 2		1 vs 2	
Trichoptera	4 vs 6		4 vs 7	*	6 vs 7		3 vs 4	*	4 vs 5		4 vs 5	
Diptera	2 vs 16	0.003	15 vs 23		16 vs 23		4 vs 25	<0.001	21 vs 37	<0.001	25 vs 37	0.002
Chironomidae	1 vs 14	0.005	14 vs 21		14 vs 21		1 vs 22	<0.001	18 vs 34	<0.001	22 vs 34	0.003
Chironomidae only	n/a		15 vs 25	*	16 vs 25	*	n/a		15 vs 28	<0.001	19 vs 28	
Intolerant (TV $\leq$ 3)[a]	6 vs 3		9 vs 7		3 vs 7	*	4 vs 2	<0.001	6 vs 4		2 vs 4	*
Intolerant (TV $\leq$ 5)[a]	9 vs 19	0.007	18 vs 27	*	19 vs 27	*	8 vs 24	<0.001	20 vs 29	*	24 vs 29	
Tolerant (TV $\geq$ 6)[a]	2 vs 4	*	3 vs 5	*	4 vs 5	*	3 vs 7	<0.001	5 vs 10	0.010	7 vs 10	*
Indices												
Shannon	1.8 vs 3.3	<0.001	2.8 vs 3.1		3.3 vs 3.1		1.5 vs 3.6	<0.001	3.1 vs 3.5	<0.001	3.6 vs 3.5	
Simpson	0.77 vs 0.94	0.007	0.89 vs 0.91		0.94 vs 0.91		0.63 vs 0.96	<0.001	0.93 vs 0.96	<0.001	0.96 vs 0.96	
Chironomidae-only Shannon	n/a		2.4 vs 2.8	*	2.7 vs 2.8		n/a		2.4 vs 3.0	<0.001	2.9 vs 3.0	
Chironomidae-only Simpson	n/a		0.88 vs 0.92		0.92 vs 0.91		n/a		0.86 vs 0.93	<0.001	0.93 vs 0.93	
HBI[a]	4.5 vs 4.1		4.1 vs 4.0		4.1 vs 4.0		5.1 vs 4.7	0.002	4.6 vs 4.6		4.7 vs 4.6	
Beck's[a]	7 vs 14	*	15 vs 23	*	14 vs 23	*	6 vs 11	0.010	12 vs 16		11 vs 16	*

[a] Uses tolerance values

Table 5. Metrics determined using amateur, expert genus, expert species, and barcode identifications (ID) for macroinvertebrates collected at 2 sites. Differences between sites (11 vs 12) were compared using paired t-tests on rarefied samples (150 individuals resampled 1000 times for all metrics except Chironomidae-only metrics for which 60 individuals were used). The significance level was adjusted with sequential Bonferroni correction for richness and index metrics and significant p values are shown. Nearly significant results ($0.05 \leq p > \alpha$ based on Bonferroni correction) are indicated by asterisks. EPT = Ephemeroptera, Plecoptera, Trichoptera; TV = tolerance value; HBI = Hilsenhoff Biotic Index; n/a = not applicable.

| | Comparisons of sites by ID method | | | | | | | |
| | Amateur | | Expert genus | | Expert species | | Barcode | |
Metric	11 vs 12	p	11 vs 12	p	11 vs 12	p	11 vs 12	p
Richness								
Total	13 vs 12		28 vs 32		33 vs 38		44 vs 52	
EPT	9 vs 6	*	11 vs 8	*	15 vs 9	0.003	18 vs 10	0.007
Ephemeroptera	3 vs 2	*	5 vs 3	*	6 vs 3	<0.001	7 vs 4	0.003
Plecoptera	1 vs 1		2 vs 1		2 vs 1		4 vs 2	*
Trichoptera	4 vs 3		4 vs 4		6 vs 4	*	7 vs 5	*
Diptera	2 vs 4		15 vs 21	*	16 vs 25	*	23 vs 37	*
Chironomidae	1 vs 1		14 vs 18		14 vs 22	*	21 vs 34	*
Chironomidae only	n/a		15 vs 15		16 vs 19		25 vs 28	
Intolerant (TV $\leq$ 3)[a]	6 vs 4		9 vs 6	*	3 vs 2		7 vs 4	*
Intolerant (TV $\leq$ 5)[a]	9 vs 8		18 vs 20		19 vs 24	*	27 vs 29	
Tolerant (TV $\geq$ 6)[a]	2 vs 3	*	3 vs 5	*	4 vs 7	0.004	5 vs 10	*
Indices								
Shannon	1.8 vs 1.5		2.8 vs 3.1	*	3.3 vs 3.6	*	3.1 vs 3.5	*
Simpson	0.77 vs 0.63	*	0.89 vs 0.94	*	0.94 vs 0.96	*	0.91 vs 0.96	0.005
Chironomidae only Shannon	n/a		2.4 vs 2.4		2.7 vs 2.9	*	2.8 vs 3.0	
Chironomidae only Simpson	n/a		0.88 vs 0.86		0.92 vs 0.93		0.91 vs 0.93	*
HBI[a]	4.5 vs 5.1	*	4.1 vs 4.6	*	4.1 vs 4.7	*	4.0 vs 4.6	0.008
Beck's[a]	7 vs 6		15 vs 12		14 vs 11		23 vs 16	*

[a] Uses tolerance values

Water-quality analysis: site 11 vs site 12

Using the same approach as above (rarefaction data, 14–17 metrics, corrected/uncorrected α values), detection of differences between sites 11 and 12 depended on both the level of taxonomy and whether we applied a Bonferroni correction (Table 5). The proportion of metrics identifying differences between sites 11 and 12 varied little when we compared data generated by amateur (0%), expert (genus [0%], species [18%]), and barcode (24%) identifications and used a Bonferroni correction. However, the proportion of metrics identifying differences for the same comparisons was much greater for an uncorrected α (amateur [36%], expert genus [47%], expert species [65%], barcode [76%]).

Of the richness metrics examined, Ephemeroptera, Plecoptera, Trichoptera (EPT); Ephemeroptera; and Tolerant were the only measures to indicate a decline in water quality from site 11 to site 12 at all 4 levels of taxonomic resolution. With Bonferroni correction, only metrics calculated with expert species (EPT, Ephemeroptera, Tolerant) and barcode (EPT, Ephemeroptera) data decreased significantly. Across all richness measures, absolute richness values increased at both sites when taxonomy changed from expert genus to expert species to barcode. However, the upward change was proportional for each site, so the statistically significant differences between sites 11 and 12 did not change substantially. Richness metrics involving pollution-tolerant groups (Diptera, Chironomidae, and Tolerant) all increased (uncorrected) from site 11 to site 12 based on expert genus, expert species, and barcode data (except expert genus Chironomidae), but with Bonferroni correction, only Tolerant richness calculated with expert species data increased significantly.

Of the diversity and pollution indices examined, only HBI indicated a decline in water quality from site 11 to site 12 with all 4 levels of taxonomic resolution, although with Bonferroni correction, the differences were significant only for barcode data. In contrast,

Shannon and Simpson Indices indicated an increase in diversity and evenness from site 11 to site 12 when calculated with expert genus, expert species, and barcode data, but with the Bonferroni correction, only Simpson Index calculated with barcode data increased significantly. Beck's Index did not discriminate between sites based on amateur, expert genus, or expert species data, but indicated a slight decrease (uncorrected α) in stream condition from site 11 to site 12 with barcode data.

The number of midge species identified at site 12 vs site 11 increased by 6 (expert genus), 11 (expert species), and 25 (barcode) species (Table 2), but rarified chironomid richness did not differ significantly when calculated using only chironomids between sites for any of these methods of identification (Table 5). Shannon and Simpson Indices calculated only for chironomids did not differ between sites based on expert genus data. When calculated only for chironomids, Shannon Index differed significantly (uncorrected) with expert species data and Simpson Index differed significantly (uncorrected) with barcode data. Based on both richness and indices using only chironomids, the general distribution of individuals among chironomid species was similar between sites, even though both rare and common species that characterized each site differed.

Discussion

Our study was designed to use DNA barcoding of stream macroinvertebrates to assess the limits of current, morphologically based taxonomic tools as applied in biodiversity inventories, community-structure studies, and stream bioassessments. To a certain extent, our study was a conservative assessment of the potential contribution of barcoding because the macroinvertebrate fauna of White Clay Creek is well known after 40 y of intensive study by the Stroud Water Research Center and because our expert taxonomists were more experienced (an average of 23 y, much of it working on the macroinvertebrate fauna of this stream) than the average ecological or water-quality technician. Overall, our results suggest that barcoding is a tool that can greatly further our understanding of stream macroinvertebrate faunas by providing a more accurate description of the species present, what they are doing ecologically, and how they respond to changes in stream habitat and water quality.

Macroinvertebrate inventory

Our results confirm that full, or nearly full, sequences of the COI gene can be obtained readily from properly preserved specimens and used to sort specimens, regardless of size or morphological condition, into fairly unique haplotype clusters representing taxa at or near the level of species. For the most part, interpretation of genetic divergence was not confounded by a lack of a barcode gap (sensu Alexander et al. 2009) for all the major groups of macroinvertebrates (except for *Perlesta* stoneflies). The levels of divergence for the various taxonomic groups (e.g., 2% for mayflies and caddisflies, 4% for chironomids, etc.) worked well within the narrow scope of the study and were comparable to findings in other studies (e.g., Avise and Walker 1999, Hebert et al. 2003, Hajibabaei et al. 2006, Stahls and Savolainen 2008, Zhou et al. 2009). For example, our chosen level of divergence for chironomids (4%) agrees well with the mean intraspecific and interspecific values of 2.3% and 14.7%, respectively, for selected chironomids reported by Sinclair and Gresens (2008). The absence of a gap between intraspecific and interspecific genetic variation for *Perlesta* specimens (Fig. 2B) may reflect the presence of ≥2 emerging or very closely related congeneric species presently co-occurring in the stream or the presence of numts (nuclear copies of mitochondrial-derived genes that act as junk DNA; Buhay 2009, Moulton et al. 2010, Pfrender et al. 2010). Regardless, it is premature to conclude that the levels of divergence used here are applicable across a broader range of species and geographic sites.

Barcoding increased species inventory across these 2 sites by 124 taxa relative to an amateur taxonomist and by 62 taxa (70%) relative to expert taxonomists (identifying to species where possible). Thus, the barcoding inventory was about 5.7× the amateur inventory, and 1.7× the expert inventory. The greater number of species delineated by barcoding relative to experts reflects, in part, the ability of barcoding to identify small, immature, or damaged specimens. For example, only 1 of 47 stonefly specimens could be identified morphologically to species because of the constraints of specimen size at this time of year and damage from collection, but barcodes placed all stonefly specimens into 1 of 8 genetically distinct taxa (Fig. 3C). The greater number of taxa delineated by barcoding also reflects the fact that some species contained within certain insect orders in this stream cannot be identified morphologically based on the currently available taxonomic keys. In some cases, our study simply confirmed the incompleteness of our taxonomic understanding for certain groups (e.g., the caddisfly *Cheumatopsyche* and the chironomid midges *O. dorenus*, *P. convictum*, and *Ta. guerlus*), which has long been acknowledged in the taxonomic literature. However, barcoding revealed some previously unrec-

ognized taxonomic issues associated with some well-known species (e.g., the ephemerellid mayfly *T. deficiens*). These results suggest a continued need to improve available taxonomic keys and that some scientific names in the literature may be confounded with cryptic species, but many taxonomic names held up under the scrutiny of barcoding. Thus, barcodes do not represent a fix or replacement for a broken system—they simply make the system better.

Our ability to sort all specimens in the samples to genetically distinct species and enumerate them has, to our knowledge, provided the most thorough taxonomic inventory (based on larvae) of riffle-inhabiting macroinvertebrates at a given point in time for any stream reach in the world. The high number of macroinvertebrate species (150 across 2 nearby sites) collected on 1 occasion (March 2008) from 1 micro-habitat (riffle) is unprecedented. The total would increase greatly if the study were expanded to other microhabitats (e.g., edges, pools, leaf packs, and debris dams) in different seasons and multiple years. Because stream macroinvertebrates were abundant and distributed relatively evenly among species, additional taxonomic resolution from barcoding gave us a new perspective on stream communities. Based on data from our composite sampling method, an average of 248 to 347 individuals representing ≥ 55 to 68 species can and do coexist in only 60 cm^2 of stream bottom. Our data confirm that stream macroinvertebrate communities are similar in structure to other floral and faunal communities in the world (Ricklefs and Miller 2000): most species are rare and only a few are common. The numerous rare species contributed greatly to the relative uniqueness in the taxonomic composition that characterized the 2 study sites. Barcodes indicated that 48% of the species at site 11 and 56% at site 12 were unique to a site. Compared to species resolution from barcoding, traditional taxonomic analyses would have underestimated the differences between the 2 sites.

Regardless of the temporal and spatial limitations in our study, the number of macroinvertebrate species (89 and 107 for sites 11 and 12, respectively) collected on 1 day in March from 1 microhabitat (riffle) is very high. These data suggest that α diversity in streams and rivers may be greatly underestimated when based on larval morphology (e.g., Jackson et al. 1994, Arscott et al. 2006). This conclusion is shared by Alexander et al. (2009) for *Ephemerella* stream mayflies and by Hebert et al. (2004) and others for terrestrial macroinvertebrates. Barcoding also has the potential to increase estimates of biodiversity for stream inventories based on adults collected over many sites, microhabitats, and years (e.g., 642 insect

species for the Breitenbach Creek near Schlitz, Germany; Allan and Castillo 2007). With high individual and species densities and with the co-occurrence of numerous congeneric species, ecological fitting (sensu Janzen 1985) appears greater in stream riffles than expected. Based on the large differences in taxonomic composition of sites 11 and 12, one might conclude that levels of β diversity along the river continuum (sensu Vannote et al. 1980) have been greatly underestimated. However, we think that, in this case, the observed differences in community composition primarily reflect a change in species composition in response to changes in land and water uses that have degraded and changed the habitat at site 12 (as indicated by physical, chemical, and biological data that span ~15 y; Fig. 1).

Macroinvertebrate community structure: amateur vs expert species vs barcode

Measures of community structure have long been part of ecology and provide simple values that summarize patterns or responses for multiple species (e.g., Washington 1984, Metcalfe 1989, Cairns and Pratt 1993). They have provided insight into a variety of basic topics ranging from disturbance to succession, from physical habitat templates to biogeographic patterns. Assessments of biological responses to pollution often involve monitoring responses of multiple species, so bioassessments also have made use of traditional measures of community structure (e.g., Wilhm and Dorris 1968, Weber 1973, Winner et al. 1980) and modified measures of community structure that incorporated information on pollution tolerance (e.g., Beck 1955, Chutter 1972, Hilsenhoff 1977, Lenat 1993). Our study used 17 community-structure indices or measures to examine the effects of different levels of taxonomic resolution on the description of the macroinvertebrate community at each site and how macroinvertebrate communities differed between sites 11 and 12 (see below). Some redundancy exists among metrics (e.g., EPT richness vs Ephemeroptera, Plecoptera, or Trichoptera richness, Diptera richness vs Chironomidae richness, Intolerant richness calculated as ≤ 3 or ≤ 5, etc.), but all metrics examined respond, at least to some degree, to different levels of taxonomic resolution (i.e., order vs family vs genus vs species) and are used routinely in stream studies.

As expected, processing samples by experts (identifying to species level where possible) rather than amateurs increased values (by 60–90% depending on α level and site) of the 10 taxon-richness metrics used to describe the 2 sites. Barcoding consistently yielded

additional species for every richness metric studied, but the statistical significance of the change in richness based on expert species vs barcode identifications depended greatly on whether we applied a Bonferroni correction. From the uncorrected comparisons, we concluded that barcode identifications significantly increased richness metrics for both sites. From the corrected comparisons (e.g., Bonferroni), we concluded that barcode identifications increased richness metrics only for site 12. However, application of the Bonferroni correction is quite subjective (Cabin and Mitchell 2000). In fact, we could argue with equal vigor that each of our metrics should be considered independent (hence, no correction is needed) and that the same null hypothesis is tested 11 times across our 11 richness metrics (hence a correction is needed). Regardless, the conclusion that barcode identifications increased richness metrics for site 12 may be closer to reality because: 1) the relative increase in richness consequent to barcoding for taxonomically well-known insect orders (e.g., Ephemeroptera, Plecoptera, and Trichoptera) was about equal at both sites (and rather modest because sites had been studied intensively for >40 y in White Clay Creek), whereas 2) large increases in species richness because of barcoding were observed for poorly known taxonomic groups (e.g., chironomid midges) that were much more abundant at site 12 than site 11 (59% of the community vs 28%). So, a reasonable conclusion from these results is that barcoding can and does significantly improve richness metrics (over expert), especially when a significant proportion of the community being evaluated consists of species that are poorly known or described.

The above conclusion seems almost trivial, but it may apply broadly across most, if not all, water-quality studies because chironomid midges are a significant proportion of the macroinvertebrate community of most streams worldwide (pristine to polluted). Midges are known to be a very diverse group worldwide (Pinder 1986), but they remain poorly known ecologically and taxonomically at the genus and species levels. In addition, they are very difficult to identify to genus and species, even with substantial training. For example, amateurs in our study could not identify chironomids beyond family and, although the experts could take most specimens to genus, only 40% of specimens could be identified to species. In contrast, the barcode technique discovered many new or unrecognized species of existing chironomid genera (and for other major groups), thereby providing greater depth to the level of assessment and understanding of the macroinvertebrate communities. In addition, because all specimens in the study were preserved and vouchered individually, they can now be sorted by haplotype cluster and provided to taxonomists as presorted groups of genetically similar specimens. This presentation makes it easier for taxonomists to distinguish intraspecific from interspecific variation in morphological characters, which in turn, increases their effectiveness at revising and extending the breadth and utility of morphological keys, an important side benefit of DNA barcoding. Another side benefit is that barcoding can be used to confirm morphological traits when revising keys (Liew et al. 2009, Monaghan and Sartori 2009). For example, in our study, barcode data indicated only 1 case in which our experts needed to synonymize 2 species (i.e., into *Eukiefferella* 1), but multiple cases in which barcodes indicated that several species were grouped under a single genus (e.g., *Cheumatopsyche*, *Stenelmis*, *Rheotanytarsus*) or species name (e.g., *T. deficiens*, *P. convictum*, *Ta. guerlus*).

Water-quality analysis: amateur vs expert species vs barcode

We combined richness, relative abundance, and taxon-specific pollution-tolerance values in various ways to produce 6 indices related to water-quality and pollution assessment (Table 4). In general, values differed in 50 to 100% (depending on α level) of these indices with the increase in taxonomic effort from amateur to expert species, but only 17 to 33% of these indices differed when taxonomy changed from expert species to barcode. For example, both of the commonly used measures of diversity that we examined (Shannon and Simpson Indices) increased with the addition of expert species taxonomy relative to amateur taxonomy largely because amateurs tended to place most individuals into a few taxonomic categories, whereas experts distributed the identifications more evenly among a larger number of categories. Neither Shannon nor Simpson Index responded greatly to the addition of barcode data (relative to expert species). Barcode data distinguished more species than expert data, but most of these species were rare, and these indices are insensitive to the contribution of rare species (DeJong 1975).

Indices that used pollution-tolerance values (e.g., HBI, Beck) probably underestimated the degree of improvement in water-quality assessment because tolerance values usually are available at the family and genus level and rarely at the species level. We were forced to assume that all congeneric species had the same tolerance value, even though the unique distribution of species at the 2 sites suggested that distinct differences in pollution tolerance existed among confamilial or congeneric species in White

Clay Creek. This problem explains why HBI and Beck's Index indicated no (HBI) or little (Beck's Index) improvement with increased taxonomic refinement from expert to barcode because many species had to be given the same tolerance value even if they were known to be unique species. Only at the more-degraded site 12 did these indices strongly indicate (i.e., corrected α) less-degraded conditions with improved taxonomy from amateur to expert. This change probably occurred because indices become more accurate when tolerance values can be used at the genus and species level (expert and barcode) vs family level (amateur). Thus, many richness measures were sensitive to changes in taxonomic resolution in our study (particularly at the more disturbed site), but the absence of accurate pollution-tolerance values reduced the usefulness of the added taxonomic resolution from barcoding when tolerance-related metrics were calculated.

Herein lies the challenge. Barcode identification is now pushing larval macroinvertebrate taxonomy beyond the currently available pollution-sensitivity and tolerance data. We think that this *pollution-tolerance gap* can be closed quickly by developing libraries relating Linnaean names to COI and other gene sequences. At the same time, it is imperative that we connect ancillary pollution data and taxon delineations at the species (or near-species) level in the form of new or revised pollution-tolerance values. Perhaps equally important is the need to develop alternative metrics designed to take advantage of the new information afforded by barcoding, especially in groups like chironomids, whose species contain a wealth of information on environmental quality that can be now accessed through molecular methods.

Our results do not suggest that barcoding will make significant contributions to all measures of community structure. Metrics known or thought to be unaffected by increased taxonomic resolution (e.g., family- or order-based metrics, such as % chironomids or % EPT; trophic measures, such as % scrapers and % shredders; and habit measures, such as % clingers or % sprawlers) were not included in our study. Similarly, multimetric indices consisting of several metrics may benefit only partially from the added taxonomic resolution associated with barcoding. For example, the Maryland Index of Biotic Integrity combines genus-level data into 6 metrics: total richness, EPT richness, Ephemeroptera richness, % pollution-intolerant individuals, % chironomids, and % clingers (individuals adapted to inhabiting the fast water of riffles) (Southerland et al. 2005). Two of these metrics (% chironomids, % clingers) largely reflect family-level characteristics. One metric (%

pollution-intolerant individuals) relies on pollution-tolerance values that presently differ little within a genus because we have not derived species-specific values. Three metrics (total, EPT and Ephemeroptera richness) would have been affected substantially by barcoding because most larvae (including very small or damaged specimens) would have been identified to species level rather than genus, and most genera were represented by ≥2 species. This conclusion could change if it were discovered that pollution-tolerance values differ among closely related barcode species (e.g., the different *T. deficiens* species at sites 11 and 12), and if multimetrics were reconstructed to make more use of the biodiversity actually in streams.

Water-quality analysis: expert genus vs barcode

The value of species identifications remains a point of contention among authors of studies looking at pollution responses and among those looking at small- and large-scale biogeographic patterns. Some have found that species-level (and in some cases genus-level) identifications have provided little additional information, whereas others observed a measurable contribution. One factor in these different conclusions is the steepness of the environmental gradients inherent in the different studies. In cases where the gradient is steep and numerous families and even orders may be affected by the environmental stressors, coarse taxonomic resolution provides much the same resolution as genus/species identifications (e.g., Bowman and Bailey 1997, Bailey et al. 2001, Arscott et al. 2006). In contrast, studies like ours in which environmental gradients are less steep and differences among sites involve species replacements within genera rather than the complete loss of genera or families, genus/species identifications have added meaningfully to efforts to differentiate among sites relative to similar efforts based on coarser taxonomic resolution (e.g., Resh and Unzicker 1975, Lenat and Resh 2001, Waite et al. 2004, Arscott et al. 2006, Pond et al. 2008). Another factor is that most (e.g., 60–80%) specimens in species studies were actually left at the genus or higher level (e.g., Waite et al. 2004, Arscott et al. 2006). Thus, most published studies have not been true family vs species or genus vs species contrasts.

We had the opportunity here to contrast directly water-quality analysis based on family vs genus vs species level identifications for 2 sites representing a shallow environmental gradient. In many ways, these contrasts provide insight into the relative change associated with water-quality assessments done by volunteers vs state/federal/professional personnel vs academicians, respectively. Our study shows that the

level of information and number of changes in water-quality metrics was notably higher when taxonomy improved from family to genus level (e.g., 158% increase in richness values and a change of up to 47% of the metrics). This finding is noteworthy but not novel and is why state, federal, and professional personnel monitor streams using genus-level macroinvertebrate data. However, that barcoding can improve the level of information provided at the genus level (e.g., richness) by another 124% and the number of metrics denoting change up to 76% is both noteworthy and novel.

Given the above, we think the potential contribution of barcode species identification to future biomonitoring efforts appears great. First, most ongoing macroinvertebrate monitoring protocols call for genus-level (or coarser) taxonomic resolution (Carter and Resh 2001), and programs that attempt species identifications are successful with only 20 to 40% of the specimens (e.g., Waite et al. 2004, Arscott et al. 2006). Thus, most ongoing protocols are dependent on detecting environmental disturbance substantial enough to cause marked changes in or local extinction of whole genera or families of macroinvertebrates. Also, identification errors appear to be relatively common, even at the genus level (e.g., 10–20%; Stribling et al. 2008). We assume that ongoing species-level identifications based solely on morphology are equally or more problematic than those at the genus level because of the inadequacy of many taxonomic keys and the skill level of the practicing taxonomist. With an adequate reference library, barcodes could attach names to almost all specimens and greatly reduce identification errors because discrimination decisions would be well documented and easily repeatable. Barcoding also can provide an invaluable service with regard to quality assurance/quality control of ongoing macroinvertebrate studies. Hence, reference specimens for each taxon in a given study could be checked quickly and accurately by either the group doing the assessment or the client or agency that contracted the work. This quality check could provide positive feedback for increasing the level of expertise of taxonomists associated with water-quality monitoring studies. Last, the need to contrast very clean with severely degraded sites will persist, but we think a need also exists to identify and manage more-subtle environmental changes (e.g., Cuffney et al. 2010), especially for studies undertaken to demonstrate level of success for environmental mitigation/restoration/conservation efforts. Barcodes can provide species identifications that will play a vital role in this effort because more coarse taxonomic data can not provide the necessary information.

Water-quality analysis: site 11 vs site 12

Many of the metrics examined in our study indicated that sites 11 and 12 differed. Loss of pollution-sensitive species and changes in biotic indices indicated that conditions at site 12 were degraded relative to site 11. Taxonomic resolution did not affect this general conclusion, although expert and barcode data were more sensitive than amateur data, and greater taxonomic resolution affected the perception of species lost because of degradation. For example, loss of EPT taxa from site 11 to site 12 was 8 species for barcode but only 6, 3, and 3 taxa based on expert species, expert genus, and amateur methods (Table 5).

Barcode data clearly convey a more accurate (and dramatic) image of the value of an ecosystem and the magnitude of species loss caused by environmental disturbance or change to the public, environmental managers, and decision makers. Barcode data finally would put biodiversity assessments for aquatic communities in terms comparable to those used for terrestrial ecosystems, in which estimates of biodiversity for herbaceous plants, shrubs, trees, birds, or mammals are never quantified or discussed at the level of genus or family. In our study, barcode and expert data provided very similar bioassessment results, but biodiversity present and lost was much higher when estimated from barcode data. This result suggests that barcode data might be more sensitive than expert or amateur data for measuring small differences in water quality. However, its effective use will require new metrics that take advantage of the information provided by DNA barcode technology.

Greater taxonomic resolution also revealed differences between sites 11 and 12 that were contrary to commonly observed patterns or assumed responses to increased environmental stress. For example, total richness was unchanged, but Shannon and Simpson Indices were greater at site 12 than at site 11. This result appears to reflect the large increase in Chironomidae richness at site 12 revealed by greater taxonomic resolution (expert and barcode). This response is not observed in most studies because investigators often identify chironomid midges to the family, subfamily, or tribe level. Lenat (1983) and Fore et al. (1996) suggested that chironomid richness may be a complicated metric to use because it can increase in response to moderate pollution, only to decrease later as environmental stress increases. Our data suggest that these responses are, at least in part, a function of the number of individual chironomids examined rather than a change in the structure of the chironomid assemblage. The two sites did not differ

statistically where the comparisons were based on fixed counts of chironomids alone. However, environmental stress was not so great that a few chironomid species would find conditions favorable while many other species would be eliminated, which is a classic response to environmental stress. We need to know more about the responses of individual chironomid species (most of which are rare) to disturbance before we can better understand and make use of their collective response to environmental stress. DNA barcode technology now makes it possible to document and quantify those responses.

Conclusions

Our study demonstrated that barcoding worked well without major changes in field and laboratory protocols. It enabled us, for the first time, to make a complete inventory of stream macroinvertebrate species collected for both ecological and bioassessment purposes, and in the process, highlighted some weaknesses in studies that use macroinvertebrates to assess stream condition. Increased taxonomic resolution by experts and barcodes provided more accurate site descriptions and better differentiation between sites than did amateur-level taxonomic resolution. The marked increase in the number of species identified, especially with barcodes, has the potential to improve our perception of the value of healthy stream and river ecosystems by better illustrating the remarkable biodiversity that they support naturally, and therefore, highlighting the significant loss of biodiversity that occurs when water and habitat are degraded.

We do not view the results of our study as a criticism of past taxonomic efforts nor do we think that present DNA technology can replace morphological taxonomy. However, the ability to identify very small specimens and to identify surprisingly common cryptic species are clear benefits of using barcodes (see also Sharley et al. 2004, Pfenninger et al. 2007, Sinclair and Gresens 2008, Stahls and Savolainen 2008, Alexander et al. 2009, Krosch et al. 2009, Zhou et al. 2010). The increasing availability and automation of barcode technology has the potential to contribute significantly to aquatic macroinvertebrate taxonomy by increasing the accuracy and rate at which new or unrecognized species are identified, described, and incorporated into morphological keys. This taxonomic growth, in turn, can lead to a better understanding of the overall biodiversity and basic ecological characteristics of these species and their role in stream and river monitoring and protection. This result was most evident for the chironomid midges (this single family

contained ~½ of species identified), but appears applicable to most, if not all, macroinvertebrate families in our study.

Last, our study confirms that the level of information and number of changes in water-quality metrics respond significantly when the level of taxonomy changes from family to genus and, more importantly, responds again when the level of taxonomy changes from genus to species. We suggest that barcoding creates the long-awaited opportunity to develop pollution-tolerance values at the species level. Our data show that we currently have both a pollution-tolerance gap (because barcoding pushed larval taxonomy beyond the available pollution-tolerance data) and a need for new or modified pollution-assessment metrics and multimetrics that can take full advantage of species-level taxonomy. Our findings suggest that water-quality monitoring projects by state, federal, and other professional organizations in the US and elsewhere can be improved significantly by barcoding because it facilitates accurate identifications beyond the family/genus level. We predict that the species-level identifications afforded by barcoding will increase interest in taxonomy and the efficiency of producing morphological keys at the species level for stream macroinvertebrates. We also predict that a combined morphological/molecular approach can and will provide a finer resolution for evaluating environmental change associated with both natural and anthropogenic processes and proactive mitigation/restoration/conservation efforts.

Literature Cited

Alexander, L. C., M. Delion, D. J. Hawthorne, W. O. Lamp, and D. H. Funk. 2009. Mitochondrial lineages and DNA barcoding of closely related species in the mayfly genus

Ephemerella (Ephemeroptera:Ephemerellidae). Journal of the North American Benthological Society 28:584–595.

ALLAN, J. D., AND M. M. CASTILLO. 2007. Stream ecology. Structure and function of running waters. Springer, Dordrecht, The Netherlands.

ARSCOTT, D. B., J. K. JACKSON, AND E. B. KRATZER. 2006. Role of rarity and taxonomic resolution in a regional and spatial analysis of stream macroinvertebrates. Journal of the North American Benthological Society 25:977–997.

AVISE, J. C., AND D. WALKER. 1999. Species realities and numbers in sexual vertebrates: perspectives from an asexually transmitted genome. Proceedings of the National Academy of Sciences of the United States of America 96:992–995.

BAILEY, R. C., R. H. NORRIS, AND T. B. REYNOLDSON. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. Journal of the North American Benthological Society 20:280–286.

BARBOUR, M. T., 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.

BECK, W. H. 1955. Suggested method for reporting biotic data. Sewage and Industrial Waste 27:1193–1197.

BOWMAN, M. F., AND R. C. BAILEY. 1997. Does taxonomic resolution affect the multivariate description of the structure of freshwater benthic macroinvertebrate communities? Canadian Journal of Fisheries and Aquatic Sciences 54:1802–1807.

BUCHWALTER, D. B., D. J. CAIN, C. A. MARTIN, L. XIE, S. N. LUOMA, AND T. GARLAND. 2008. Aquatic insect ecophysiological traits reveal phylogenetically based differences in dissolved cadmium susceptibility. Proceedings of the National Academy of Sciences of the United States of America 105:8321–8326.

BUCHWALTER, D. B., AND S. N. LUOMA. 2005. Differences in dissolved cadmium and zinc uptake among stream insects: mechanistic explanations. Environmental Science and Technology 39:498–504.

BUHAY, J. E. 2009. "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. Journal of Crustacean Biology 29:96–110.

CABIN, R. J., AND R. J. MITCHELL. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. Bulletin Ecological Society of America 81:246–248.

CAIRNS, J., AND J. R. PRATT. 1993. A history of biological monitoring using benthic macroinvertebrates. Pages 10–27 in D. M. Rosenberg and V. H. Resh (editors). Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall, New York.

CAO, Y., D. P. LARSEN, R. M. HUGHES, P. L. ANGERMEIER, AND T. M. PATTON. 2002. Sampling effort affects multivariate comparisons of stream assemblages. Journal of the North American Benthological Society 21:701–714.

CARR, C. M. 2010. The Polychaeta of Canada: exploring diversity and distribution patterns using DNA bar-

codes. MS Thesis, University of Guelph, Guelph, Ontario.

CARTER, J. L., AND V. H. RESH. 2001. After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. Journal of the North American Benthological Society 20:658–682.

CHUTTER, F. M. 1972. An empirical biotic index of the quality of water in South African streams and rivers. Water Research 6:19–30.

CUFFNEY, T., R. BRIGHTBILL, J. MAY, AND I. WAITE. 2010. Responses of benthic macroinvertebrates to environmental changes associated with urbanization in nine metropolitan areas. Ecological Applications 20:1384–1401.

DAILY, G., AND K. ELLISON. 2001. The new economy of nature. Island Press, Washington, DC.

DEJONG, T. M. 1975. A comparison of three diversity indices based on their components of richness and evenness. Oikos 26:222–227.

DEWAARD, J. R., N. V. IVANOVA, M. HAJIBABAEI, AND P. D. N. HEBERT. 2008. Assembling DNA barcodes: analytical protocols. Pages 275–283 in C. C. Martin (editor). Environmental genomics, methods in molecular biology. Volume 410. Humana Press, Totowa, New Jersey.

FOLMER, O., M. BLACK, W. HOEH, R. LUTZ, AND R. VRIJENHOEK. 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294–299.

FORE, L. S., J. R. KARR, AND R. W. WISSEMAN. 1996. Assessing invertebrate responses to human activities: evaluating alternatives approaches. Journal of the North American Benthological Society 15:212–231.

FUNK, D. H., AND B. W. SWEENEY. 1990. Electrophoretic analysis of species boundaries and phytogenetic relationships of taeniopterygid stoneflies (Plecoptera). Transactions of the American Entomological Society 116:727–751.

FUNK, D. H., B. W. SWEENEY, AND R. L. VANNOTE. 1988. Electrophoretic study of Eastern North American *Eurylophella* (Ephemeroptera: Ephemerellidae) with the discovery of morphologically cryptic species. Annals of the Entomological Society of America 81:174–186.

GILLER, P. S., AND G. O'DONOVAN. 2002. Biodiversity and ecosystem function: do species matter? Proceedings of the Royal Irish Academy 102B:129–139.

GOTELLI, N. J., AND R. K. COLWELL. 2001. Quantifying biodiversity: procedures and pitfalls in the measurement and comparisons of species richness. Ecological Letters 4:379–391.

GRESENS, S. E., K. T. BELT, J. A. TANG, D. C. GWINN, AND P. A. BANKS. 2007. Temporal and spatial responses of Chironomidae (Diptera) and other benthic invertebrates to urban stormwater runoff. Hydrobiologia 575:173–190.

HAJIBABAEI, M., J. R. DEWAARD, N. V. IVANOVA, S. RATNASINGHAM, R. T. DOOH, S. L. KIRK, P. M. MACKIE, AND P. D. N. HEBERT. 2005. Critical factors for assembling a high volume of DNA barcodes. Philosophical Transactions of

the Royal Society of London Series B: Biological Sciences 360:1959–1967.

HAJIBABAEI, M., D. H. JANZEN, J. M. BURNS, W. HALLWACHS, AND P. D. N. HEBERT. 2006. DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences of the United States of America 103:968–971.

HEBERT, P. D. N., A. CYWINSKA, S. L. BALL, AND J. R. DEWAARD. 2003. Biological identification through DNA barcodes. Proceedings of the Royal Society of London Series B: Biological Sciences 270:313–321.

HEBERT, P. D. N., E. H. PENTON, J. M. BURNS, D. H. JANZEN, AND W. HALLWACHS. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proceedings of the National Academy of Science of the United States of America 101:14812–14817.

HILSENHOFF, W. L. 1977. Use of arthropods to evaluate water quality of streams. Department of Natural Resources Technical Bulletin No. 100. Department of Natural Resources, Madison, Wisconsin. (Available from: http://digicoll.library.wisc.edu/cgi-bin/EcoNatRes/EcoNatRes-idx?id=EcoNatRes.DNRBull100).

IVANOVA, N. V., J. R. DEWAARD, AND P. D. N. HEBERT. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6:998–1002.

JACKSON, J. K., AND V. H. RESH. 1992. Variation in genetic structure among populations of the caddisfly *Helicopsyche borealis* from three streams in northern California, U.S.A. Freshwater Biology 27:29–42.

JACKSON, J. K., AND V. H. RESH. 1998. Morphologically cryptic species confounding ecological studies of the caddisfly genus *Gumaga* (Sericostomatidae) in northern California. Aquatic Insects 20:69–84.

JACKSON, J. K., B. W. SWEENEY, T. L. BOTT, J. D. NEWBOLD, AND L. A. KAPLAN. 1994. Transport of *Bacillus thuringiensis* var. *israelensis* and its effect on drift and benthic densities of nontarget macroinvertebrates in the Susquehanna River, northern Pennsylvania. Canadian Journal of Fisheries and Aquatic Sciences 51:295–314.

JANZEN, D. H. 1985. On ecological fitting. Oikos 45:308–310.

KENNEY, M. A., A. E. SUTTON-GRIER, R. F. SMITH, AND S. E. GRESENS. 2009. Benthic macroinvertebrates as indicators of water quality: the intersection of science and policy. Terrestrial Arthropod Reviews 2:99–128.

KNOPP, M., AND R. CORMIER. 1997. Mayflies – an angler's study of trout water Ephemeroptera. Greycliff Publishing Company, Helena, Montana.

KROSCH, M. N., A. M. BAKER, B. G. MCKIE, P. B. MATHER, AND P. S. CRANSTON. 2009. Deeply divergent mitochondrial lineages reveal patterns of local endemism in chironomids of the Australian Wet Tropics. Austral Ecology 34:317–328.

LENAT, D. R. 1983. Chironomid taxa richness: natural variation and use in pollution assessment. Freshwater Invertebrate Biology 2:192–198.

LENAT, D. R. 1993. A biotic index for the southeastern United States – derivation and list of tolerance values, with criteria for assigning water quality ratings. Journal of the North American Benthological Society 12:279–290.

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.

LIEW, T., M. SCHILTHUIZEN, AND J. J. VERMEULEN. 2009. Systematic revision of the genus *Everettia* Godwin-Austen, 1891 (Mollusca: Gastropoda: Dyakiidae) in Sabah, northern Borneo. Zoological Journal of the Linnean Society 157:515–550.

METCALFE, J. L. 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. Environmental Pollution 60:101–139.

MONAGHAN, M. T., AND M. SARTORI. 2009. Genetic contributions to the study of taxonomy, ecology, and evolution of mayflies (Ephemeroptera): review and future perspectives. Aquatic Insects 31:19–39.

MOULTON, M. J., H. SONG, AND W. F. WHITING. 2010. Assessing the effects of primer specificity on eliminating numt coamplification in DNA barcoding: a case study from Orthoptera (Arthropoda: Insecta). Molecular Ecology 10:615–627.

PA DEP (PENNSYLVANIA DEPARTMENT OF ENVIRONMENTAL PROTECTION). 2009. Assessment and listing methodology for integrated water quality monitoring and assessment reporting. Clean Water Act Sections 305(b)/303(D). Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania. (Available from: <http://www.depweb.state.pa.us/watersupply/cwp/view.asp?a=1261&q=545541)

PFENNINGER, M., C. NOWAK, K. C. STEINKE, AND B. STREIT. 2007. Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic *Chironomus* (Diptera) species. Molecular Ecology 16:1957–1968.

PFRENDER, M. E., C. P. HAWKINS, M. BAGLEY, G. COURTNEY, B. R. CREUTZBURG, J. H. EPLER, S. FEND, L. C. FERRINGTON, P. L. HARTZELL, S. JACKSON, D. P. LARSEN, A. LEVESQUE, J. C. MORSE, M. PETERSEN, D. RUITER, D. SCHINDEL, AND M. WHITING. 2010. Assessing macroinvertebrate biodiversity in freshwater ecosystems: advances and challenges in DNA based approaches. Quarterly Review of Biology 85:319–340.

PINDER, L. C. V. 1986. Biology of freshwater Chironomidae. Annual Review of Entomology 31:1–23.

POND, G. J., M. E. PASSMORE, F. A. BORSUK, L. REYNOLDS, AND C. J. ROSE. 2008. Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. Journal of the North American Benthological Society 27:717–737.

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., R. H. Norris, and M. T. Barbour. 1995. Design and implementation of rapid assessment approaches for water-resource monitoring using benthic macroinvertebrates. Australian Journal of Ecology 20:108–121.

Resh, V. H., and J. D. Unzicker. 1975. Water quality monitoring and aquatic organisms: the importance of species identification. Journal Water Pollution Control Federation 47:6–18.

Ricklefs, R. E., and G. L. Miller. 2000. Ecology. 4th edition. W. H. Freeman and Company, New York.

Sharley, D. J., V. Pettigrove, and Y. M. Parsons. 2004. Molecular identification of *Chironomus* spp. (Diptera) for biomonitoring of aquatic ecosystems. Australian Journal of Entomology 43:359–365.

Simpson, E. H. 1949. Measurement of diversity. Nature 163: 688.

Sinclair, C. S., and S. E. Gresens. 2008. Discrimination of *Cricotopus* species (Diptera: Chironomidae) by DNA barcoding. Bulletin of Entomological Research 98: 555–563.

Smith, E. P., and J. R. Voshell. 1997. Studies of benthic macroinvertebrates and fish in streams within EPA Region 3 for development of biological indicators of ecological condition. Part 1. Benthic macroinvertebrates. Final report for Cooperative Agreement CF821462010. US Environmental Protection Agency, Washington, DC. (Available from: R. Voshell, Dept. of Entomology, 302 Price Hall, Virginia Tech, Blacksburg, Virginia 24061 USA.)

Southerland, M. T., G. M. Rogers, M. J. Kline, R. P. Morgan, D. M. Boward, P. F. Kazyak, R. J. Klauda, and S. A. Stranko. 2005. New biological indicators to better assess the condition of Maryland streams. DNR-12-0305-0100. Maryland Department of Natural Resources Monitoring and Non-Tidal Assessment Division, Annapolis, Maryland. (Available from: http://www.dnr.state.md.us/irc/digitaldocs.html)

Stahls, G., and E. Savolainen. 2008. MtDNA COI barcodes reveal crytic diversity in the *Baetis vernus* group (Ephemeroptera, Baetidae). Molecular Phylogenetics and Evolution 46:82–87.

Stribling, J. B., K. L. Pavlik, S. M. Holdsworth, and E. W. Leppo. 2008. Data quality, performance, and uncertainty in taxonomic identification for biological assessments. Journal of the North American Benthological Society 27: 906–919.

Sweeney, B. W., and D. H. Funk. 1991. Population genetics of the burrowing mayfly *Dolania americana*: Geographic variation and the presence of a cryptic species. Aquatic Insects 13:17–27.

Sweeney, B. W., D. H. Funk, and R. L. Vannote. 1987. Genetic variation in stream mayfly (Insecta: Ephemeroptera) populations in eastern North America. Annals of the Entomological Society of America 80:600–612.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular evolutionary genetic analysis (MEGA) software. Version 4.0. Molecular Biology and Evolution 24: 1596–1599.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130–137.

Waite, I. R., A. T. Herilhy, D. P. Larsen, N. S. Urquhart, and D. J. Klemm. 2004. The effects of macroinvertebrate taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands, USA. Freshwater Biology 49:474–489.

Washington, H. G. 1984. Diversity, biotic and similarity indices. A review with special relevance to aquatic ecosystems. Water Research 18:653–694.

Waters, T. F. 2000. Wildstream: a natural history of the free flowing river. Riparian Press, Saint Paul, Minnesota.

Weaver, W., and C. E. Shannon. 1949. The mathematical theory of communication. University of Illinois, Urbana, Illinois.

Weber, C. I. (editor). 1973. Biological field and laboratory methods for measuring the quality of surface water and effluents. EPA 670-4-73-001. Office of Research and Development, US Environmental Protection Agency, Cincinnati, Ohio.

Wilhm, J. L., and T. C. Dorris. 1968. Biological parameters for water quality criteria. BioScience 18:477–481.

Winner, R. W., M. W. Boesel, and M. P. Farrell. 1980. Insect community structure as an index of heavy-metal pollution in lotic ecosystems. Canadian Journal of Fisheries and Aquatic Sciences 37:647–655.

Zhou, X., K. M. Kjer, and J. C. Morse. 2007. Associating larvae and adults of Chinese Hydropsychidae caddisflies (Insecta:Trichoptera) using DNA sequences. Journal of the North American Benthological Society 26: 719–742.

Zhou, X., S. J. Adamowicz, L. M. Jacobus, R. E. Dewalt, and P. D. N. Hebert. 2009. Towards a comprehensive barcode library for artic life – Ephemeroptera, Plecoptera, and Trichoptera of Churchill, Manitoba, Canada. Frontiers in Zoology 6:30.

Zhou, X., L. M. Jacobus, R. E. Dewalt, S. J. Adamowicz, and P. D. N. Hebert. 2010. Ephemeroptera, Plecoptera, and Trichoptera fauna of Churchill (Manitoba, Canada): insights into biodiversity patterns from DNA barcoding. Journal of the North American Benthological Society 29: 814–837.