Combining data from multiple agencies to assess benthic macroinvertebrate communities in a large gravel-bed river

Matthew J. Wilson[1,4], Matthew E. McTammany[1,5], Michael D. Bilger[2,6], Sean P. Reese[3,7], and Benjamin R. Hayes[3,8]

[1]Biology Department, Bucknell University, Lewisburg, Pennsylvania 17837 USA
[2]EcoAnalysts, Selinsgrove, Pennsylvania 17842 USA
[3]Environmental Center, Bucknell University, Lewisburg, Pennsylvania 17837 USA

Abstract: Government agencies frequently conduct benthic macroinvertebrate surveys for bioassessment at large spatial scales in a variety of aquatic habitats, including large rivers. However, these data are rarely used by investigators outside the specific regulatory agency. We used data from 150 benthic macroinvertebrate samples collected over a period of 20 y from 10 locations in a large, shallow river system (the Susquehanna River and 2 major tributaries) by personnel in 4 government agencies to explore broad spatial and temporal patterns in benthic assemblages. We standardized sample size and taxonomy to account for differences in sampling, processing, and identification methods among agencies. Invertebrate assemblages were dominated by mayflies and caddisflies (46–83%). Percent Ephemeroptera, Plecoptera, Trichoptera (EPT) and standard diversity measures were inversely correlated, indicating that traditional macroinvertebrate Index of Biotic Integrity (IBI) approaches might not be applicable to large rivers. These data showed differences in assemblage composition across sub-basins and revealed effects of the spread of invasive Asian clams and of black fly management on benthic assemblage structure in the river. Large-river invertebrates are understudied and, even with challenges of combining data sets from multiple agencies, we showed the potential utility of applying data from a large river system to reveal ecological patterns across space and time.

Key words: benthic macroinvertebrate, large river, biomonitoring, bioassessment, Susquehanna River, taxonomy, methods

Limited published information regarding benthic macroinvertebrate assemblage structure is available from large, shallow, nonnavigable rivers (Rempel et al. 2000), but benthic macroinvertebrates are collected routinely by federal, state, and tribal agencies to assess the ecological health of streams and rivers for a variety of regulatory and monitoring purposes (e.g., PADEP 2012). Moreover, far more sites are sampled over time in the course of government monitoring programs than could be collected by private individuals or collaborative groups. For example, the Pennsylvania Department of Environmental Protection (PADEP) collects samples from ~1000 sites/y (G. Walters, Chief, Assessment Section, Division of Water Quality Standards, PADEP, personal communication). Data from government agencies also might provide information about benthic macroinvertebrate assemblages from streams not typically used for academic research because agencies often select sites based on spatially unbiased sampling designs (e.g., US Environmental Protection Agency [EPA] Environmental Monitoring and Assessment Program [EMAP]; USEPA 2002). These surveys could provide valuable data for large-scale ecological studies and for assessing changes in the 'biotic integrity' of streams and rivers over time. However, despite the considerable investment of time, funds, and energy, data from these surveys are rarely compared among agencies or over time. Thus, they do not contribute greatly to our understanding of river ecosystem structure and function. However, publicly available agency data could be used to provide basic ecological information about river macroinvertebrates, examine effects of basin-wide management practices (e.g., spraying *Bacillus thuringiensis israelensis* [*Bti*] to control black fly populations), track invasive species, or detect rare, threatened, or endangered species.

Significant challenges must be overcome to compare information from benthic macroinvertebrate surveys among agencies and over time. Inconsistent taxonomic resolution,

E-mail addresses: [4]mjw029@bucknell.edu; [5]mmctamma@bucknell.edu; [6]Present address: Susquehanna University, 514 University Avenue, Selinsgrove, Pennsylvania 17870 USA, bilgerm@susqu.edu; [7]spr016@bucknell.edu; [8]brh010@bucknell.edu

DOI: 10.1086/680962. Received 14 June 2013; Accepted 24 July 2014; Published online 24 Febuary 2015.
Freshwater Science. 2015. 34(2):593–605.

quality assurance, and variable expertise and experience of taxonomists are major hurdles to making basin-wide assessments with data from multiple agencies (Cuffney et al. 2007). Other issues include differences in sampling methods, processing methods, timing, frequency, and site location (Carter and Resh 2001). Moreover, standard sampling and sorting methods have changed over the past 20 y. For example, in 1999, the EPA standard protocols for rapid bioassessment of macroinvertebrates increased from a 100-count (Plafkin et al. 1989) to a 200-count (Barbour et al. 1999) procedure. Evaluating the quality of identification may be difficult when archived or reference specimens are discarded because of limited funds and storage space. This situation arises, in part, from sampling done to meet agency-specific priorities rather than to facilitate interagency comparisons or ecological research.

Cuffney et al. (2007) pioneered methods for resolving inconsistent taxonomy across data sets by addressing the issue of identification of organisms to different taxonomic levels (e.g., family and genus) in single or multiple surveys. These methods have been used to compare data collected by a single agency over time (Hamilton et al. 2010, Kennen et al. 2010, Stamp et al. 2010, King et al. 2011). To our knowledge, no one has attempted to standardize invertebrate survey data collected by multiple agencies at the same agency-specific sampling sites over time.

We expanded on the methods of Cuffney et al. (2007) to integrate benthic macroinvertebrate data collected by 4 government agencies from river sites across the Susquehanna River Basin (New York and Pennsylvania, USA) over a 20-y period into a single data set. Our goal was to make these data comparable by developing a method for standardizing surveys across multiple agencies and through time. We used these data to identify spatial and temporal patterns in invertebrate assemblages, and we assessed whether these data could be used to reveal ecological patterns of invertebrate assemblages in large rivers or to develop metrics for IBI development. We predicted that the West Branch and North Branch of the Susquehanna would have different macroinvertebrate assemblages because the branches differ in geology, hydrology, habitat structure, and dominant land use (coal mining in the West Branch and agriculture in the North Branch). We also expected the benthic macroinvertebrate assemblages in the Lower Susquehanna to be most similar to those in the North Branch, which provides ~60% of the flow to the lower river.

METHODS

Study sites

The Susquehanna River drains 71,227 km^2 and is 714 km long (SRBC 2006b) with an average discharge of 1061 m^3/s at Marietta, Pennsylvania, USA (USGS 2012). The watershed includes portions of New York, Pennsylvania, and Maryland. The river flows through the Northern Appalachian Plateau and Uplands, the Central Appalachian Ridges and Valleys, and the Northern Piedmont before reaching Chesapeake Bay. Basin geology is highly variable but consists primarily of shale, sandstone, and limestone formations (SRBC 2006a). The watershed is mostly forested (67%; 1.6% water and wetland), but the dominant land uses are agriculture (29%) and urban development (2.4%) (Boyer et al. 2002).

US governmental agencies have collected benthic macroinvertebrate samples from the Susquehanna River since 1968, after passage of the Pollution Control Act of 1948 (USEPA 1972). The New York Department of Environmental Conservation (NYDEC), PADEP, US Geological Survey (USGS), and Susquehanna River Basin Commission (SRBC) all routinely sample macroinvertebrates from the Susquehanna River Basin for bioassessment purposes. We assembled a geographical information system (GIS) database of macroinvertebrate sampling locations and dates from the Susquehanna River and its major tributaries between 1990 and 2010 (Fig. 1). We selected 10 of the 117 sites on the Susquehanna or major tributaries that have been sampled frequently, cover a broad spatial area, and were collected at 6^{th}–8^{th}-order reaches: Amity Hall on the Juniata River (a large tributary of the Lower Susquehanna), Wellsburg on the Chemung River (a large tributary of the North Branch), 3 sites (Bower, Williamsport, and Lewisburg) on the West Branch, 2 sites (Towanda and Danville) on the North Branch, and 3 sites (Sunbury, Fort Hunter, and Marietta) on the Lower Susquehanna below the confluence of the North and West Branches. These sites range in drainage area from 816 km^2 (Bower) to 67,314 km^2 (Marietta) and have been sampled multiple times since 1990 (Table 1).

Sampling methods of agencies

Agencies that contributed to our data set used different sampling and processing methods (Appendix S1). Personnel from all agencies collected kick samples from riffles or runs, but the mesh size used varied among agencies from 500 to 800 μm (PADER 1988, Plafkin et al. 1989, Barbour et al. 1999, Moulton et al. 2002, PADEP 2006, 2009, Smith et al. 2009). The greatest variation among protocols was in sample-processing methods. NYDEC personnel sorted all individuals >1.5 mm in length with the aid of a stereomicroscope until 100 individuals had been picked (Smith et al. 2009), whereas USGS personnel sorted with the aid of a stereomicroscope and removed all individuals encountered in 2 h of sorting time (Moulton et al. 2000). Personnel from other agencies used gridded sorting trays and picked randomly selected grid squares until 100 or 200 individuals were found (PADER 1988, Plafkin et al. 1989, Barbour et al. 1999, PADEP 2006). Most agency protocols required iden-

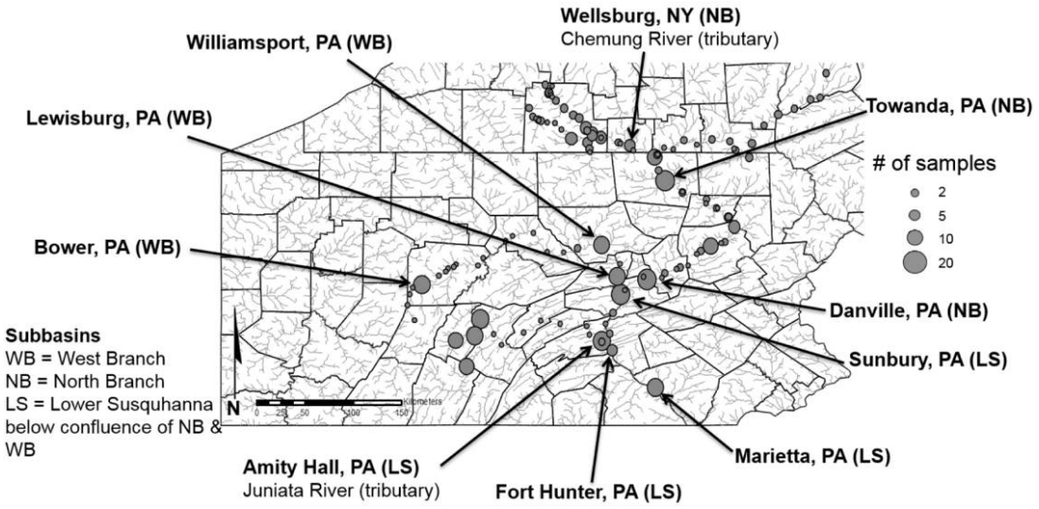

Figure 1. Sites in the Susquehanna River and tributaries for which benthic macroinvertebrate data are available from state and federal agency monitoring programs. Dots are scaled by the number of years a site was sampled between 1990 and 2010. Arrows indicate sites used in our study. NY = New York, PA = Pennsylvania.

tification of individuals to family or genus, but USGS protocol called for identification to the lowest possible level based on condition of the specimen.

Taxonomic standardization

We attempted to use genus-level identification for all taxa. However, many individuals were identified to an ambiguous level of classification (i.e., some individuals were identified to a more inclusive taxonomic level than other individuals in the same sample or data set). To standardize taxonomic resolution at the genus level, ambiguous 'parents' (individuals identified to a higher taxonomic level than other individuals within that taxon) were distributed among their 'children' (lower taxonomic levels within each

Table 1. Description of sampling sites on the Susquehanna River (USA) and information about benthic macroinvertebrate data. SRBC = Susquehanna River Basin Commission, PADEP = Pennsylvania Department of Environmental Protection, NYDEC = New York Department of Environmental Conservation, USGS = US Geological Survey.

Site	Latitude (°N)	Longitude (°W)	Sub-basin	Distance upstream from mouth (km)	Drainage area (km^2)	Number of samples (agency)
Bower	40.895	−78.676	West Branch	521	816	14 (1 SRBC, 13 PADEP)
Williamsport	41.226	−77.107	West Branch	262	14,716	15 (All PADEP)
Lewisburg	40.965	−76.877	West Branch	212	17,734	19 (1 USGS, 3 SRBC, 15 PADEP)
Wellsburg	42.0208	−76.7283	Chemung River (North Branch)	476	6491	13 (All SRBC, NYDEC Chironomidae at Chemung, NY)
Towanda	41.7631	−76.4375	North Branch	436	20,194	15 (1 SRBC, 14 PADEP)
Danville	40.9422	−76.6011	North Branch	224	29,060	16 (1 USGS, 2 SRBC, 13 PADEP)
Sunbury	40.8182	−76.8420	Lower Susquehanna	196	47,397	18 (1 USGS, 2 SRBC, 15 PADEP
Amity Hall	40.4258	−77.0159	Juniata River (Lower Susquehanna)	140	8687	19 (1 USGS, 2 SRBC, 16 PADEP)
Fort Hunter	40.3456	−76.9204	Lower Susquehanna	124	62,419	6 (2 SRBC, 4 PADEP)
Marietta	40.0372	−76.5236	Lower Susquehanna	72	67,314	15 (1 USGS, 2 SRBC, 12 PADEP)

ambiguous parent taxon). We used 'distribute parent among children' (DPAC) methods that were a combination of the 'single sample liberal' and 'grouped sample liberal' (DPAC-SL and DPAC-GL, respectively) methods described by Cuffney et al. (2007; see below for details). The liberal approach to DPAC distributes ambiguous individuals among child taxa based strictly on the relative abundances of the child taxa. This method is based on the assumption that individuals identified to genus accurately represent the relative abundances of genera within a family.

The taxonomy of several freshwater invertebrate taxa changed substantially from 1990 to 2010. For example, Brigham et al. (1982) recognized the caddisfly *Ceratopsyche* as a subgenus of *Symphitopsyche* (which is no longer recognized from North America) and included the *bifida* group of *Hydropsyche* (as defined by Ross 1944), but Wiggins (1996) considered *Ceratopsyche* as a subgenus of *Hydropsyche*. We observed a distinct change in our data set that corresponded to this taxonomic change. *Ceratopsyche* was relatively common in samples collected before 1997 yet was nearly absent in samples collected after 1997, probably because biologists began identifying *Ceratopsyche* as *Hydropsyche* to be consistent with taxonomic revisions. Therefore, we included individuals identified as *Ceratopsyche* with *Hydropsyche* across our entire data set. We accounted for other similar taxonomic issues by standardizing nomenclature in the data set based on the current classification for freshwater invertebrates by Wiggins (1996) for Trichoptera and Merritt et al. (2008) for other aquatic insect orders.

Macroinvertebrates were not always identified to the same taxonomic level (e.g., Chironomidae) by all agencies or in all samples processed by a single agency. Our goal was taxonomic standardization at the genus level, but individuals in some phyla and classes (i.e., Annelida and Arachnida, respectively) were not identified to the genus level in most surveys or were identified into groups that could not be classified to the genus level based on our methods (below). Therefore, Annelida were grouped as Oligochaeta or Hirudinea and all Arachnida were grouped as Hydracarina.

Sample size standardization The most common processing method was a fixed count of 100 or 200 individuals (mean and median = 120 individuals for all samples with <200 individuals). The USGS sorting method yielded samples with very large numbers of invertebrates (>7000 in 1 sample; Fig. 2A). Use of raw data would have given too much weight to samples and sites with extremely high numbers and would have biased richness estimates. Therefore, we subsampled every sample with >200 individuals to attain a sample size of 120 individuals. We subsampled by comparing the relative abundance of each taxon to a random number between 0 and 1. If the random number was less than the proportional abundance of the taxon in the sample, 1 individual was counted. If the random number was greater than the taxon's relative abundance, no individuals were counted. We repeated the simulation 120 times and summed the results to generate a 120-count subsample in which taxa with higher relative abundances in the original sample were counted more frequently. Rare taxa were less likely to be, but could be, included in the subsample. We did not bootstrap because we were attempting to represent a known sample with a smaller number of individuals, not to generate a randomized sample to compare with a known sample. We did not simply reduce the number of each taxon proportionally and round to the nearest whole number because this approach would have eliminated rare taxa. Not all samples required subsampling, so we used this virtual subsampling method to mimic field collection methods used by the agencies as closely as possible. Sampling protocols used for rapid bioassessments based on fixed-counts do not involve collection of several samples and subsequent calculation of average num-

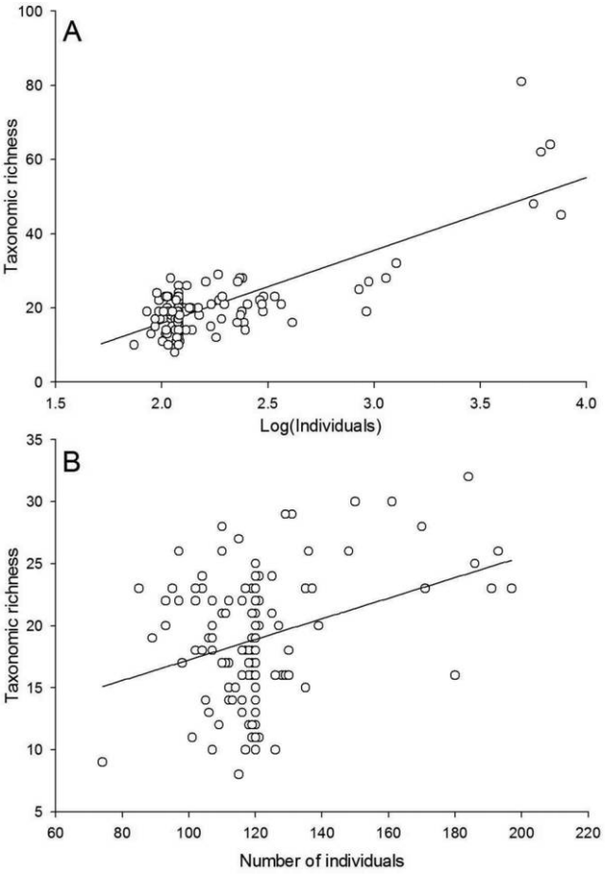

Figure 2. Taxonomic richness as a function of number of individuals identified in each benthic macroinvertebrate sample prior to subsampling (R^2 = 0.66, F = 292.3, n = 150, p < 0.0001; logged values plotted for clarity only) (A) and after subsampling samples containing >200 individuals to 120 individuals (R^2 = 0.10, F = 17.12, n = 150, p < 0.0001) (B).

bers of individuals in each taxon. Therefore, we generated a single subsample to make the subsampled data as comparable as possible to raw counts. We used a 200-count lower limit for subsampling because it is conceivable when doing a 100-count to pick 90 individuals from 1 grid square and 100 (or more) from the next, resulting in a 100-count sample with nearly 200 individuals.

DPAC-SL We used the SL approach to resolve ambiguous taxa in samples with >200 individuals before subsampling. For samples with multiple levels of ambiguity (e.g., order- and family-level ambiguous parents of genera and all chironomid taxa), we distributed the lowest ambiguous parent taxonomic level among its children first, and then used those relative abundances to distribute the higher ('grandparent') level of ambiguity; i.e., we distributed ambiguous families among genera, then ambiguous orders among all genera found in that sample. We formulated equation(s) in Excel (Office 2007; Microsoft, Redmond, California) to resolve all possible ambiguities in each sample containing >200 individuals. In each case, we distributed ambiguous parents among their children by multiplying the parental numerical abundance by its children's relative abundances and adding the total to the child genus' numerical abundance in that sample. However, we were unable to use DPAC-SL to resolve all ambiguous taxa because some samples contained ambiguous parents without children. We used DPAC-GL to resolve these ambiguous taxa after subsampling samples with >200 individuals and all ambiguous taxa in samples with <200 individuals.

DPAC-GL DPAC-SL distributes ambiguous parents among their children in a single sample, whereas DPAC-GL distributes ambiguous parents among their children based on the relative abundance of individuals in those child taxa across all samples at a single site. First, we subsampled all samples with >200 individuals to a 120-count (see S*ample size standardization*) to prevent samples with large numbers of individuals from overweighting relative abundances. Next, we summed total abundance of all taxa across all samples within a site (but not across sites). We treated this summed sample as a single sample, and we used the resulting abundances to calculate the relative abundance of each genus within its tribe/subfamily/family/order in the same manner as DPAC-SL. We then used these relative abundances to distribute ambiguous parents within samples. This method removed nearly all ambiguous parents from the data set.

SRBC biologists did not identify Chironomidae larvae from the Chemung River at Wellsburg, New York, to genus level. However, we wanted to include this site in our data set because it had a longer data record than other sites on the Chemung River. Chironomid larvae were identified to genus by NYDEC biologists for 2 sampling sites <8 km upstream and downstream of the SRBC sampling site in the same sampling years used by the SRBC. We calculated relative abundance of each chironomid genus for these NYDEC data sets and used them with the DPAC-GL method to distribute chironomids to genus at the SRBC sampling site.

DPAC-GC Eight ambiguous parent taxa with no identified child taxa within their respective sites remained after application of DPAC-SL and -GL. Other sites in our data set had identified child taxa for these ambiguous parent taxa. These 8 ambiguous taxa across all 10 sites accounted for only 41 of the 18,100 individuals in this data set after subsampling. We used the 'grouped conservative' (DPAC-GC) method (Cuffney et al. 2007) across all samples in the data set to remove these 8 ambiguous taxa. This method places all ambiguous parents in the most common child genus. These 8 taxa had no children at their respective sampling sites, so we used the total abundance of their children across all 150 samples after DPAC-SL and -GL and subsampling to distribute the last 41 ambiguous individuals among their children.

Partial individual distribution We generated partial individuals in some data sets when we multiplied the number of parent individuals by relative abundances of children. These partial individuals would unfairly weight diversity measures (e.g., several genera with 0.05 individuals in a single sample), so all partial organisms were rounded to whole individuals based on the decimal value of partial organisms in a parent taxa group. Individuals were distributed among several genera, so rounding traditionally (down if <0.5, up if ≥0.5) would have removed or added individuals to the sample in some instances. Instead, we rounded the highest decimals up and the rest down to keep the same count in each sample. For example, if 7 individuals from a parent taxon were distributed among 4 children taxa as 3.2, 2.2, 1.4, and 0.2 they would be rounded to 3, 2, 2, and 0 individuals, respectively. In Excel, use of a 'RIGHT()' command to remove whole integers followed by a series of conditional equations (e.g., 'IF()' and 'NOT()') with 'FLOOR()' and 'CEILING()' could be used for this rounding process. However, we typically did not have >3 or 4 values in a series, and manual rounding was more efficient than creating a new conditional equation for each group of taxa to round. This process prevented creation or removal of individuals from the data set via rounding, but subsampling and DPAC methods removed 27 genera and 30 ambiguous taxa from the data set. We acknowledge that losing taxa from the data set might affect assessment of overall biodiversity or the ability to track rare but important taxa, but these steps were essential to enable comparison of benthic macroinvertebrate assemblages in the river over a 20-y period based on data from multiple agencies with various sampling, sorting, and identification protocols.

Data analysis

We calculated Shannon–Wiener diversity (H′), Shannon evenness (E), Simpson's reciprocal index (1/D), mean richness per site, and the Hilsenhoff Biotic Index (HBI) as measures of diversity and potential changes in assemblage health and pollution tolerance (Barbour et al. 1999, Merritt et al. 2008). We accounted for potential bias caused by small individual samples and variable numbers of samples collected at each location by averaging diversity and pollution metrics for each sample across all samples within a site. We compared these diversity and tolerance metrics to general community composition as relative abundance of each insect order, Ephemeroptera, Plecoptera, Trichoptera (EPT) taxa, and functional feeding groups (FFGs) to investigate trends of potential importance in an IBI. We based habit and FFG classifications on information published by McCafferty and Bae (1992), Wiggins (1996), Barbour et al. (1999), and Merritt et al. (2008). We calculated % *Corbicula* spp. and % Simuliidae to evaluate whether we could use this data set to discern spatial or temporal trends in relative abundance related to invasion (*Corbicula* spp.) and pest management (Simuliidae).

We subsampled to avoid overweighting rare taxa and inflating diversity measures. Therefore, the resulting data were not quantitative, and we were very hesitant to compare raw counts as densities. Therefore, we converted data to relative abundance to account for variability among number of samples per site by summing all taxonomic information across time for each site and then calculating relative abundance for compositional metrics. This step was necessary because some sites were sampled 19 times, whereas others were sampled only 6 times.

We used nonmetric multidimensional scaling (NMDS) ordination to compare samples within and across subbasins (PC-ORD, version 5; MjM Software Design, Gleneden Beach, Oregon). We ran the NMDS analysis with the *Slow and Thorough* Autopilot NMDS option, which consists of 40 real and 50 randomized runs of data with 6 initial axes (McCune and Mefford 1999). We used Sørensen distances with random starting coordinates and 400 iterations, final stress of 18.3, instability < 0.0005, and $p = 0.0196$ (Monte Carlo, 50 runs). We did not down-weight or remove rare taxa because subsampling had already eliminated 27 rare taxa from the data set. We also used NMDS to compare assemblage composition metrics over time within sampling sites and across the entire basin to reveal temporal trends.

RESULTS

Subsampling and ambiguous taxa

Before subsampling the samples with >200 individuals, the data set contained 174 unambiguous taxa. Subsampling led to removal of 27 rare taxa only found in large samples, leaving 147 taxa in the final data set, with 41 (Fort Hunter) to 72 (Lewisburg) taxa found at any 1 site. The greatest proportion of total richness captured in a single 100-count sample was 56% (Fort Hunter) and the lowest proportion was 15% (Sunbury). Sample size and richness were linearly related ($R^2 = 0.66$). The relationship between richness and sample size was a consequence of the high abundances in USGS samples, and removal of these samples from the regression analysis yielded $R^2 = 0.15$. In contrast, distribution of all parent taxa and subsampling reduced the strength of the regression (R^2) between sample size and richness from 0.66 to 0.10 (Fig. 2A, B). Thus, our approach reduced variability in richness attributable to sample size and enabled us to retain all samples in the data set.

We compared site NMDS scores among agencies to assess whether assemblage composition was affected by the collecting agency. The 95% confidence intervals (CIs) for all agencies overlapped on Axis 1, and 95% CIs for all agencies except PADEP and SRBC overlapped on Axis 2 (axis scores from Fig. 3 were used in this analysis). However, the separation between PADEP and SRBC disappeared when sites sampled by only one agency (Wellsburg and Williamsport) were removed. Thus, the collecting agency did not influence taxonomic composition of samples.

Diversity and tolerance

Diversity was highest in North Branch sites and lowest in West Branch sites, particularly upstream at Williamsport and Bower, regardless of metric (Table 2). Richness ranged from 14 (Williamsport) to 22 (Marietta) taxa/sample, was positively related to H′ ($R^2 = 0.75$), E ($R^2 = 0.32$), and 1/D ($R^2 = 0.53$), was not related to the Hilsenhoff Biotic Index (HBI; $R^2 = 0.01$), and was negatively related to %EPT taxa ($R^2 = 0.18$). We expected sites with low diversity to have fewer pollution-sensitive individuals and, therefore, to have lower %EPT taxa and higher HBI scores. However, Williamsport had the highest %EPT taxa (driven in part by abundant Hydropsychidae) and Bower had the lowest HBI score (Table 2), even though these sites had the lowest richness. HBI scores were higher in the North Branch than in other sub-basins, but the North Branch had the highest values for diversity metrics. The Chemung River (Wellsburg) had the highest average HBI score and the highest average score for all 3 diversity measures.

Assemblage composition

Assemblages were dominated by insects in the orders Ephemeroptera and Trichoptera. Eight of the 10 most-abundant genera across sites were from these orders. Four genera (*Hydropsyche*, *Isonychia*, *Cheumatopsyche*, and *Maccaffertium*) were among the top 10 most-abundant taxa in all 3 sub-basins (Table 3). Taxonomic differences among sub-basins were evident at the order level. Coleoptera were nearly absent from the West Branch (1.3%), but were common in the North Branch (14%), and were abundant in the Lower Susquehanna (21%). Trichoptera were most abun-

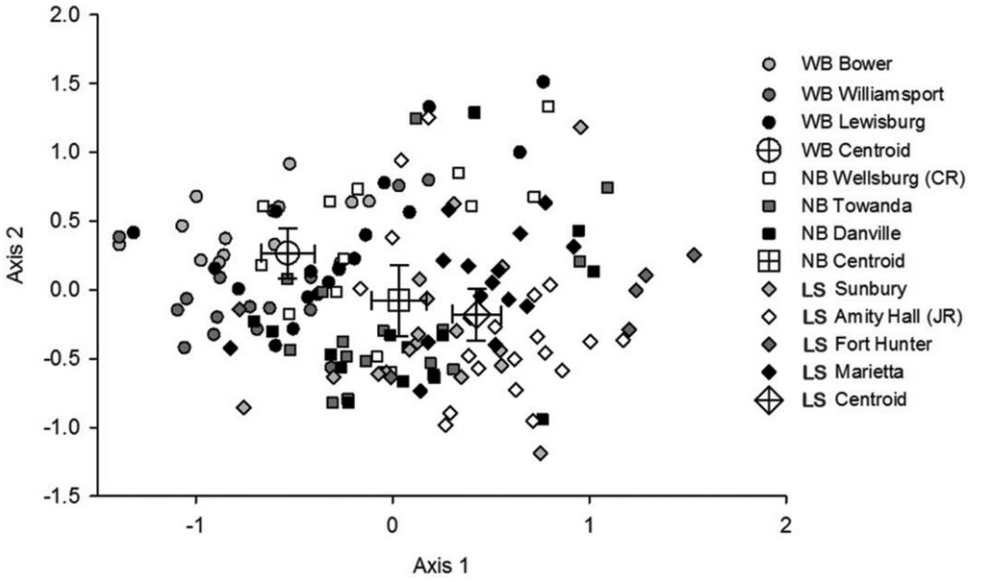

Figure 3. Nonmetric multidimensional scaling (NMDS) ordination of benthic macroinvertebrate assemblages from sites in the Susquehanna River. Error bars on centroids represent 95% confidence intervals for axis 1 and axis 2 scores. LS = Lower Susquehanna below the confluence of the North Branch (NB) and West Branch (WB).

dant in the West Branch (41%), whereas Ephemeroptera were most abundant in the Lower Susquehanna (44% at Fort Hunter and 43% at Marietta) and least abundant in the North Branch (29%). Diptera were most abundant in the North Branch (12%) and at Sunbury (14%), the uppermost site on the Lower Susquehanna.

The 95% CIs for the sub-basin centroids of site scores did not overlap on axis 1 of the NMDS ordination (Fig. 3). Sites were grouped by sub-basin along axis 1 (Fig. 3), primarily based on relative abundances of *Anthopotamus*, *Stenelmis*, and *Hydropsyche* (Table 4). *Anthopotamus* comprised 3.9% and 7.6% of individuals from the North Branch and Lower Susquehanna, respectively, but was not found in the West Branch. *Stenelmis* was rare in the West Branch but was the most common taxon in the North Branch and Lower Susquehanna. *Hydropsyche* was abundant in all sub-basins but was the most abundant taxon in the West Branch and was relatively less abundant in the North Branch and Lower Susquehanna (Table 3). Site distributions along axis 2 were related to relative abundances of less-common taxa (e.g., *Stenacron*) or taxa that were also strongly correlated with axis 1 and showed sub-basin specificity (e.g., *Stenelmis* and *Macrostemum*; Table 4). However, relative abundance of *Maccaffertium* (10.6% in

Table 2. Diversity and biotic index values for benthic macroinvertebrate communities at each site in the Susquehanna River based on government agency data from 1990–2010. Richness, Shannon diversity (H′), evenness (E), reciprocal Simpson diversity (1/D), and the Hilsenhoff Biotic Index (HBI) are means from all samples collected at a site. Percent Ephemeroptera, Plecoptera, Trichoptera (EPT) and %Diptera were calculated from data pooled across all samples from each site. WB = West Branch, NB = North Branch, LS = Lower Susquehanna, CR = Chemung River, JR = Juniata River.

Site	Richness	H′	E	1/D	HBI	%EPT	%Diptera
WB Bower	16	2.0	0.72	5.7	4.5	82	12
WB Williamsport	14	1.9	0.72	5.0	4.9	84	7
WB Lewisburg	18	2.2	0.76	6.6	5.0	73	7
NB Wellsburg (CR)	20	2.5	0.83	10.1	5.1	55	18
NB Towanda	21	2.5	0.82	9.2	4.9	62	6
NB Danville	21	2.4	0.80	8.8	5.1	56	13
LS Sunbury	19	2.2	0.76	7.4	5.0	53	14
LS Amity Hall (JR)	20	2.3	0.78	7.6	4.8	47	10
LS Fort Hunter	18	2.2	0.77	6.8	4.5	54	3
LS Marietta	22	2.5	0.81	9.1	4.8	59	6

Table 3. Mean (±1 SE) relative abundances of the 10 most-abundant taxa in Susquehanna River sites combined by sub-basin. Taxa shown in bold were among the top 10 taxa in all 3 sub-basins. West Branch: n = 3 sites, North Branch: n = 3 sites, Lower Susquehanna: n = 4 sites.

West Branch		North Branch		Lower Susquehanna	
Taxon	%	Taxon	%	Taxon	%
Hydropsyche	18.7 ± 4.4	*Stenelmis*	9.2 ± 1.9	*Stenelmis*	18.2 ± 2.4
Isonychia	14.4 ± 2.0	***Isonychia***	9.2 ± 1.9	***Isonychia***	8.0 ± 1.2
Cheumatopsyche	14.2 ± 3.3	***Cheumatopsyche***	7.6 ± 2.2	*Anthopotamus*	7.6 ± 2.6
Maccaffertium	10.6 ± 2.4	*Macrostemum*	6.4 ± 1.4	***Maccaffertium***	6.9 ± 2.9
Baetis	5.5 ± 0.2	*Chimarra*	5.0 ± 2.4	***Cheumatopsyche***	6.0 ± 2.1
Macrostemum	3.3 ± 2.2	***Hydropsyche***	4.4 ± 0.6	*Leucrocuta*	4.2 ± 1.4
Rheotanytarsus	2.0 ± 0.7	***Maccaffertium***	4.4 ± 0.3	*Corbicula*	3.9 ± 1.6
Simulium	1.9 ± 1.8	*Polypedilum*	4.3 ± 2.1	***Hydropsyche***	3.3 ± 0.6
Oligochaeta	1.7 ± 0.6	*Anthopotamus*	3.9 ± 0.8	*Baetis*	3.2 ± 0.6
Taeniopteryx	1.7 ± 1.2	*Sphaerium*	3.8 ± 1.9	*Chimarra*	3.0 ± 1.3

the West Branch, 6.9% in the Lower Susquehanna, and 4.4% in the North Branch) was strongly positively correlated with axis 2, and *Chimarra* (5.0% in the North Branch, 3.0% in the Lower Susquehanna, and 1.0% in the West Branch) was strongly negatively correlated with axis 2 (Table 4).

Across the watershed, relative abundances of filter-feeding taxa decreased with distance downstream, and relative abundances of scraper taxa increased (Fig. 4). This shift corresponded to declines in relative abundances of *Cheumatopsyche*, *Hydropsyche*, and *Isonychia* (filterer-collectors) with distance downstream and a sharp increase in *Stenelmis* (scrapers) in Lower Susquehanna sites (Table 3). Relative abundance of black flies (primarily *Simulium*) decreased markedly below Bower on the West Branch and below Danville on the North Branch (Fig. 5A) and was markedly lower at sites treated with *Bti* than at untreated sites (Fig. 5B).

The only taxon that showed a temporal trend was the invasive clam *Corbicula* spp. (Fig. 6). *Corbicula* spp. relative abundance was high shortly after initial colonization of a site and declined steadily for 10 to 15 y after colonization. The data indicated a possible increase in relative abundance at sites where it has been long established.

Table 4. Taxa correlated most strongly (Pearson product moment correlation, $r > 0.25$ or < -0.25, $p < 0.005$) with non-metric multidimensional scaling ordination axis scores. Each taxon made up >3% of benthic samples except *Polypedilum* (2.6%), *Sphaerium* (1.9%), *Gammarus* (1.5%), *Taeniopteryx* (0.9%), *Stenacron* (0.9%), and *Amnicola* (0.5%).

Axis 1		Axis 2	
Taxon	r	Taxon	r
Hydropsyche	−0.74	*Stenelmis*	−0.63
Anthopotamus	0.57	*Macrostemum*	−0.47
Stenelmis	0.47	*Stenacron*	0.45
Isonychia	−0.43	*Chimarra*	−0.36
Amnicola	0.36	*Maccaffertium*	0.36
Gammarus	0.36	*Polypedilum*	0.35
Baetis	−0.28	*Taeniopteryx*	0.35
Macrostemum	−0.26	*Sphaerium*	−0.35

DISCUSSION

Benefits of multiple-agency data

Federal and state agencies commonly collect biomonitoring samples to meet intra-agency goals, but when these data are combined among agencies, they cover a much broader scope. Thus, the combined data can be used to address new questions or treat questions regarding water quality and bioassessment more thoroughly. For example, in our data set, Williamsport was sampled only by PADEP and Wellsburg only by SRBC. Without combining data from these 2 agencies, we would have been unable to compare assemblages from these sites, which were in separate sub-basins. The multiple-agency data set increased the spatial extent of our study, and increased fine-scale spatial and temporal resolution by including samples collected from the same site(s) by multiple agencies over time. This increased resolution could improve the ability of the resulting data set to inform a regional IBI or to answer ecological questions. Our analysis showed that NMDS site scores did not differ among agencies. Therefore, we are confident in recommending this multiple-agency approach as a way to address gaps in our knowledge of large river ecology and assessment.

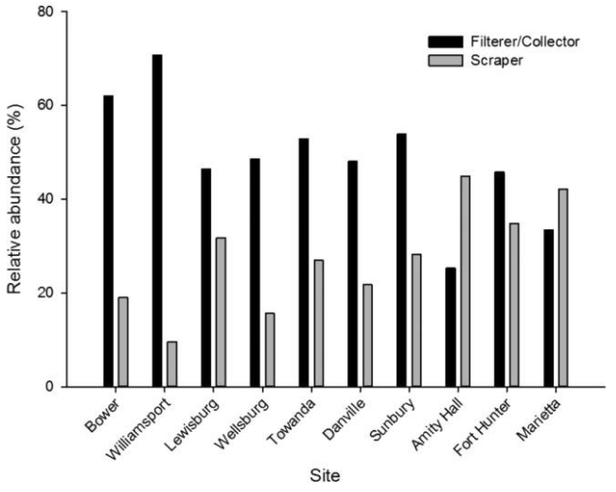

Figure 4. Relative abundance of filterer-collector and scraper macroinvertebrate functional feeding groups for each site in the Susquehanna River. Sites are grouped by sub-basin from upstream to downstream. Data were pooled across all sampling dates for each site prior to computing relative abundances.

Challenges of multiple-agency data

Resolving ambiguous taxa Generating this type of data set from multiple agencies over a 20-y period is challenging. Some subjective decisions were required to make data comparable across agencies and locations (e.g., grouping Annelida as Oligochaeta or Hirudinea, rather than using higher taxonomic resolution). The methods outlined here would have been difficult to implement without samples identified to a fine level of taxonomic resolution (at least genus-level identification). Thus, the USGS samples played a crucial role in resolving ambiguous taxa. Without samples for which invertebrates were identified to a fine level of taxonomic resolution, a more conservative approach would be best to resolve taxonomic ambiguities (Cuffney et al. 2007). However, grouping taxa at the family (or higher) level would result in a great loss of information, particularly for genus-rich families, such as Hydropsychidae and Heptageniidae.

The DPAC-GL method can dampen temporal trends at a single location, but it does not affect comparisons of data across sites. Moreover, the NMDS ordination did not reveal temporal trends in assemblage composition. Therefore, we are confident that the DPAC-GL method did not significantly alter temporal patterns of assemblage composition. One potential alternative when samples with high taxonomic resolution (e.g., USGS) are not available would be to use the 'Remove Parent Keep Child' (RPKC) method (Cuffney et al. 2007). However, we think this method is less desirable than the DPAC-GL/GC method because an inconsistent number of individuals are discarded, depending on the taxonomic resolution of each sample (Cuffney et al. 2007). Use of Operational Taxonomic Units (OTUs) is another option to prevent elimination of individuals from the data set. However, this method may artificially inflate taxonomic richness and diversity estimates (Cuffney et al. 2007). Moreover, in situations in which specific OTUs differed among agencies, data from different agencies could not be combined. Another potential option for resolving ambiguities would be to collect new samples and identify new specimens to inform the distribution of ambiguous parent taxa in historical samples. This approach would rely on consistent community structure and persis-

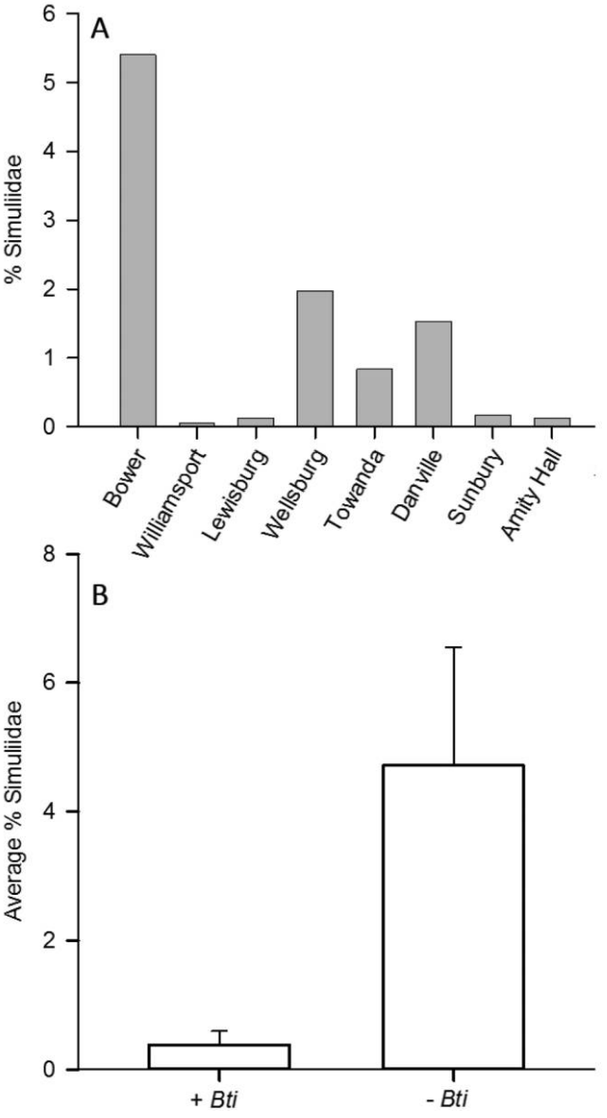

Figure 5. Relative abundance of Simuliidae at each site (A) and at sites grouped by presence (+) or absence (–) of *Bacillus thuringiensis israelensis* (*Bti*) application on the river near sampling locations (B). No Simuliidae were found at Fort Hunter or Marietta, the downstream-most sites, and these sites are omitted from the figure.

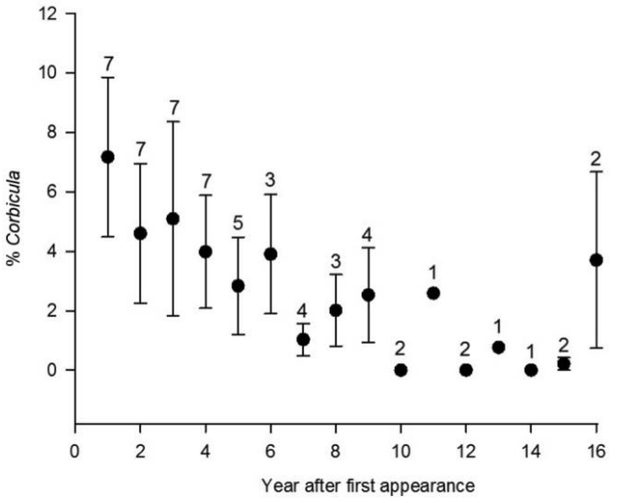

Figure 6. Mean (±1 SE) relative abundance of *Corbicula* spp. (Asian clams) in benthic samples from the Susquehanna River as a function of time since they first appeared at a site. Only locations where *Corbicula* spp. were found as of 2010 are included. Numbers above points are number of samples.

tent populations of all taxa over time, which might depend on stability of the particular taxon or aquatic ecosystem being studied.

Taxonomic changes Changes in taxonomy over time can be problematic even after samples have been standardized across agencies. For example, the presumed differences in identification of *Ceratopsyche* and *Hydropsyche* over time compelled us to combine individuals from these 2 genera. In addition, the reclassification of all but 1 species in the genus *Stenonema* to the genus *Maccaffertium* in 2004 (Wang and McCafferty 2004) occurred after publication of the EPA's Rapid Bioassessment Protocols for use in wadeable streams and rivers (Barbour et al. 1999). Therefore, we combined individuals of both genera into 1 genus (*Maccaffertium*) and used the FFG, HBI, and habit designation of *Stenonema* to categorize this abundant genus (7.3% of all individuals in this data set) based on the Rapid Bioassessment Protocol (Barbour et al. 1999).

Difficulty distinguishing between particular taxa poses taxonomic challenges for the practitioner, but conflicts among systematists are even more problematic for agency biologists to track. These examples show how important it is to record exactly which taxonomic keys are used in scientific studies and bioassessment programs.

Limitations resulting from sampling methods The small number of individuals identified for most samples in this data set (~120) posed challenges for answering ecologically relevant questions. Ecological diversity tends to scale with ecosystem size and resource availability (Blakely and Didham 2010), and a large productive and heterogeneous river system like the Susquehanna may require more intensive sampling than would be needed in smaller streams to reliably characterize the benthic assemblage. Sampling in a 500-m wide river is daunting, but a sampling protocol using 5 kicks to form a composite sample, from which a fixed count of 120 organisms is taken, seems inadequate and is expected to capture <½ of the local richness based on the average of rarefaction curves of samples by subbasin (Fig. 7).

Use of larger sample counts and quantitative sampling methods would improve ability to compare densities of organisms across a basin. Quantitative samples would allow agencies to assess whether absolute abundances of particular taxa have changed (rather than relative abundances, which can be affected by changes in other taxa), a clarification that could be very beneficial in the development of an IBI. However, increasing sample counts would increase the time burden on agency biologists. We suggest sampling fewer sites more intensively to compensate. Sampling of fewer locations by an individual agency could be offset if agencies collected from different sites and the data sets were combined. We also recommend a multihabitat sampling approach in which samples from several microhabitats are subsequently pooled and subsampled (sensu Moulton et al. 2000) with all samples collected from similar habitats (e.g., riffles). We think this greater initial cost in effort (and money) would be offset by the short- and long-term usefulness of the resulting data.

IBI implications

Our data set revealed important considerations when developing an IBI for the Susquehanna River and potentially for other large rivers. For example, a basin-wide IBI focused on dominant taxa would not be optimal, a point also made for the Mississippi River (Jackson et al. 2010).

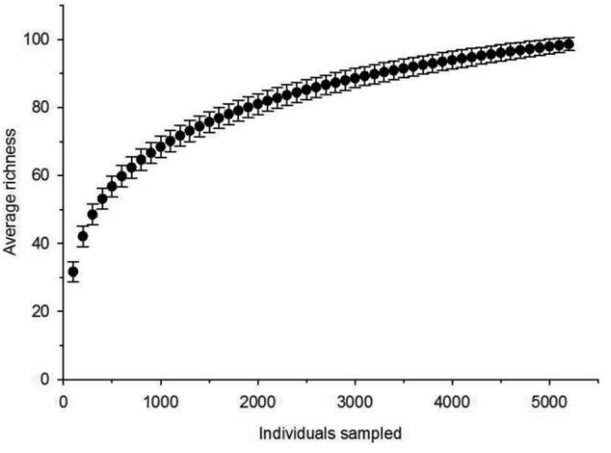

Figure 7. Mean (±1 SE) richness from rarefaction curves for individual sub-basins. Each point represents the expected richness from an additional 100 individuals added to the sample.

We found clear evidence for sub-basin specificity in dominant taxa. The predominant taxa in the West Branch are hydropsychid caddisflies, but riffle beetles are predominant in the North Branch and Lower Susquehanna. In addition, the North Branch has higher diversity and abundance of dipterans than the West Branch and Lower Susquehanna, and some genera (e.g., *Anthopotamus* and *Taeniopteryx*) show strong sub-basin specificity in their distributions. Thus, an IBI developed with data from all basins might misclassify sites because of differences in assemblage structure among sub-basins that are unrelated to impairment. We recommend taking a sub-basin or reach-specific approach similar to that used by Angradi et al. (2009) to develop condition indices for the Lower Missouri, Upper Mississippi, and Ohio River. Moreover, the high % EPT taxa (46–83%) abundance is composed mainly of relatively tolerant Hydropsychidae (average of 22%) and Heptageniidae (average of 10%). Thus, multimetric IBIs for large rivers might need to account for the dominance of pollution-tolerant mayflies and caddisflies. For example, Harrington et al. (1999) used %EPT minus Hydropsychidae.

We also found strong taxon-specific effects of a pesticide (*Bti*), which we present as an example of an anthropogenic effect on the assemblage that is unrelated to water quality but could confound an IBI. Simuliidae, particularly *Simulium*, were most abundant in the North and West Branches instead of in downstream Lower Susquehanna locations, where filter-feeding taxa might be expected to be more abundant based on River Continuum Concept predictions (Vannote et al. 1980). Black flies also were much less abundant at sites during years when PADEP personnel sprayed *Bti* to control their abundance (Fig. 5B). In addition, care should be taken when using relative abundances of functional feeding groups as IBI indicators. For example, *Corbicula* spp. and Simuliidae are both filter feeders, but *Corbicula* spp. has high relative abundance at downstream sites where Simuliidae is least abundant. Assessing sites based on relative abundance of filter feeders would obscure this pattern. In this particular case, the pattern might be important to track the progression of an invasive species. *Corbicula fluminea* was introduced to the Susquehanna below Conowingo Dam by 1980 (Mangan 2002). Other surveys suggest *C. fluminea* established populations throughout the entire Susquehanna River basin by 1995 (Foster et al. 2013), but samples from the most upriver sites in our data set did not contain *C. fluminea*.

Basin-wide trends

The taxonomic similarity between the North Branch and the Lower Susquehanna was expected because ~60% of the water feeding the Lower Susquehanna below their confluence comes from the North Branch (USGS 2012). In addition, the Lower Susquehanna sampling site at Sunbury was on the same (east) side of the river as the North Branch, ~28 km downstream of the Danville sampling location and <4 km downriver from the confluence of the 2 branches (Fig. 1). Dipterans were as abundant in Sunbury (14.0%) as the North Branch (12.1%), which seems logical given the proximity of this sampling site to sampling sites on the North Branch.

Macroinvertebrate assemblages in the Chemung River were more similar to assemblages in the West Branch than the North Branch even though the Chemung River is in the North Branch sub-basin and geographically distant from the West Branch. This result may indicate that river size or condition (e.g., high fine-sediment loading) may be a stronger driver of assemblage composition than geographic proximity. However, macroinvertebrate assemblages from the Juniata River clustered with nearby Lower Susquehanna sites in the ordination instead of with sites in the West Branch or Chemung River, which are more similar in size. The taxonomic similarity of the Juniata River to the Lower Susquehanna and between the West Branch and Chemung River might reflect the importance of tributary position in the river network, as well as geographic distance and stream size, in determining assemblage similarity between sampling locations.

We have shown the potential of long-term biomonitoring data from multiple agencies to broaden the scope and spatial extent of our understanding of macroinvertebrate assemblages within a region. Such data sets also have implications for future biomonitoring protocols and IBI development. The methods we have presented provide a pathway for similar analyses in other systems and across other agencies at a regional or continental scale.

ACKNOWLEDGEMENTS

We thank biologists from the Susquehanna River Basin Commission, NYDEC, PADEP, and USGS for providing data acquired through countless hours of sorting and identification of benthic macroinvertebrate samples. This study was supported by funding from the Chesapeake Bay Commission as part of the "State of the Susquehanna" report in 2010 with SRBC and the Henry Luce Foundation for funding of the Susquehanna River Initiative.

LITERATURE CITED

Angradi, T. R., M. S. Pearson, D. W. Bolgrien, T. M. Jicha, D. L. Taylor, and B. H. Hill. 2009. Multimetric macroinvertebrate indices for mid-continent US great rivers. Journal of the North American Benthological Society 28:785–804.

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. Office of Water, US Environmental Protection Agency, Washington, DC.

Blakely, T. J., and R. K. Didham. 2010. Disentangling the mechanistic drivers of ecosystem-size effects on species diversity. Journal of Animal Ecology 79:1204–1214.

Boyer, E. W., C. L. Goodale, N. A. Jaworski, and R. W. Howarth. 2002. Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern USA. Biogeochemistry 57:137–169.

Brigham, A. R., W. U. Brigham, and A. Gnilka. 1982. Aquatic insects and oligochaetes of North and South Carolina. Midwest Aquatic Enterprises, Mahomet, Illinois.

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.

Cuffney, T., M. Bilger, and A. Haigler. 2007. Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. Journal of the North American Benthological Society 26:286–307.

Foster, A. M., P. Fuller, A. Benson, S. Constant, D. Raikow, J. Larson, and A. Fusaro. 2013. *Corbicula fluminea*. USGS Nonindigenous Aquatic Species Database. US Geological Survey, Reston, Virginia. (Available from: http://nas.er.usgs.gov)

Hamilton, A. T., J. D. Stamp, and B. G. Bierwagen. 2010. Vulnerability of biological metrics and multimetric indices to effects of climate change. Journal of the North American Benthological Society 29:1379–1396.

Harrington, J. M., P. Ode, A. Montalvo, D. Post, C. Sheehy, and M. Dawson. 1999. Biological and physical/habitat assessment of California water bodies. California Department of Fish and Game, Rancho Cordova, California. (Available from: http://www.waterboards.ca.gov/water_issues/programs/tmdl/records/region_4/2003/ref1506.pdf)

Jackson, J. K., J. M. Battle, and B. W. Sweeney. 2010. Monitoring the health of large rivers with macroinvertebrates: do dominant taxa help or hinder the assessment? River Research and Applications 26:931–947.

Kennen, J. G., K. Riva-Murray, and K. M. Beaulieu. 2010. Determining hydrologic factors that influence stream macroinvertebrate assemblages in the northeastern US. Ecohydrology 3:88–106.

King, R., M. Baker, P. Kazyak, and D. Weller. 2011. How novel is too novel? Stream community thresholds at exceptionally low levels of catchment urbanization. Ecological Applications 21:1659–1678.

Mangan, B. P. 2002. Range expansion of the Asiatic clam, *Corbicula fluminea*, into the North Branch of the Susquehanna River. Journal of the Pennsylvania Academy of Science 76:40–42.

McCafferty, W., and Y. Bae. 1992. Filter-feeding habits of the larvae of *Anthopotamus* (Ephemeroptera: Potamanthidae). Annales De Limnologie–International Journal of Limnology 28: 27–34.

McCune, B., and M. J. Mefford. 1999. PC-ORD. Multivariate analysis of ecological data, Version 4. MjM Software Design, Gleneden Beach, Oregon.

Merritt, R. W., K. W. Cummins, and M. B. Berg (editors). 2008. An introduction to the aquatic insects of North America. 4th edition. Kendall/Hunt, 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. Open File Report 00-212. US Geological Survey, Reston, Virginia.

Moulton, S. R., J. G. Kennen, R. M. Goldstein, and J. A. Hambrook. 2002. Revised protocols for sampling algal, invertebrate, and fish communities as part of the National Water-Quality Assessment Program. Open File Report 02-150. US Geological Survey, Reston, Virginia.

PADEP (Pennsylvania Department of Environmental Protection). 2006. Standardized biological field collection and laboratory methods. Document Number 391-3200-015. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

PADEP (Pennsylvania Department of Environmental Protection). 2009. A benthic index of biotic integrity for wadeable freestone riffle-run streams in Pennsylvania. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania. (Available from: http://files.dep.state.pa.us/Water/Drinking%20Water%20and%20Facility%20Regulation/WaterQualityPortalFiles/Technical%20Documentation/freestoneIBImarch2012.pdf)

PADEP (Pennsylvania Department of Environmental Protection). 2012. 2012 Pennsylvania integrated water quality monitoring and assessment report. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania. (Available from: http://www.epa.gov/reg3wapd/pdf/pdf_nps/nps_annualreports/2012/PA2012AnnualReportFinal.pdf)

PADER (Pennsylvania Department of Environmental Resources). 1988. Standardized biological field collection methods. Pennsylvania Department of Environmental Resources, Harrisburg, Pennsylvania. (Available from: Pennsylvania Department of Environmental Resources Headquarters, Rachel Carson State Office Building, 400 Market Street, Harrisburg, Pennsylvania 17101 USA.)

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers: benthic macroinvertebrates and fish. EPA 440-4-89-001. US Environmental Protection Agency, Washington, DC.

Rempel, L. L., J. S. Richardson, and M. C. Healey. 2000. Macroinvertebrate community structure along gradients of hydraulic and sedimentary conditions in a large gravel-bed river. Freshwater Biology 45:57–73.

Ross, H. H. 1944. The caddis flies, or Trichoptera, of Illinois. Bulletin of the Illinois Natural History Survey 23:1–326.

Smith, A. J., D. L. Heitzman, and B. T. Duffy. 2009. Standard operating procedure: biological monitoring of surface waters in New York State. SOP 208-09:159. New York Department of Environmental Conservation, Albany, New York.

SRBC (Susquehanna River Basin Commission). 2006a. Basinwide maps and GIS data. Susquehanna River Basin Commission, Harrisburg, Pennsylvania. (Available from: http://www.srbc.net/atlas/bwmg.asp)

SRBC (Susquehanna River Basin Commission). 2006b. Susquehanna River Basin Commission information sheet. Susquehanna River Basin Commission, Harrisburg, Pennsylvania. (Available from: www.srbc.net/programs/docs/floodffwsgeneral(2_07).pdf)

Stamp, J. D., A. T. Hamilton, L. Zheng, and B. G. Bierwagen. 2010. Use of thermal preference metrics to examine state

biomonitoring data for climate change effects. Journal of the North American Benthological Society 29:1410–1423.

USEPA (US Environmental Protection Agency). 1972. Federal Water Pollution Control Act. 33 U.S.C. §1251. US Environmental Protection Agency, Washington, DC.

USEPA (US Environmental Protection Agency). 2002. Research strategy – Environmental Monitoring and Assessment Program. EPA 620/R-02/002. Office of Research and Development, US Environmental Protection Agency, Research Triangle Park, North Carolina.

USGS (US Geological Survey). 2012. USGS current water data for Pennsylvania. US Department of the Interior and US Geological Survey, Washington, DC. (Available from: http://waterdata.usgs.gov/pa/nwis/rt)

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.

Wang, T. Q., and W. P. McCafferty. 2004. Heptageniidae (Ephemeroptera) of the world. Part I: phylogenetic higher classification. Transactions of the American Entomological Society 130:11–45.

Wiggins, G. B. 1996. Larvae of the North American caddisfly genera (Trichoptera). University of Toronto Press, Toronto, Ontario.