Environ Monit Assess (2020) 192:531
https://doi.org/10.1007/s10661-020-08489-7

To composite or replicate: how sampling method and protocol differences alter collected stream invertebrates and associated bioassessment metrics

Lusha Tronstad · Oliver Wilmot · Darren Thornbrugh · Scott Hotaling

Received: 21 May 2019 / Accepted: 13 July 2020

Abstract Aquatic invertebrates are excellent indicators of ecosystem quality; however, choosing a sampling method can be difficult. Each method and associated protocol has advantages and disadvantages, and finding the approach that minimizes biases yet fulfills management objectives is crucial. To test the effects of both sampling methods and sample handling—i.e., to composite samples or leave them as replicates—we collected aquatic invertebrates from the Niobrara River at Agate Fossil Beds National Monument, Nebraska, using three methods and two sample handling protocols. We compared aquatic invertebrate assemblages collected with a Hester-Dendy multi-plate sampler, Hess sampler, and a D-frame dipnet. We calculated six common bioassessment metrics from composite (combined) and replicate (separate) samples. Hess samples contained the highest taxonomic richness (capturing 77% of all taxa observed) and dipnet samples the least (47%). Hester-Dendy samples had the greatest proportion of Ephemeroptera, and Ephemeroptera, Plecoptera, and Trichoptera (EPT). Dipnet samples had the lowest evenness values. In terms of sample handling, composite samples had inflated richness, diversity, and evenness compared with replicate samples, but bioassessment metrics calculated from proportions or averages (i.e., Hilsenhoff's Biotic Index and the proportion of EPT taxa) did not differ between them. The proportion of invertebrate groups from composite samples were not statistically different among sampling methods, but several groups differed between replicate samples collected by different methods. Ultimately, we recommend collecting replicate samples with a Hess sampler when the goal of the study is to detect ecosystem change, among locations or differences in variables of interest.

Keywords Aquatic invertebrates · Hess · Hester-Dendy · Dipnet · Method comparison · Stream monitoring · Biomonitoring · Stream ecology

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10661-020-08489-7) contains supplementary material, which is available to authorized users.

L. Tronstad (✉) · O. Wilmot
Wyoming Natural Diversity Database, University of Wyoming, 1000 E. University Ave, Laramie, WY 82071, USA
e-mail: tronstad@uwyo.edu

D. Thornbrugh
Northern Great Plains Inventory and Monitoring Program, National Park Service, Rapid City, SD, USA

S. Hotaling
School of Biological Sciences, Washington State University, Pullman, WA, USA

Present Address:
D. Thornbrugh
USDA Forest Service, Eastern Region, Milwaukee, WI 53202, USA

Introduction

Aquatic invertebrates have been used to monitor ecosystem quality for over 150 years (Cairns and Pratt 1993), largely because they have several characteristics

Published online: 18 July 2020

that make them ideal for the task. Aquatic invertebrates are relatively long lived (weeks to > 100 years, Rosenberg and Resh 1993a) and unlike water samples that are collected periodically, aquatic invertebrates spend most or all of their life cycle in water and therefore, their presence or absence reflects long-term conditions at a site. For instance, water samples may miss discrete, short-lived discharges of pollution, but aquatic invertebrate communities will respond to such an event (Rosenberg and Resh 1993b; Pallottini et al. 2017a). Furthermore, aquatic invertebrates are relatively sedentary and diverse and are relatively inexpensive to collect and identify. Most importantly, lower ecosystem quality in a stream can increase mortality and decrease reproduction, survival, and fitness of sensitive aquatic invertebrates (e.g., Ephemeroptera) while others are more tolerant of disturbances (e.g., Diptera; Johnson et al. 1993; Barbour et al. 1999). Changes in the diversity or assemblage structure of aquatic invertebrates can inform managers of stream ecosystem quality (Rosenberg and Resh 1993b) and altered trophic roles (Vannote et al. 1980).

Choosing a sampling method for aquatic invertebrate monitoring is difficult and depends on many variables. All approaches have advantages and disadvantages (e.g., cost to implement, time, bias toward specific taxa, or life histories; e.g., Macanowicz et al. 2013, Tronstad and Hotaling 2017). Therefore, identifying a method that is cost-effective, minimizes bias, and fulfills management objectives, is critical. Two types of indices are used to evaluate streams, multimetric (e.g., Kerans and Karr 1994; Mondy et al. 2012), and predictive models (e.g., Hawkins et al. 2000), but both use metrics developed from richness, abundance, tolerance, and more recently, functional traits (Ussegilo-Polatera et al. 2000; Pallottini et al. 2017b). Bioassessment studies use a variety of sampling methods, including kicknets, fixed-area samplers (e.g., Hess sampler), artificial substrates (e.g., Hester-Dendy samplers), and dipnets (Carter and Resh 2001). However, some sampling methods are not well suited to all stream habitats. For example, artificial substrates (e.g., Hester-Dendy plates) are ideal for large, deep rivers that are otherwise difficult to sample (De Pauw et al. 1986), and Hess samplers and dipnets are used in wadeable streams. However, artificial substrates rely on colonization and therefore do not represent natural assemblages or densities and can be biased toward certain insect orders (Letovsky et al. 2012). The type of information being collected also matters. For example, qualitative data may be sufficient if the study is estimating ecosystem health to meet federal standards, but more rigorous quantitative sampling is needed to assess change over time (e.g., Slavik et al. 2004). Qualitative samples only report proportional data, while fixed-area samplers provide quantitative information on the density and biomass for each taxon in the assemblage.

Laboratory protocols can alter the taxa identified and the bioassessment metrics calculated. Previous studies (e.g., Vinson and Hawkins 1996) investigated what type of subsampling method is best for bioassessment studies to minimize cost and produce reliable results. The two main types of subsampling—fixed area (e.g., 25% of the sample) and fixed count (e.g., 300 individuals; e.g., King and Richardson 2002)—have been compared for many data types (e.g., Vinson and Hawkins 1996). However, the question of how replicate samples should be handled, i.e., whether combined into composites or processed as replicates, remains largely unaddressed. Most bioassessment protocols (e.g., US Environmental Protection Agency) direct users to composite samples in the field. That is, individual samples are combined into one large sample which is assumed to homogenize variance (Carey and Keough 2002); however, we are not aware of any studies investigating that assumption. Alternatively, replicate samples can be kept and analyzed separately with the potential for added insight at relatively little additional cost. DiFranco (2014) recommends collecting three replicate samples in wetland habitats, but some methods straddle a grey area between replicate and composite samples by directing users to pool microhabitat samples (e.g., pools and riffles) so that variance among microhabitats is estimated (Lazorchak et al. 1998; Hering et al. 2004; Mondy et al. 2012).

The National Park Service (NPS) has been monitoring aquatic invertebrates in the Niobrara River at Agate Fossil Beds National Monument since 1989 using Hester-Dendy samplers. However, due to the inherent complications of collecting samples using artificial substrates and an inability to make direct comparisons to other streams, a change in the monitoring approach is under consideration. In this study, we used the opportunity to address an applied issue in stream biomonitoring and answer three questions: (1) How does the sampling method affect the invertebrate assemblage collected in the Niobrara River? (2) How do the corresponding bioassessment metrics compare among sampling methods? (3) To what degree do composite vs. replicate samples alter the assemblage and bioassessment metrics?

Materials and methods

Study area

The headwaters of the Niobrara River are located near Lusk, Wyoming, and the river flows eastward into Nebraska and eventually into the Missouri River near Niobrara, Nebraska (Fig. 1). The Niobrara River Basin covers 32,600 km^2 of which the majority is grassland in northern Nebraska (Galat et al. 2005). Over 95% of the land within the basin is used for agriculture. The Niobrara River flows through Agate Fossil Beds National Monument in western Nebraska about 23 km from the Wyoming border. Here, the Niobrara River is a low-order stream flowing through the grassland. Agate Fossil Beds National Monument includes ~ 10.9 km^2 in a valley bottom and ~ 18 km of river flows through the park (Fig. 1). The river's riparian vegetation is dominated by cattails (*Typha* sp.) and the invasive yellow flag iris (*Iris pseudacorus* L. 1753), and its substrate is predominantly fine particles (e.g., sand, silt, and clay). Currently, northern pike (*Esox lucius* Linnaeus 1758), white suckers (*Catostomus commersonii* Lacepede 1803), and green sunfish (*Lepomis cyanellus* Rafinesque 1819) inhabit the river within the park (Spurgeon et al. 2014); however, nine other fish species were collected at Agate Fossil Beds National Monument prior to 1990 (Spurgeon et al. 2014).

We sampled three long-term monitoring sites along the Niobrara River (Fig. 1; Tronstad and Hotaling 2017) in 2016. We deployed Hester-Dendy samplers in mid-July and returned to collect them as well as Hess and dipnet samples in mid-August (see below). The most upstream site (Agate Springs Ranch) is located near the western park boundary. Agate Springs Ranch has an overstory of plains cottonwood (*Populus deltoides* W. Bartram ex Marshall) and cattails are more abundant than iris. The central site, Agate Middle, lacks an overstory and has gravel substrate with abundant iris and cattails surrounding the river.

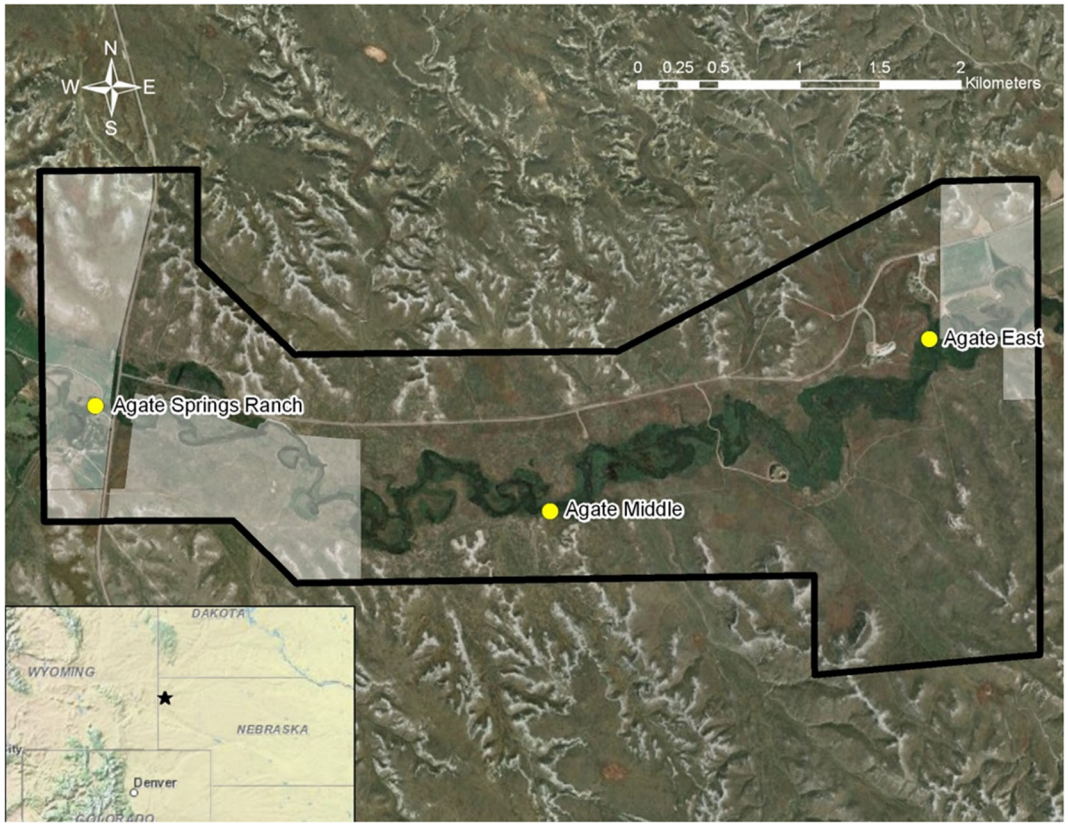

Fig. 1 We sampled three sites along the Niobrara River at Agate Fossil Beds National Monument in Nebraska, USA. The black line is the Monument boundary and the transparent white areas are private land within the Monument. The inset shows the location of Agate Fossil Beds National Monument in Nebraska (star)

Finally, Agate East is located before the Niobrara River flows out of the park and is the deepest site with riparian vegetation dominated by iris and a few willows (*Salix* spp.).

General measurements

To assess the general environmental characteristics of our study sites, we measured a number of standard variables (e.g., temperature), as well as water quality and clarity, sediment composition, water depth, and discharge. We measured dissolved oxygen (percent saturation and mg/L), pH, water temperature, specific conductivity, and oxidation-reduction potential using a Yellow Springs Instruments (YSI) Professional Plus. The YSI was calibrated on site before use. We measured water clarity by estimating the depth at which a Secchi disk disappeared from sight. The dominant substrate was recorded in the main channel of all sites and where each Hess sample was taken using soil texture tests (Thien 1979). Clay was defined as fine particles forming a ribbon after removing water, whereas silt did not form a ribbon. The sand was characterized by particles 0.06–2 mm in diameter, gravel was 2–64 mm in diameter, cobble was 64–256 mm in diameter, boulders were 25–400 cm in diameter, bedrock was > 4 m in diameter, and hardpan/shale was identified by the firm, consolidated fine substrate. We recorded the location of each site using a global positioning system (GPS; Garmin eTrex Vista HCx). Finally, we estimated stream discharge (Q; m^3/s) by measuring water depth (d; m) and velocity (v; m/s) using a Marsh-McBirney Flo-Mate 2000 at 0.3-m intervals across the stream's width (w; m) and summing each interval using Eq. (1):

$$Q = \sum d_i \times v_i \times w_i \quad (1)$$

Hester-Dendy sample collection

We deployed seven Hester-Dendy samplers (76 mm × 76 mm, 9 plates, Wildlife Supply Company) at each site. For each sampler, we strung a rope across the stream between two fixed posts with evenly spaced loops to separate the Hester-Dendy multi-plate samplers. The Hester-Dendy samplers were suspended in the water column at least 15 cm above the substrate. Debris dams were cleared weekly and we retrieved the samplers after 30 days of colonization by approaching the site from downstream, placing a dipnet (150-μm mesh) under it and cutting the rope. Hester-Dendy samplers were immediately placed in a container with ~ 80% ethanol and any organisms in the dipnet were removed and placed in the same container. In the laboratory, we dismantled and scrubbed the Hester-Dendy samplers to remove invertebrates that colonized the plates; then, we rinsed the samplers through a 212-μm sieve and preserved all specimens in ~ 80% ethanol. The middle five Hester-Dendy samples were used for analysis except when one of the samplers was compromised (e.g., touching the bottom).

Hess sample collection

We collected five Hess samples (500-μm mesh, 860-cm^2 sampling area, Wildlife Supply Company) within 50 m of the Hester-Dendy samples at each site. Samples were taken along the shallower margins of the stream where emergent vegetation is abundant. We placed the Hess sampler over vegetation to collect invertebrates living on it and in the surrounding benthic sediment. The vegetation and sediment were vigorously agitated and invertebrates were captured in the net. Samples were preserved in 80% ethanol and returned to the laboratory for analysis.

Dipnet sample collection

We collected dipnet samples along a 150-m stream reach using standard protocols in wadeable streams (US EPA 2013). We sampled invertebrates along 11 evenly spaced transects that were 15 m apart using a D-frame net (243-μm mesh, 30.5 × 25.4 cm opening, Wildlife Supply Company). At each transect, we sampled the right, left, and center of the stream systematically. Multiple habitats were sampled including benthic substrate, woody debris, macrophytes, and leaf packs. All samples were composited and preserved in the field with 95% ethanol.

Sample processing—Hester-Dendy and Hess

Invertebrates collected with Hester-Dendy and Hess samplers were sorted from debris in white trays and identified under a dissecting microscope. We rinsed samples through a large (2 mm) and small (212 μm for Hester-Dendy and 500 μm for Hess) sieve. All

invertebrates were sorted from the debris, counted, and identified in the samples; however, we subsampled the small fraction if individuals were numerous (> 1000 individuals) using the record player method (Waters 1969). Invertebrates were identified according to Merritt et al. (2008) for insects and Thorp and Covich (2010) and Smith (2001) for non-insect invertebrates. Invertebrate tolerance values were assigned to each taxon from Barbour et al. (1999).

Sample processing—Dipnet

We processed dipnet samples following the official EPA protocol (US EPA 2013). We elutriated all dipnet samples to remove inorganic substrate with a 500-μm mesh sieve. In the laboratory, we spread the sample evenly over a 30 × 36 cm sorting tray that was divided into 30 numbered grids (6 cm^2 each). Using a random number generator in R (R Development Core Team 2013), we selected six of the 30 grids, removed the invertebrates, and counted them. If the first six grids did not contain a minimum of 500 individuals, we randomly selected additional grids until the minimum threshold was reached. We removed and identified large or rare invertebrates defined as longer than 1.2 cm (Vinson and Hawkins 1996). All invertebrates were identified to the lowest taxonomic level possible, typically genus, and we normalized our abundance estimates for each site based upon the number of grids that were counted.

Statistical analyses

We used R (R Development Core Team 2013) and the packages *plyr* (Wickham 2011), *Matrix* (Bates and Maechler 2013), and *vegan* (Oksanen et al. 2013) to calculate invertebrate abundances, proportions, bioassessment metrics, and perform statistical tests. To estimate ecosystem quality, we calculated six common bioassessment metrics: Hilsenhoff's Biotic Index (HBI); Ephemeroptera, Plecoptera, and Trichoptera (EPT) richness; proportion of EPT taxa (number of EPT taxa divided by the total number of taxa collected); taxonomic diversity (Shannon's index); taxonomic richness and taxonomic evenness. Replicate Hester-Dendy samplers and Hess samples at each site were used to better calculate parameters (e.g., means) as is commonly done in stream invertebrate studies, but we recognize that these samples have limitations in statistical inference and must be interpreted carefully.

We compared invertebrate proportions and bioassessment metrics among sites and sampling methods with ANOVAs. If sites or methods were significantly different, we used Tukey's honest significant difference (HSD) to verify which sites or methods differed from one another with pair-wise comparisons. To compare invertebrate assemblages recovered with Hester-Dendy and Hess samples to dipnet samples, we electronically composited replicates at each site. However, to explore how compositing samples affect bioassessment metrics, we also calculated bioassessment metrics separately for each Hester-Dendy and Hess replicate at each site.

We evaluated differences in the aquatic invertebrate assemblage across sites and sampling method with nonmetric multidimensional scaling (NMDS) implemented in the R package *vegan* (Oksanen et al. 2013). NMDS provides an ordination-based approach to rank distances between objects and has been shown to perform well with non-normally distributed data (Legendre and Legendre 1998). To prepare our data for NMDS analysis, we removed rare taxa (as defined as any taxon that was unique to a single site + method combination). Next, we calculated the mean and standard deviation (SD) for each taxon and removed two species which were present at more than two deviations above the mean. Finally, we removed any taxon present at less than 0.1% of the overall abundance (after the first two filtering steps were completed). NMDS analyses were performed using Bray-Curtis distances on composite samples with default settings. To test whether the assemblages recovered were different depending on the sampling method or site, we performed an analysis of similarities (ANOSIM) with default settings (including 999 permutations). Next, we investigated differences in multivariate dispersion for each method by calculating the mean distance of each sample to the group's centroid in multivariate space with the function *betadisper*. We assessed pair-wise differences in dispersion with Tukey's HSD. To better visualize taxonomic differences in invertebrate assemblages collected with each sampling method, we constructed a ternary plot using the R package *ggtern* (Hamilton 2015). For ternary plot construction, we only removed rare taxa (as described above) before averaging the abundances of each taxon in composite samples across sites for each method.

Results

Environmental variation

Sites were environmentally similar to one another with little variation between our July and August sampling dates (Table 1). Water temperatures ranged from ~ 21–24 °C. Dissolved oxygen concentrations were near saturation. Specific conductivity was approximately 350 μS/cm and pH was consistently highest at Agate Springs Ranch. The oxidation-reduction potential was highest at Agate Springs Ranch (169–197 mV) and we measured reducing conditions (< 200 mV) at all sites. The discharge was higher in August and Agate East had the lowest flow. Agate East was the deepest site (1.2–1.5 m). Agate Springs Ranch was the narrowest (3–3.8 m) and shallowest (0.5–0.7 m; Table 1) site. The substrate at all sites was dominated by fine sediment (i.e., clay, sand, and silt) and gravel.

Hester-Dendy sampling

Across all methods and sites, Hester-Dendy samples contained 52% of the total invertebrate community we observed. Insecta and Crustacea (90% of individuals) were the most abundant taxa in Hester-Dendy samples. Of the insects, Diptera and Ephemeroptera were the most abundant followed by Trichoptera and Odonata (Supplementary Material 1). Hester-Dendy samples from Agate Middle (909 ind/sample) contained more invertebrates than both Agate Springs Ranch (217 ind/sample) and Agate East (279 ind/sample; $F = 7.1$, $df = 2$, $p = 0.009$; Tukey's HSD, $p < 0.025$; calculated with replicate samples). Taxonomic richness was lowest at Agate Springs Ranch (Table 2; $F = 28.7$, $df = 2$, $p < 0.001$). Taxonomic diversity ($F = 0.35$, $df = 2$, $p = 0.71$), taxonomic evenness ($F = 0.25$, $df = 2$, $p = 0.78$), EPT richness (Table 2; $F = 2.1$, $df = 2$, $p = 0.16$), and the proportion of EPT taxa did not differ among sites (Table 2; $F = 1.8$, $df = 2$, $p = 0.2$). The average tolerance value for an invertebrate collected with Hester-Dendy sampling was lowest at Agate Springs Ranch (HBI; Table 2; $F = 18.9$, $df = 2$, $p < 0.001$; Tukey's HSD, $p \leq 0.05$).

Hess sampling

We collected 77% of all observed taxa with Hess sampling. Overall, Insecta, Crustacea, and Annelida (98% of individuals) were the most numerous groups in Hess

Table 1 Water quality and site characteristics measured when Hester-Dendy samplers were deployed (July) and when Hester-Dendy, Hess, and dipnet samples were collected (August). A "B" after the Secchi disk depth indicated that the bottom of the stream was visible and the number is the maximum depth at the site. Stream width was measured with emergent vegetation excluded. Abbreviations and units include T_{WATER}, water temperature; T_{AIR}, air temperature; *DO*, dissolved oxygen; *SPC*, specific conductivity; and *ORP*, oxidation-reduction potential

Parameter/site	Ranch	Middle	East	Ranch	Middle	East
Date in 2016	18 July	18 July	19 July	19 Aug	17 Aug	17 Aug
Time	13:50	18:00	11:15	13:15	15:30	17:15
T_{WATER} (°C)	23.8	21.1	21.6	21.7	21.1	22.9
T_{AIR} (°C)	30	28	30	34	36	28
DO (% sat.)	NA	NA	NA	107.0	98.0	107.0
DO (mg/L)	NA	NA	NA	8.0	7.3	7.9
SPC (μS/cm)	357.2	352.4	364.9	347.2	354.4	358.6
pH	8.5	8.1	7.9	8.5	8.0	8.2
ORP (mV)	168.7	45.2	32.5	196.6	72.6	81.1
Secchi depth (cm)	47 (B)	82 (B)	67 (B)	58.5 (B)	73 (B)	149.0
Max. depth (m)	1.6	2.7	4.0	2.2	2.4	4.9
Width (m)	12.4	14.0	12.7	9.7	13.5	16.4
Discharge (m^3/s)	0.18	0.21	0.13	0.22	0.27	0.17
Dominant substrate	Sand	Gravel	Silt	Sand	Gravel	Silt/sand

Table 2 Invertebrate bioassessment metrics calculated from Hester-Dendy, Hess, and dipnet samples collected in the Niobrara River. Metrics for Hester-Dendy and Hess samples were calculated from replicate samples (i.e., mean metrics ± standard error) and composited samples (all replicate samples combined for each site and sampler). Dipnet samples were composited in the field and therefore no replicate samples are available for comparison

	Replicate			Composite		
Hester-Dendy	Ranch	Middle	East	Ranch	Middle	East
Richness	11 ± 0.75	17 ± 0.77	19 ± 0.80	14	24	29
Diversity	1.80 ± 0.13	1.90 ± 0.07	1.89 ± 0.11	3.37	3.37	3.41
Evenness	0.78 ± 0.04	0.69 ± 0.02	0.65 ± 0.03	1.28	1.06	1.01
EPT richness	5.4 ± 0.24	4.0 ± 0.55	4.8 ± 0.58	6	6	8
No. EPT/no. taxa	0.53 ± 0.01	0.25 ± 0.02	0.26 ± 0.03	0.43	0.25	0.28
HBI	3.9 ± 0.44	5.3 ± 0.17	6.4 ± 0.11	4.0	5.3	6.4
Hess	Ranch	Middle	East	Ranch	Middle	East
Richness	10 ± 1.86	24 ± 2.5	19 ± 1.8	19	41	34
Diversity	1.66 ± 0.41	2.00 ± 0.13	2.22 ± 0.14	2.24	3.64	3.80
Evenness	0.73 ± 0.14	0.65 ± 0.03	0.46 ± 0.03	0.76	0.98	1.08
EPT richness	2.4 ± 0.40	3.0 ± 0.32	1.6 ± 0.40	4	4	4
No. EPT/no. taxa	0.26 ± 0.03	0.14 ± 0.02	0.08 ± 0.01	0.21	0.10	0.12
HBI	5.4 ± 0.20	6.5 ± 0.45	6.8 ± 0.16	5.1	6.5	6.8
Dipnet	Ranch	Middle	East	Ranch	Middle	East
Richness	-	-	-	20	20	12
Diversity	-	-	-	2.31	1.79	0.69
Evenness	-	-	-	0.77	0.60	0.27
EPT richness	-	-	-	6	3	2
No. EPT/no. taxa	-	-	-	0.30	0.15	0.17
HBI	-	-	-	5.7	6.7	7.7

samples. Of the insects, Diptera were most abundant followed by Ephemeroptera, Odonata, and Trichoptera (Supplementary Material 2). Hess samples from Agate Middle (926 ind/sample) had higher abundances of invertebrates compared with both Agate East (465 ind/sample) and Agate Springs Ranch (282 ind/sample; $F = 8.7$, $df = 2$, $p = 0.005$; Tukey's HSD, $p \leq 0.035$; calculated from replicate samples). Taxonomic richness was lowest at Agate Springs Ranch (Table 2; $F = 11.7$, $df = 2$, $p = 0.001$; Tukey's HSD, $p < 0.02$), but taxonomic diversity did not differ among sites (Table 2; $F = 5.3$, $df = 2$, $p = 0.02$). Taxonomic evenness was highest at Agate Springs Ranch (Table 2; $F = 14.6$, df = 2, $p < 0.001$; Tukey's HSD, $p \leq 0.01$). Agate Springs Ranch also had a higher proportion of EPT taxa than both other sites (Table 2; $F = 3.8$, df = 2, $p = 0.05$). Additionally, invertebrates at Agate Springs Ranch had the lowest mean tolerance value (HBI; Table 2; $F = 24$, df = 2, $p < 0.0001$; Tukey's HSD, $p < 0.001$).

Dipnet sampling

Of all the invertebrate taxa observed in this study, 47% were found in dipnet samples. Overall, Insecta and Crustacea (99% of individuals) were the most numerous invertebrates. Within insects, Diptera were the most abundant order followed by Ephemeroptera, Odonata, and Coleoptera (Supplementary Material 3). We collected the most individuals from Agate Middle (~ 2685 ind/sample; composited) and fewer individuals from Agate East (~ 1260 ind/sample) and Agate Springs Ranch (~ 400 ind/sample). Taxonomic richness and diversity were lowest at Agate East (Table 2). Taxonomic evenness was highest at Agate East (Table 2). Agate Springs Ranch had the highest number of EPT as well as the highest EPT proportion (Table 2). As a result, invertebrates at Agate Springs Ranch had the lowest mean tolerance value (HBI). No statistical comparisons among sites are reported due to the lack of replicates for the dipnet sampling.

Community composition

We identified 73 invertebrate taxa representing six phyla (Annelida, Arthropoda, Mollusca, Nematoda, Nematomorpha, and Platyhelminthes) in the Niobrara River when all samplers were combined (Supplementary Material 1-3). Hester-Dendy samples contained nine taxa not found in Hess samples, 18 taxa not collected with the dipnet, and 8 taxa unique to Hester-Dendy samples. Hess samples contained 30 taxa not collected with Hester-Dendy samplers, 31 taxa not collected with the dipnet, and 21 taxa unique to Hess samples. Dipnet samples included 16 taxa not collected with Hester-Dendy samplers, 10 taxa not present in Hess samples, and 8 taxa unique to dipnet samples.

When composited, proportions of insects (Fig. 2a; $F = 0.3$, $df = 1$, $p = 0.75$) and non-insects (Fig. 2b; $F = 0.3$, $df = 1$, $p = 0.75$) did not differ among sampling methods. Proportions of Annelida, Crustacea, Coleoptera, Diptera, Ephemeroptera, Hemiptera, Mollusca, Odonata, and Trichoptera also did not differ when composited ($p \geq 0.25$; Fig. 2). Conversely, when treated as replicates (Hester-Dendy vs. Hess only), the proportion of insects (Fig. 2a; $F = 4.8$, $df = 1$, $p = 0.04$), non-insects (Fig. 2b; $F = 4.8$, $df = 1$, $p = 0.04$), Annelida (Fig. 2c; $F = 11.8$, $df = 1$, $p = 0.002$), Ephemeroptera (Fig. 2d; $F = 4.6$, $df = 1$, $p = 0.04$), Odonata (Fig. 2e; $F = 4.6$, $df = 1$, $p = 0.04$), and Trichoptera (Fig. 2f; $F = 6.9$, $df = 1$, $p = 0.01$) differed between Hester-Dendy and Hess samples. The proportion of Mollusca ($F = 3.7$, $df = 1$, $p = 0.065$), Crustacea ($F = 0.43$, $df = 1$, $p = 0.52$), Coleoptera ($F = 0.2$, $df = 1$, $p = 0.65$), Diptera ($F = 0.79$, $df = 1$, $p = 0.38$), and Hemiptera ($F = 2.5$, $df = 1$, $p = 0.13$) did not differ between replicate Hester-Dendy and Hess samples.

Additionally, NMDS analyses indicated that sampling methods collected different aquatic invertebrate assemblages (p, ANOSIM = 0.008; Fig. 3a), but assemblages did not differ among sites (p, ANOSIM = 0.408; Fig. 3b). While different sampling methods yielded distinct assemblages, the amount of multivariate space occupied by each method, a proxy for the taxonomic breadth of invertebrates recovered, did not differ (p, Tukey's HSD ≥ 0.94; Fig. 3a). Visualization of the assemblage recovered by each method via a ternary plot highlighted that Hess and Hester-Dendy sampling collected the majority of taxa (Fig. 4). After filtering rare taxa as described above, only one taxon (*Ceratopogon*, Ceratopogonidae) was observed in dipnet samples yet was largely absent elsewhere. Both Hess (13 taxa) and Hester-Dendy (7 taxa) sampling recovered taxa that were either rare or completely absent in the other methods. A few taxa were relatively equally represented across all three methods including *Anax*, Collembola, *Hyallela*, and Lymnaeidae (Fig. 4).

Bioassessment metrics

Bioassessment metrics differed among composited sampling methods, but most comparisons were not significant without incorporating replicates. Taxonomic richness (Fig. 5a; $F = 2.6$, $df = 2$, $p = 0.19$), diversity (Fig. 5b; $F = 4.4$, $df = 2$, $p = 0.10$), evenness (Fig. 5c; $F = 5.4$, $df = 2$, $p = 0.07$), and EPT richness (Fig. 5d; $F = 3.3$, $df = 2$, $p = 0.14$) did not differ among sampling methods. The proportion of EPT taxa (Fig. 5e; $F = 63$, $df = 2$, $p = 0.0009$) was highest in Hester-Dendy samples and lowest in Hess samples (Tukey's HSD, $p < 0.05$). HBI values (Fig. 5f; $F = 28$, $df = 2$, $p = 0.005$) were lowest in Hester-Dendy samples (Tukey's HSD, $p < 0.02$).

Most bioassessment metrics calculated from electronically composited samples were higher than those estimated from replicate samples. When composited, 40% more taxa were observed in Hester-Dendy and 80% more taxa were identified in Hess samples compared with replicate samples (Table 2). Similarly, EPT richness was 43% higher in composite Hester-Dendy and 83% higher in composite Hess samples compared with replicates. Taxonomic diversity was 82% higher in composite Hester-Dendy samples and 63% higher in composite Hess samples compared with replicate samples. Finally, composite Hester-Dendy (58%) and Hess samples (54%) had higher evenness values than replicate samples. Conversely, the proportion of EPT taxa and HBI values did not differ between composite and replicate samples.

Discussion

Both sampling method and processing (whether replicate or composite) influence the invertebrate assemblage collected and bioassessment metrics calculated. Hess samples yielded more unique taxa and the most complete picture of the stream invertebrate assemblage. Hester-Dendy samples were biased toward EPT taxa and dipnet sampling emphasized the most common taxa and thus had the lowest evenness values. Compositing samples yields elevated taxonomic richness, diversity,

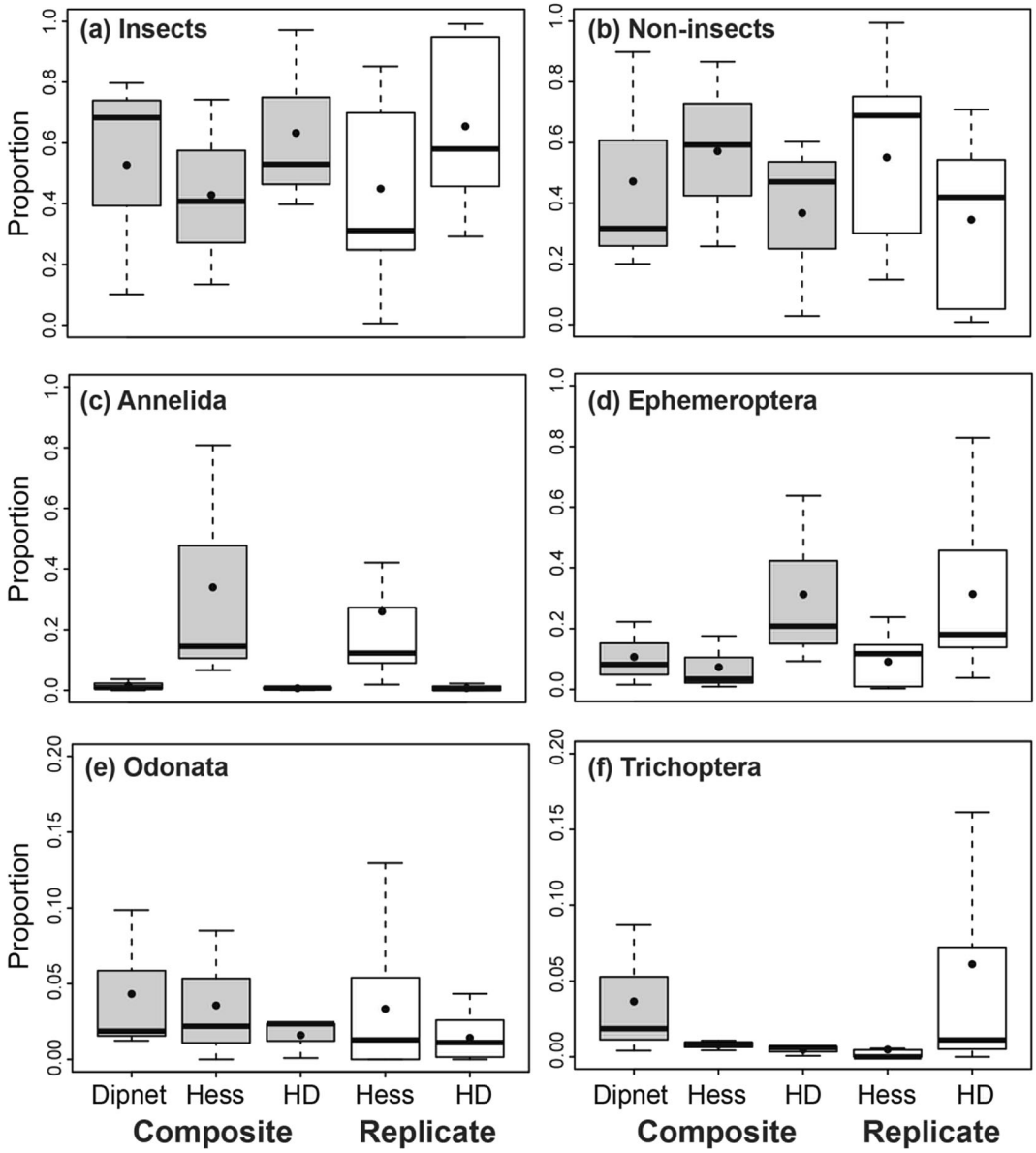

Fig. 2 Proportions of insects (**a**), non-insect invertebrates (**b**), Annelida (**c**), Ephemeroptera (**d**), Odonata (**e**), and Trichoptera (**f**) in dipnet, Hess, and Hester-Dendy (HD) samples that were composited (grey boxes) or kept separate as replicates (white boxes; HD and Hess only) collected from the Niobrara River, Nebraska, USA. Black circles are mean values, bold lines are median values, lower and upper limits are the 25th and 75th percentiles, and whiskers indicate the lower and upper limits of the data

and evenness compared with the same metrics calculated from individual replicates; however, metrics based on proportions or averaging (e.g., HBI) did not differ. Our results add another line of evidence that different sampling methods collect different portions of the invertebrate community and care must be taken when choosing an approach. For example, many studies compared the aquatic invertebrates captured using different samplers in a variety of habitats, such as streams, wetlands, vegetation, and sinkholes (e.g., Macanowics et al. 2013; Turner and Trexler 1997; Buss and Borges 2008); however, we are unaware of any studies comparing Hess, Hester-Dendy, and dipnet sampling directly. While managers should be aware of the potential bias of different methods, some approaches may be more useful than others under certain conditions. For

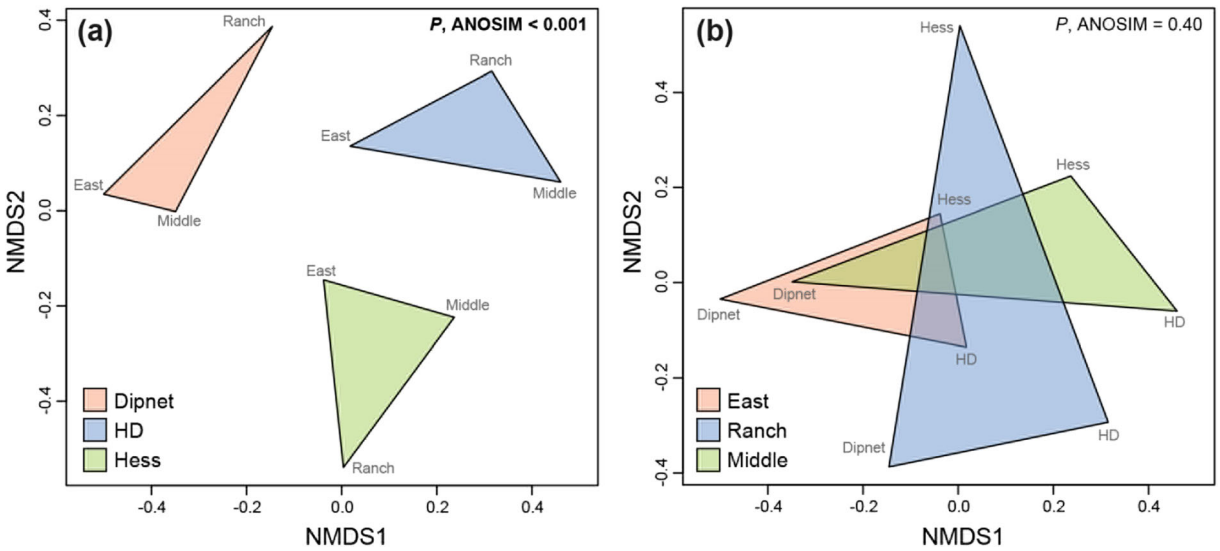

Fig. 3 Comparisons of invertebrate assemblages recovered by **a** sampling method and **b** site with non-metric multidimensional scaling (NMDS). Collected assemblages differed with the sampling method but not the site. HD, Hester-Dendy

example, funnel traps, dipnets, and stovepipe corers captured the most taxa in emergent vegetation of the Florida Everglades while Hester-Dendy sampling collected fewer taxa (Turner and Trexler 1997). Similar to the Niobrara River, quantitative Surber samplers (an analog of Hess sampling) collected 95–98% of taxa in two Australian rivers where qualitative kicknet samples only captured 63–66% of the community (Gillies et al. 2009).

Bioassessment metrics are influenced by sampling method (e.g., Bouchard et al. 2014), sorting technique (e.g., Nichols and Norris 2006), subsampling method (e.g., Nichols and Norris 2006; King and Richardson 2002), mesh size (e.g., Battle et al. 2007), and the taxonomic level specimens are identified to (e.g., King and Richardson 2002; Jones 2008). Despite the fact that compositing samples are common in stream bioassessment (e.g., US EPA 2013, RIVPACS), few studies have

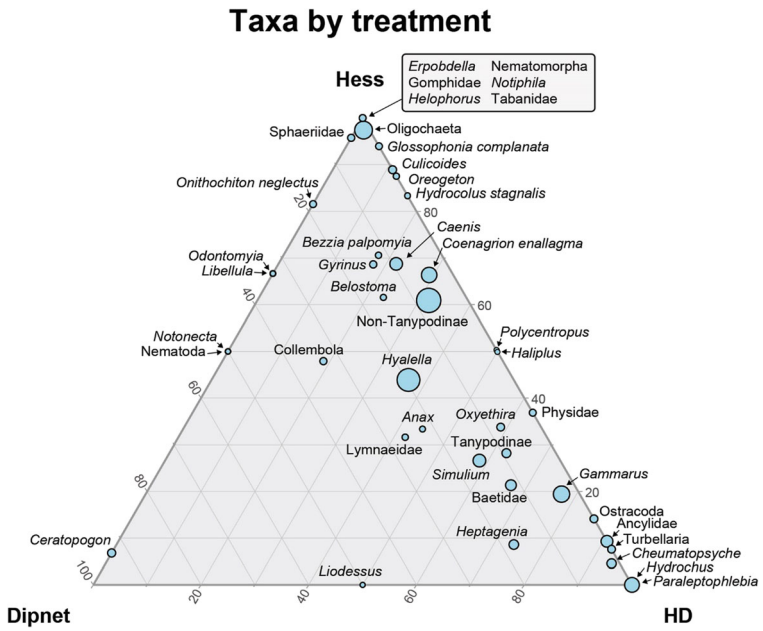

Fig. 4 Distribution of taxa recovered by Hess, Hester-Dendy, and dipnet sampling in the Niobrara River. The position of a given point indicates the percentage of the associated taxon with each sampling method. Circle size indicates the relative abundance of each taxon overall

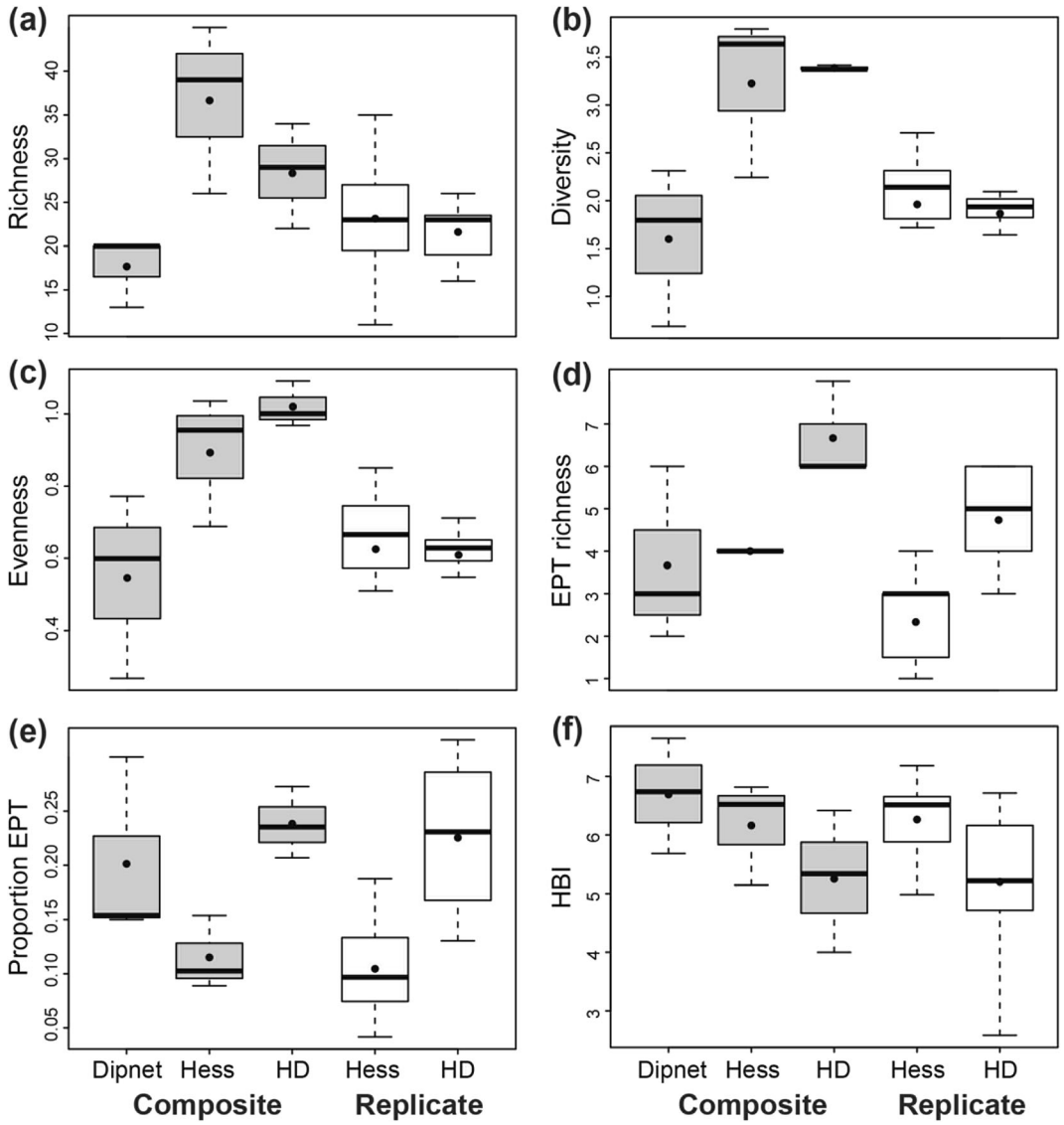

Fig. 5 **a** Richness; **b** diversity; **c** evenness; **d** Ephemeroptera; Plecoptera, and Trichoptera (EPT) richness; **e** proportion of EPT taxa; and **f** Hilsenhoff's biotic index (HBI) calculated from dipnet, Hester-Dendy (HD), and Hess samples for this study. Metrics calculated from composited samples are in grey and those calculated from five replicate samples are in the white boxes. For all metrics, except HBI, higher values indicate better ecosystem quality. Black circles represent mean values and bold lines are median values, lower and upper edges of the box are the 25th and 75th percentiles, and whiskers indicate the lower and upper limits of the data

investigated how compositing samples may alter metrics. We show that compositing alters bioassessment metrics (e.g., taxonomic richness, diversity, and evenness) and therefore, metrics calculated from composite samples should not be compared with those calculated from replicate samples. Indeed, only metrics calculated from proportions or averages should be compared between composite and replicate samples.

Composite samples are typically used as a cost-efficient method to assess conditions in aquatic ecosystems when estimating variance is not critical (Downes 2010). Most bioassessment protocols (e.g., RIVPACS and US EPA) recommend compositing samples to calculate a single estimate of metrics per site. Collecting a large composite sample is presumed to homogenize the variance and therefore produce a single, reliable value

(Carey and Keough 2002; B. Marshall, personal communication). One study discovered that metrics calculated using composite samples varied by 30% within a site (B. Marshall, personal communication). Vlek et al. (2006) compared the ecological quality class (a measure of stream ecosystem health) from bioassessment metrics calculated with replicate and composite samples and found that 8% were in different classes when five replicate samples were collected. In our study, composite samples from all methods produced a different result for each site using Hilsenhoff's Biotic Index (Hilsenhoff 1987). Composited Hester-Dendy samples had the highest ratings (fair to very good) and dipnet samples the lowest (poor to fair). Bradley and Ormerod (2002) reported that rare taxa were the largest source of error when sampling streams with kicknets. Another source of error likely lies in the subsampling of large composite samples which may introduce variance compared with replicate samples. Regardless of the subsampling method (i.e., fixed area or fixed counts), fewer individuals are removed and analyzed in composite samples versus replicate samples. Ultimately, more individuals analyzed will always yield more accurate estimates of conditions, but increasing the number of individuals also requires more resources. More studies designed to estimate differences between composite and replicate samples and their associated bioassessment metrics are needed to understand the consequences of sampling designs and when it is appropriate to use them.

Unlike composite samples, replicates enable managers to calculate variance which provides additional power to estimate differences among variables and/or sites of interest while simultaneously improving bioassessment accuracy (Quinn and Keough 2002). A key to effective use of replicate samples lies in identifying the variables for which knowledge of the variance is valuable and collecting replicates for them, while also identifying when to composite samples for other variables to save resources (Downes 2010). Replicate samples are recommended for monitoring data where statistical power is needed to detect changes over time (e.g., Slavik et al. 2004); however, studies should be carefully planned, analyzed, and interpreted when replicate samples are used (Hurlbert 1984; Davies and Gray 2015). Replicates are also necessary when the goal of a study is to detect differences among variables (e.g., sites, substrate) because replicates provide vital statistical power. For example, when replicates were composited in our study, we did not detect statistically significant differences in the proportion of invertebrate groups or the calculated metrics (e.g., taxonomic richness); however, when replicates for Hester-Dendy and Hess samples were compared, many groups yielded statistically different results. For best practices in stream biomonitoring, we recommend collecting replicate samples that are analyzed separately and electronically composited later if the need arises. While an argument could be made that collecting one composited sample in the field reduces the number of samples to manage in transit and process, in our experience, replicate samples are easier to process in the laboratory as they reduce the amount of material per sample, especially in areas with a lot of organic matter.

We also showed that different sampling methods yield very different perspectives on the aquatic invertebrate community being studied which will influence calculated bioassessment metrics. Previous studies reported that Hester-Dendy sampling tends to select for EPT taxa (Canton and Chadwick 1983; Letovsky et al. 2012). Because EPT richness is a common metric in biomonitoring, Hester-Dendy samples can bias bioassessment metrics toward lower values, indicating better ecosystem health. Our results support this as Hester-Dendy samples in the Niobrara River had the largest proportion of Ephemeroptera, the highest EPT, and the largest proportion of EPT taxa. As a result, HBI values were lowest for Hester-Dendy samples because Ephemeroptera tend to be sensitive taxa with low tolerance values. Beyond a single season, we have shown that Hess samples collected more taxa than Hester-Dendy samples across five consecutive years of sampling in the Niobrara River (Tronstad and Hotaling 2017). Dipnets performed consistently poorer than both Hester-Dendy and Hess samples in terms of the number of unique taxa recovered. Similarly, Hester-Dendy samples collected lower taxonomic diversity compared with kicknet samples (McCabe et al. 2012; Letovsky et al. 2012), sweep nets, and stovepipe cores (Turner and Trexler 1997) in other aquatic ecosystems. Quantitative samplers (e.g., Surber and Hess samplers) collected similar (Buss and Brges 2008) or more taxa than kicknets (Gillies et al. 2009) and box samplers (O'Connor et al. 2004). In the Niobrara River, Hess samples contained more than twice as many taxa as dipnets at two of the sites. Thus, our study lends additional support to previous findings that quantitative sampling (e.g., Hess or Surber) out performs other methods by collecting more taxa overall, more unique taxa, and by sampling natural features, a

more representative view of the natural community (Tronstad and Hotaling 2017).

Hester-Dendy and Hess samples suggested that invertebrates were fairly evenly distributed in the sampled assemblage based on taxonomic evenness. We calculated taxonomic evenness as Shannon's diversity index divided by the $\log_{10}$ of richness. A value near zero indicates that the assemblage is dominated by a few taxa, whereas a value near one indicates that the abundance of each taxon is similar. Mean richness for composited samples were close to one for both Hess and Hester-Dendy samples; however, dipnet samples had a mean value of 0.55, suggesting substantial bias in the assemblage toward high-density taxa (Table 1). Specifically, our dipnet samples had a high abundance of Amphipoda. Our results indicated that taxonomic evenness should only be compared with other dipnet samples and dipnets likely underestimate the evenness of the invertebrate community being studied.

We recommend sampling quantitatively (e.g., Hess, stovepipe core) for aquatic invertebrate biomonitoring studies when streams are wadeable. In our study, Hess samples collected the most taxa overall, yielded an intermediate HBI value that we expect most closely reflected the natural community because we sampled natural, benthic features in the stream. Either a Hess sampler or stovepipe core can be used in most habitats (e.g., streams, wetlands, and lakes) and substrate types (e.g., fines, gravel, and cobble) that are wadeable. For sample processing, we recommend collecting replicate samples in the field, especially when the variance is important for detecting changes (e.g., over time or differences among variables of interest). Generally, composite samples lack the statistical power to detect changes in variables of interest. Choosing the most appropriate sampling method paired with processing each replicate individually will provide the most valuable experimental design in most cases, particularly because replicates can always be electronically combined after the fact but the reciprocal is not true.

Acknowledgments We thank Katrina Cook, Linda Cooper, Isaac Dority, Heather Hicks, Ariele Johnson, Alexis Lester, Tresize Tronstad, and Sarah Wannemuehler for field and laboratory assistance. Robert Manasek and James Hill of the National Park Service provided logistical and field support, as well as the opportunity to work at Agate Fossil Beds National Monument. The project was supported by the National Park Service. Discussions with Brett Marshall were helpful in developing the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Barbour, M. T., Gerritsen, J., Snyder, B. D., & Stribling, J. B. (1999). *Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish*. Washington, D.C.: U.S. Environmental Protection Agency.

Bates, D., & Maechler, M. (2013). Matrix: sparse and dense matrix classes and methods. R package version 1.0-12.

Battle, J. M., Jackson, J. K., & Sweeney, B. W. (2007). Mesh size affects macroinvertebrate descriptions in large rivers: Examples from the Savannah and Mississippi Rivers. *Hydrobiologia, 592*, 329–343. https://doi.org/10.1007/s10750-007-0771-x.

Bouchard, R. W., Genet, J. A., & Chirhart, J. W. (2014). Does supplementing dipnet samples with activity traps improve the ability to assess the biological integrity of macroinvertebrate communities in depressional wetlands? *Wetlands, 34*(4), 699–711. https://doi.org/10.1007/s13157-014-0535-0.

Bradley, D. C., & Ormerod, S. J. (2002). Evaluating the precision of kick-sampling in upland streams for assessments of long-term change: The effects of sampling effort, habitat and rarity. *Archiv Fur Hydrobiologie, 155*(2), 199–221.

Buss, D. F., & Borges, E. L. (2008). Application of Rapid Bioassessment Protocols (RBP) for benthic macroinvertebrates in Brazil: Comparison between sampling techniques and mesh sizes. *Neotropical Entomology, 37*(3), 288–295. https://doi.org/10.1590/s1519-566x2008000300007.

Cairns, J., & Pratt, J. R. (1993). A history of biological monitoring using benthic macroinvertebrates. In D. M. Rosenberg & V. H. Resh (Eds.), *Freshwater biomonitoring and bBenthic macroinvertebrates* (pp. 10–27). New York, NY: Chapman and Hall.

Canton, S. P., & Chadwick, J. W. (1983). Aquatic insect communities of natural and artificial substrates in a montane stream. *Journal of Freshwater Ecology, 2*(2), 153–158.

Carey, J., & Keough, M. (2002). The variability of estimates of variance, and its effect on power analysis in monitoring design. *Environmental Monitoring and Assessment, 74*(3), 225–241.

Carter, J. L., & Resh, V. H. (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*(4), 658–682.

Davies, G. M., & Gray, A. (2015). Don't let spurious accusations of pseudoreplication limit our ability to learn from natural experiments (and other messy kinds of ecological monitoring). *Ecology and Evolution, 5*, 5295–5304.

De Pauw, N., Roels, D., & Fontoura, A. P. (1986). Use of artificial substrates for standardized sampling of macroinvertebrates in

the assessment of water-quality by the Belgian Biotic Index. *Hydrobiologia, 133*(3), 237–258.

DiFranco, J. L. (2014). Protocols for sampling aquatic macroinvertebrates in freshwater wetlands. Maine Department of Environmental Protection, Portland, Maine, DEPLW0640A-2014.

Downes, B. J. (2010). Back to the future: Little-used tools and principles of scientific inference can help disentangle effects of multiple stressors on freshwater ecosystems. *Freshwater Biology, 55*(Supplement 1), 60–79.

Galat, D. L., Berry, C. R., Peters, E. J., & White, R. G. (2005). Missouri River Basin. In A. C. Benke & C. E. Cushing (Eds.), *Rivers of North America* (pp. 427–480). New York, NY: Elsevier.

Gillies, C. L., Hose, G. C., & Turak, E. (2009). What do qualitative rapid assessment collections of macroinvertebrates represent? A comparison with extensive quantitative sampling. *Environmental Monitoring and Assessment, 149*, 99–112.

Hamilton, N. (2015). ggtern: An extension to ggplot2, for the creation of ternary Diagrams. (R package version, 1 ed.).

Hawkins, C. P., Norris, R. H., Hogue, J. N., & Feminella, J. W. (2000). Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications, 10*, 1456–1477.

Hering, D., Moog, O., Sandin, L., & Verdonschot, P. F. M. (2004). Overview and application of the AQEM assessment system. *Hydrobiologia, 516*, 1–20.

Hilsenhoff, W. L. (1987). An improved biotic index of organic stream pollution. *Great Lakes Entomologist, 20*, 31–39.

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs, 54*, 187–211.

Johnson, R. K., Wiederholm, T., & Rosenberg, D. M. (1993). Freshwater biomonitoring using individual organisms, populations, and species assemblages of benthic macroinvertebrates. In D. M. Rosenberg & V. H. Resh (Eds.), *Freshwater biomonitoring and benthic macroinvertebrates* (pp. 40–158). New York: Chapman and Hall.

Jones, F. C. (2008). Taxonomic sufficiency: The influence of taxonomic resolution on freshwater bioassessments using benthic macroinvertebrates. *Environmental Reviews, 16*, 45–69. https://doi.org/10.1139/a07-010.

Kerans, B. L., & Karr, J. R. (1994). A bethic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications, 4*, 768–785.

King, R. S., & Richardson, C. J. (2002). Evaluating subsampling approaches and macro invertebrate taxonomic resolution for wetland bioassessment. *Journal of the North American Benthological Society, 21*(1), 150–171. https://doi.org/10.2307/1468306.

Lazorchak, J. M., Klemm, D. J., & Peck, D. V. (1998). Environmental monitoring and assessment program-surface waters: Field operations and methods for measuring the ecological condition of wadeable streams. US Environmental Protection Agency Report EPA/620/R-94/004F.

Legendre, P., & Legendre, L. (1998). *Numerical Ecology*. Amsterdam: Elsevier Science.

Letovsky, E., Myers, I. E., Canepa, A., & McCabe, D. J. (2012). Differences between kick sampling techniques and short-term Hester-Dendy sampling for stream macroinvertebrates. *Bios, 83*(2), 47–55.

Macanowics, N., Boeing, W. J., & Gould, W. R. (2013). Evaluation of methods to assess benthic biodiversity of desert sinkholes. *Freshwater Science, 32*(4), 1101–1110.

McCabe, D. J., Hayes-Pontius, E. M., Canepa, A., Berry, K. S., & Levine, B. C. (2012). Measuring standardized effect size improves interpretation of biomonitoring studies and facilitates meta-analysis. *Freshwater Science, 31*(3), 800–812.

Merritt, R. W., Cummins, K. W., & Berg, M. B. (Eds.). (2008). *An Introduction to the Aquatic Insects of North America* (4th ed.). Dubuque, IA: Kendall Hunt Publishing.

Mondy, C. P., Villeneuve, B., Archaimbault, V., & Usseglio-Polatera, P. (2012). A new macroinvertebrate-based multimetric index (I2M2) to evaluate ecological quality of French wadeable streams fulfilling the WFD demands: A taxonomical and trait approach. *Ecological Indicators, 18*, 452–467.

Nichols, S. J., & Norris, R. H. (2006). River condition assessment may depend on the sub-sampling method: field live-sort versus laboratory sub-sampling of invertebrates for bioassessment. *Hydrobiologia, 572*, 195–213. https://doi.org/10.1007/s10750-006-0253-6.

O'Connor, A. O., Bradish, S., Reed, T. E., Moran, J., Regan, E. C., Visser, M., et al. (2004). A comparison of the efficacy of pond-net and box sampling methods in Turloughs – Irish Ephemeral Aquatic Systems. *Hydrobiologia, 524*(1), 133–144.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., et al. (2013). Vegan: Community Ecologye Package.

Pallottini, M., Cappelletti, D., Fabrizi, E., Gaino, E., Goretti, E., Selvaggi, R., & Céréghino, R. (2017a). Macroinvertebrate functional trait responses to chemical pollution in agricultural-industrial landscapes. *River Research and Applications, 33*, 505–513.

Pallottini, M., Goretti, E., Selvaggi, R., Cappelletti, D., Dedieu, N., & Cereghino, R. (2017b). An efficient semi-quantitative macroinvertebrate multimetric index for the assessment of water and sediment contamination in streams. *Inland Waters, 7*, 314–322.

Quinn, G., & Keough, M. (2002). *Experimental design and data analysis for biologist*. Cambridge: Cambridge University Press.

R Core DevelopmentTeam. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rosenberg, D. M., & Resh, V. H. (1993a). Introduction to freshwater biomonitoring and benthic macroinvertebrates. In D. M. Rosenberg & V. H. Resh (Eds.), *Freshwater biomonitoring and benthic macroinvertebrates* (pp. 1–9). New York: Chapman and Hall.

Rosenberg, D. M., & Resh, V. H. (Eds.). (1993b). *Freshwater biomonitoring and benthic macroinvertebrates*. New York: Chapman and Hall.

Slavik, K., Peterson, B. J., Deegan, L. A., Bowden, W. B., Hershey, A. E., & Hobbie, J. E. (2004). Long-term responses of the Kuparuk River Ecosystem to phosphorus fertilization. *Ecology, 85*(*4*), 939–954.

Smith, D. G. (2001). *Pennak's freshwater invertebrates of the United States* (4th ed.). New York: John Wiley and Sons, Inc..

Spurgeon, J. J., Stasiak, R. H., Cunningham, G. R., Pope, K. L., & Pegg, M. A. (2014). Status of native fishes withing selected protected areas of the Niobrara River in western Nebraska. *Great Plains Research, 24*, 71–78.

Thien, S. (1979). A flow diagram for teaching texture by feel analysis. *Journal of Agronomic Education, 8*, 54–55.

Thorp, J. H., & Covich, A. P. (Eds.). (2010). *Ecology and Classification of North American Freshwater Invertebrates* (3rd ed.). New York: Elsevier.

Tronstad, L. M., & Hotaling, S. (2017). Long-term trends in aquatic ecosystem bioassessment metrics are not influences by sampling method: Empirical evidence from the Niobrara River. *Knowledge and Managment of Aquatic Ecosystems, 418*(28). https://doi.org/10.1051/kmae/2017020.

Turner, A. M., & Trexler, J. C. (1997). Sampling aquatic invertebrates from marshes: Evaluating the options. *Journal of the North American Benthological Society, 16*(3), 694–709. https://doi.org/10.2307/1468154.

US Environmental Protection Agency. (2013). National rivers and streams assessment 2013-2014: fIeld operations manual-wadeable. (pp. 177). Washington DC: United States Environmental Protection Agency, Office of Water.

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

Vinson, M., & Hawkins, C. P. (1996). Effects of sampling area and subsampling procedure on comparisons of taxa richness among streams. *Journal of the North American Benthological Society, 15*(3), 392–399.

Vlek, H. E., Sporka, F., & Krno, I. (2006). Influence of macroinvertebrate sample size on bioassessment of streams. *Hydrobiologia, 566*, 523–542.

Waters, T. F. (1969). Subsampler for dividing large samples of stream invertebrate drift. *Limnology and Oceanography, 14*(5), 813–815.

Wickham, H. (2011). The Split-Apply_Combine Strategy for Data Analysis. *Journal of Statistical Software, 40*, 1–29.