LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 2024

doi: 10.1002/lom3.10601

Combined benthic and stream edge sampling better represent macroinvertebrate assemblages than benthic sampling alone along an aridity gradient

Brian A. Gill,[1]* Daniel C. Allen,[2] Meryl C. Mims,[3] Thomas M. Neeson,[4] Albert Ruhi,[5] Carla L. Atkinson,[6] Arial J. Shogren,[6] Travis M. Apgar,[5] Zacchaeus G. Compson,[7] Stephen Cook,[8] Daryl R. Trumbo,[3] Michelle H. Busch,[8] Kelsey D. Hollien,[1] Kyle Leathers,[5] Megan C. Malish,[4] Grace L. O'Malley,[3] Samuel Silknetter,[3] Chelsea R. Smith,[6] Howard Dunleavy,[3] Michael T. Bogan[1]

[1]School of Natural Resources and the Environment, The University of Arizona, Tucson, Arizona, USA
[2]Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, Pennsylvania, USA
[3]Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA
[4]Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, Oklahoma, USA
[5]Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, California, USA
[6]Department of Biological Sciences, University of Alabama, Tuscaloosa, Alabama, USA
[7]Department of Biological Sciences, Advanced Environmental Research Institute, University of North Texas, Denton, Texas, USA
[8]Department of Biology, University of Oklahoma, Norman, Oklahoma, USA

Abstract

Studies of stream macroinvertebrates traditionally use sampling methods that target benthic habitats. These methods could underestimate biodiversity if important assemblage components exist outside of the benthic zone. To test the efficacy of different sampling methods, we collected paired reach-wide benthic and edge samples from up to 10 study reaches in nine basins spanning an aridity gradient across the United States. Edge sampling targeted riparian-adjacent microhabitats not typically sampled, including submerged vegetation, roots, and overhanging banks. We compared observed richness, asymptotic richness, and assemblage dissimilarity between benthic samples alone and different combinations of benthic and edge samples to determine the magnitude of increased diversity and assemblage dissimilarity values with the addition of edge sampling. We also examined how differences in richness and assemblage composition varied across an aridity gradient. The addition of edge sampling significantly increased observed richness (median increase = 29%) and asymptotic richness (median increase = 173%). Similarly, median Bray–Curtis dissimilarity values increased by as much as 0.178 when benthic and edge samples were combined. Differences in richness metrics were generally higher in arid basins, but assemblage dissimilarity either increased or decreased across the aridity gradient depending on how benthic and edge samples were

*Correspondence: briangillphd@gmail.com

Author Contribution Statement: B.A.G.: Conceptualization, methodology, validation, formal analysis, data acquisition, data curation, writing—original draft, visualization. D.C.A.: Data acquisition, writing—review/editing, supervision, project administration, funding acquisition. M.C.M: Data acquisition, writing—review/editing, supervision, project administration, funding acquisition. T.M.N.: Data acquisition, writing—review/editing, supervision, project administration, funding acquisition. A.R.: Data acquisition, writing—review/editing, supervision, project administration, funding acquisition. C.L.A.: Data acquisition, writing—review/editing, supervision, project administration. A.J.S.: Data acquisition, writing—review/editing, supervision, project administration. T.M.A.: Data acquisition, writing—review/editing. Z.G.C.: Data acquisition, writing—review/editing. S.C.: Data acquisition, writing—review/editing. D.R.T.: Data acquisition, writing—review/editing. M.H.B.: Data acquisition, writing—review/editing. K.D.H.: Data acquisition, writing—review/editing. K.L.: Data acquisition, writing—review/editing. M.C.M.: Data acquisition, writing—review/editing. G.L.O.: Data acquisition, writing—review/editing. S.S.: Data acquisition, writing—review/editing. C.R.S.: Data acquisition, writing—review/editing. H.D.: Data acquisition, writing—review/editing. M.T.B.: Conceptualization, methodology, validation, formal analysis, data acquisition, data curation, writing—original draft, visualization, supervision, project administration, funding acquisition.

Additional Supporting Information may be found in the online version of this article.

combined. Our results suggest that studies that do not sample stream edges may significantly underestimate reach diversity and misrepresent assemblage compositions, with effects that can vary across climates. We urge researchers to carefully consider sampling methods in field studies spanning climatic zones and the comparability of existing data sets when conducting data synthesis studies.

Studies of stream macroinvertebrates require thoughtful matching of research questions and sampling methods (Jackson et al. 2019; Callisto et al. 2021). For fundamental questions, such as how macroinvertebrate assemblages change across environmental gradients, collection methods should strive to sample entire assemblages to enable the detection of drivers of composition. Alternatively, for biomonitoring applications, collection methods should target a subset of common taxa with limited mobility, long lifespans, and known responses to stressors that will signal environmental impairment (Rosenberg and Resh 1993; Rehn et al. 2007). Despite differences in targets and goals, macroinvertebrate studies commonly rely on established biomonitoring protocols for collecting primary data or conducting data syntheses. This preference for standardized methods over customizing approaches to address specific research questions can introduce biases, potentially distorting results and leading to erroneous inferences.

Established methods for sampling stream macroinvertebrates involve collecting benthic material across a study reach, but diverge in the extent to which different microhabitats are sampled (Rosenberg and Resh 1993; Jackson et al. 2019; Ligeiro et al. 2020; Callisto et al. 2021). In traditional quantitative approaches, a specific microhabitat is targeted with the goal of reducing inter-sample variability and increasing sensitivity (Resh and McElravy 1993; Rehn et al. 2007). In other bioassessment approaches, multiple microhabitats, such as riffles, runs, and pools, are sampled to represent a larger proportion of the macroinvertebrate assemblage (Resh and Jackson 1993; Gerth and Herlihy 2006; Blocksom et al. 2008). In general, these methods focus on benthic microhabitats (e.g., Ode et al. 2016), but some protocols have expanded sampling to include other often overlooked microhabitats, such as snags, vegetated banks, and submerged macrophytes (e.g., Massachusetts Department of Environmental Protection 1995; Florida Department of Environmental Protection 1996).

Focusing on benthic habitats alone when sampling aquatic macroinvertebrate assemblages could negatively bias sample collection (i.e., underrepresenting or overrepresenting certain assemblage members). Special concern may arise in the context of non-perennial streams (Larned et al. 2010; Acuña et al. 2014; Datry et al. 2014*b*), where key macroinvertebrate species exist outside the traditionally sampled benthic zone (Stubbington et al. 2017). Cycles of flooding and drying are common to these systems, causing stream habitats to shift from lotic (flowing) to lentic (non-flowing) to dry states (no surface water), with associated faunal changes (Boulton et al. 1992; Stanley et al. 1997; Bonada et al. 2007; Dewson et al. 2007). During non-lotic states, these systems require macroinvertebrate sampling methods that can accommodate a broad range of habitats, including those lacking substantial flow (Gerth and Herlihy 2006; Blanchette and Pearson 2012; Yeardley et al. 2020). Stream drying and reach fragmentation also cause the proportion of edge microhabitat to increase relative to the benthic area. Edge microhabitats are riparian-adjacent stream microhabitats not typically sampled using standard aquatic macroinvertebrate methods. They encompass various features such as submerged vegetation, roots, rocky stream margins, and overhanging banks and are known to harbor unique faunas (Moore 1987; Parsons and Norris 1996; Blanchette and Pearson 2012). Consequently, edge microhabitats may be particularly important in arid systems where flow intermittency increases the value of sampling edge habitats to gain a complete picture of aquatic macroinvertebrate assemblages.

Stream edge sampling becomes increasingly important when estimates of biodiversity at the reach scale are utilized in large-scale analyses across datasets because data compatibility is critical (Cao and Hawkins 2011). In recent years, advances have been made in our understanding of the effects of aridity and flow intermittency on the biodiversity of aquatic invertebrate assemblages by comparing independent data sets collected from different climatic regions (Datry et al. 2014*a*; Crabot et al. 2020; Vander Vorste et al. 2021). When benthic sampling methods are similar across data sets, the implicit assumption is that each data set adequately represents reach-scale biodiversity for its study. If benthic sampling methods systematically underestimate the number of taxa in arid-land stream reaches by excluding their dynamic edge habitats (Blanchette and Pearson 2012), then large-scale comparisons across climatic zones may become problematic.

Here, we ask: (1) to what magnitude does sampling of edge microhabitats increase our estimates of reach-level macroinvertebrate richness and community dissimilarity? and (2) do any observed effects of including edge samples with benthic samples vary geographically across an aridity gradient? To address these questions, we compared paired samples collected from the same reach using an established method for benthic sampling (Ode et al. 2016) and a method combining benthic and edge microhabitat sampling (Eppehimer et al. 2020). Samples were collected from replicate reaches across basins spanning nearly the entire East–West climatic aridity gradient across the United States. For each reach, we calculated differences in macroinvertebrate richness and composition between paired samples. To determine if the effects of edge sampling vary with aridity, we tested for relationships between paired sample differences in taxon richness and assemblage composition and aridity. We hypothesized that sampling edge microhabitats would significantly increase the number of taxa detected and change the composition of macroinvertebrate assemblages. We expected that the effects of edge sampling would be greater in more arid than in more mesic locations where edge habitat should be periodically abundant and species are accustomed to episodic drying.

Materials and procedures

Study basins, reaches, and aridity gradient

We established nine study basins across the United States spanning a wide range of aridity conditions (Fig. 1; Supporting Information Fig. S1; Supporting Information Dataset S1). Within each basin, we selected ten 150 m study reaches that were wadeable, minimally impacted, and separated by a minimum Euclidian distance of ~ 1 km. Using the Global Aridity Index and Potential Evapotranspiration Climate Database (Zomer et al. 2022), we determined the mean aridity index value from 1970 to 2000 for each reach (Supporting Information Fig. S2; Supporting Information Dataset S1). Reach aridity index values ranged from 0.17 to 1.31, with lower values corresponding to arid conditions and higher values corresponding to mesic conditions.

Sample collection, processing, and identification

We collected benthic macroinvertebrates following a modified version of California's Surface Water Ambient Monitoring Program protocol (SWAMP v2; Ode et al. 2016). At each reach within each basin, we collected two samples: (1) a reach-wide benthic (RWB) sample (following the SWAMP protocol) and (2) an edge sample (Eppehimer et al. 2020). Samples were not collected if the stream reach was completely dry.

For RWB sampling, we collected and composited 11 benthic subsamples from across each 150 m reach. Starting downstream and proceeding upstream, we selected subsample locations every 15 m, alternating channels left, center, and right (25%, 50%, and 75% of the wetted channel width, respectively). For streams > 30 cm wide and > 5 cm deep, we collected benthic samples using either a 500 μm mesh D-net or Surber sampler. In two cases, for streams < 30 cm wide or < 5 cm deep, we used a mini-Surber sampler. In reaches with appreciable flow, nets were placed immediately downstream of the subsample collection area such that dislodged material would flow downstream and be captured by the net. In cases of stagnant water or dense vegetation, we continuously moved the net through the areas while collecting. Each subsample collection area was one net-width long and wide (i.e., for a 0.3048 m wide net, the area sampled would be 0.0929 m^2). Within each subsample collection area, we cleaned all substrates larger than a golf ball into the net and set it aside. For larger rocks or bedrock, we thoroughly scoured the rock surface with our hands. When possible, we agitated the top ~ 10 cm of substrate. If a subsample location within reach was dry, no subsample was taken. At each reach, we recorded the total number of subsamples collected and the device used for sampling, allowing for the calculation of the total area sampled.

For edge sampling, we used a 500 μm mesh D-net to take and composite five sweeps (~ 0.33 m^2 each) along each reach's wetted edge in riparian-adjacent microhabitats. We targeted submerged vegetation, roots, rocky stream margins, and overhanging banks, spacing sweeps roughly evenly across the 150 m reach. The goal of the edge sampling was to detect taxa present in the reach missed by RWB sampling.

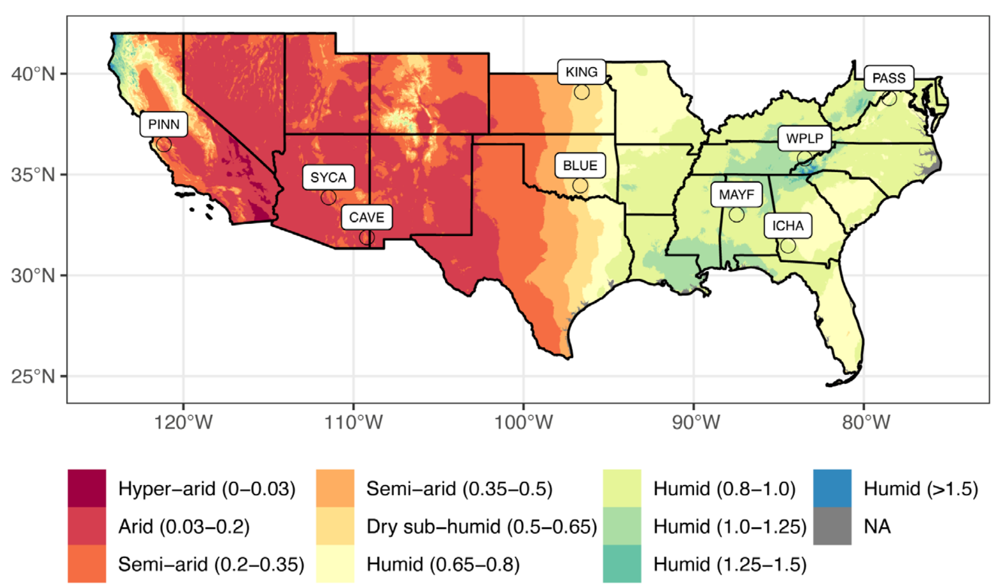

Fig. 1. Map of basin locations and mean aridity index from 1970 to 2000 (Supporting Information Dataset S1) across the United States (30 arc-second resolution). We chose basins spanning a wide range of aridity conditions from arid to mesic (range of reach values: 0.17–1.31; Supporting Information Fig. S2).

Following both RWB and edge sampling, we field elutriated and stored samples separately in 95% ETOH, which was replaced at least once within 24 h or until the color was stable.

In the laboratory, we examined a minimum of 500 macroinvertebrate specimens per sample type per reach. After visually assessing specimen densities, we used a Folsom plankton splitter to produce subsamples (hereafter, "splits") that we thought would contain at least 500 specimens, processing additional splits if we found fewer than 500 specimens. For samples with less than 500 specimens total, we processed the entire sample collected. We then identified specimens to genus-level for insects and to order- or family- for non-insects using available keys (Thorp and Covich 2001; Stewart and Stark 2002; Wiggins 2015; Merritt et al. 2019; Tennessen 2019).

Prior to analyses, we adjusted counts for each taxon from each sample based on the split processed in the lab to estimate the number of individuals in an entire sample (e.g., we adjusted a count of 15 individuals for a ½ sample split to 30). In addition, we sorted through any available unprocessed remnant material (i.e., unsorted splits) and identified any large and rare taxa not found in the processed split(s). Counts for these specimens were added unadjusted to RWB and edge counts that had been adjusted based on the lab split processed. For RWB samples, we also adjusted counts for each taxon by dividing them by the total area sampled during RWB sampling at each reach (e.g., for an area sampled of 0.8361 m^2, we adjusted a count of 10 to 11.96 and rounded up to 12 individuals m^{-2}). Finally, we created a "combined sample" from RWB samples and edge samples using a "plus one" method (Supporting Information Dataset S2; Eppehimer et al. 2020). The "plus one" method assumes that most of the assemblage is in the benthic zone, and so it heavily weighs the RWB sample in creating a combined sample. Taxa that were detected in the edge sample but not the RWB sample were added to the RWB sample with an abundance count of one regardless of the number of individuals present in the edge samples. This approach allows those edge-only taxa to be considered in overall reach-scale diversity and composition analyses, but does not weigh them as heavily in abundance-based considerations (Eppehimer et al. 2020). We focus on this "plus one" combination method in main text analyses, but we considered alternative combination treatments in additional sensitivity analyses (Fig. 2).

Data analysis

Observed taxon richness

For each reach, we determined the number of taxa observed in RWB and combined ("plus one") samples (Supporting Information Dataset S3). To determine the effect of edge sampling on richness, we subtracted values of the observed richness of RWB samples from those of combined samples. Because absolute (gamma) richness values varied across basins, we scaled differences in richness between sample types by calculating percentage increases in richness in combined samples vs. RWB samples. Depending on the distributions of the data, we ran one-sample *t*-tests or Wilcoxon signed-rank tests to determine if percentage increases in richness were significantly greater than zero for all basins together and for each basin individually.

Asymptotic taxon richness

For both RWB and combined samples, we calculated estimated asymptotic taxon richness values using the ChaoRichness function (Chao 1984, 1987) from the R package iNEXT (Hsieh et al. 2016; R Core Team 2020; Supporting Information Dataset S3). As for observed richness, we subtracted the taxon richness of RWB samples from those of combined samples, scaled increases in richness as percentages, and ran one-sample *t*-tests or Wilcoxon signed-rank tests as appropriate. We assessed the sensitivity of analyses of asymptotic taxon richness to the "plus one" method by also running analyses with combined sample counts that were created by summing all abundance counts from RWB and edge samples (hereafter the "sum" combination method; Fig. 2; Supporting Information Dataset S4).

Composition

We tested for differences in the taxonomic composition of combined and RWB samples by calculating Bray–Curtis dissimilarity between sample types for each reach (Supporting Information Dataset S3). For all dissimilarities across basins and those from each basin individually, we ran one-sample *t*-tests or Wilcoxon signed-rank tests. We assessed the sensitivity of these results to the "plus one" method by also running analyses with combined sample taxa counts determined by the "sum" method and by converting compositional matrices to presence/absence data (hereafter "presence/absence" combination method; Fig. 2; Supporting Information Dataset S5).

Relationships between richness, composition, and aridity

For richness, we used linear regressions to assess how aridity was associated with differences between RWB and combined samples. For percent increases in observed and asymptotic taxon richness, we used reach aridity as a predictor and percent increases in observed and asymptotic richness as responses. To evaluate regression assumptions, we visually inspected plots of residuals, calculated Cook's *D*, and reran analyses without high-influence data. For asymptotic richness, we evaluated the sensitivity of trends in percent increases by also testing for relationships using linear regression with taxa counts determined by the "sum" method.

To assess trends in differences in assemblage composition of RWB and combined samples with aridity, we ran beta regressions with aridity as a predictor and Bray–Curtis dissimilarity as a response in the R package betareg (Ferrari and Cribari-Neto 2004; Cribari-Neto and Zeileis 2010; Simas et al. 2010). We used beta regression because it can be used to model variables that assume values in the standard unit interval (0, 1). We fit beta regressions with and without variable dispersions and chose the best model based on a

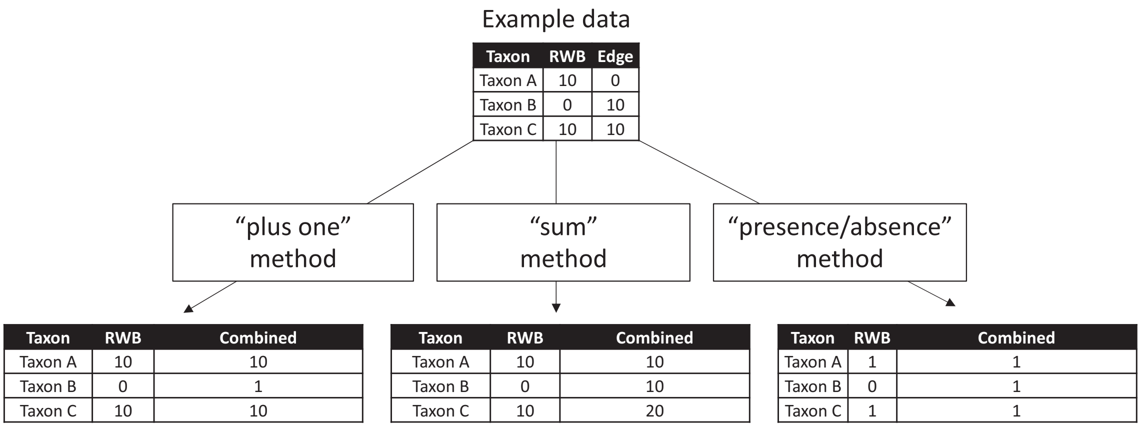

Fig. 2. Diagram showing the methods used to combine reach-wide benthic (RWB) and edge samples for sensitivity analyses. The top table in the figure shows example data collected from a study reach (a set of paired RWB and edge samples). The three tables below show how we used the "plus one," "sum," and "presence/absence" methods to combine RWB and edge data into a "combined" sample type. Comparisons in analyses were based on RWB and combined samples, but combination methods employed varied by analysis type based on expected sensitivity.

log-likelihood test. We also ran beta regression with taxa counts calculated using the "sum" and "presence/absence" methods.

Assessment

Observed richness

Significant percent increases in observed taxon richness were found for combined samples relative to RWB samples for all basins together and for each basin independently (Fig. 3A). Increases in observed richness ranged from 1 to 27 taxa (median = 12; Supporting Information Dataset S3), whereas percent increases in observed richness ranged from three to 129% (median = 29%).

Asymptotic richness

Significant percent increases in estimated asymptotic taxon richness overall and for each basin independently were found when comparing the combined to RWB samples (Fig. 3B). Estimated increases in asymptotic richness ranged from 1 to 774 taxa (median = 87 taxa), whereas estimated percent increases in asymptotic richness ranged from 3% to 1488% (median = 173%). We also found significant increases in estimated asymptotic taxon richness overall and for all basins using the alternate "sum" method, except for basin WPLP (*t*-test $p = 0.069$; Supporting Information Fig. S3). Changes in asymptotic richness ranged from −64 to 132 taxa (median = 14 taxa), whereas percent changes in asymptotic richness ranged from −53% to 188% (median = 31%). Negative values represent cases where the estimated combined sample asymptotic richness was less than that of the RWB sample alone.

Differences in taxonomic composition

Small but significant differences in taxonomic composition overall and for each basin independently were found when comparing RWB and combined samples (Fig. 3C). Bray–Curtis dissimilarities ranged from 0.000 to 0.103 (median = 0.005). Significant compositional differences overall and by basin were also observed using the alternate "sum" and "presence/absence" methods (Supporting Information Fig. S4). Utilizing the "sum" method, Bray–Curtis dissimilarities ranged from 0.015 to 0.529 (median = 0.174; Supporting Information Fig. S4A), whereas using the "presence/absence" method dissimilarities ranged from 0.017 to 0.393 (median = 0.128; Supporting Information Fig. S4B).

Relationships between richness, composition, and aridity

Percent increases in observed and asymptotic richness between RWB and combined samples were largest in arid basins and decreased in more mesic basins (Fig. 4A,B). However, using the alternate "sum" method, we did not find evidence for any relationship between asymptotic richness and aridity (Supporting Information Fig. S5). These results were insensitive to high-influence data.

Assemblage dissimilarities determined using the "plus one" method were smallest in arid basins and increased in more mesic basins, but changes were small (Fig. 4C). Using the alternate "sum" and "presence/absence" methods, we found the opposite trend: assemblage dissimilarities were largest in arid basins, decreased in more mesic basins, and changes were large (Supporting Information Fig. S6A,B).

Discussion

The addition of stream edge sampling to an established RWB sampling protocol had significant effects on the characterizations of macroinvertebrate assemblage structure. Adding edge sampling resulted in median increases in observed and asymptotic richness of 29% and 173%, respectively. Significant differences in assemblage composition were found using different sampling methods, with median dissimilarity values increasing by as much as 0.178. Differences in richness significantly increased with or were unrelated to aridity, whereas differences

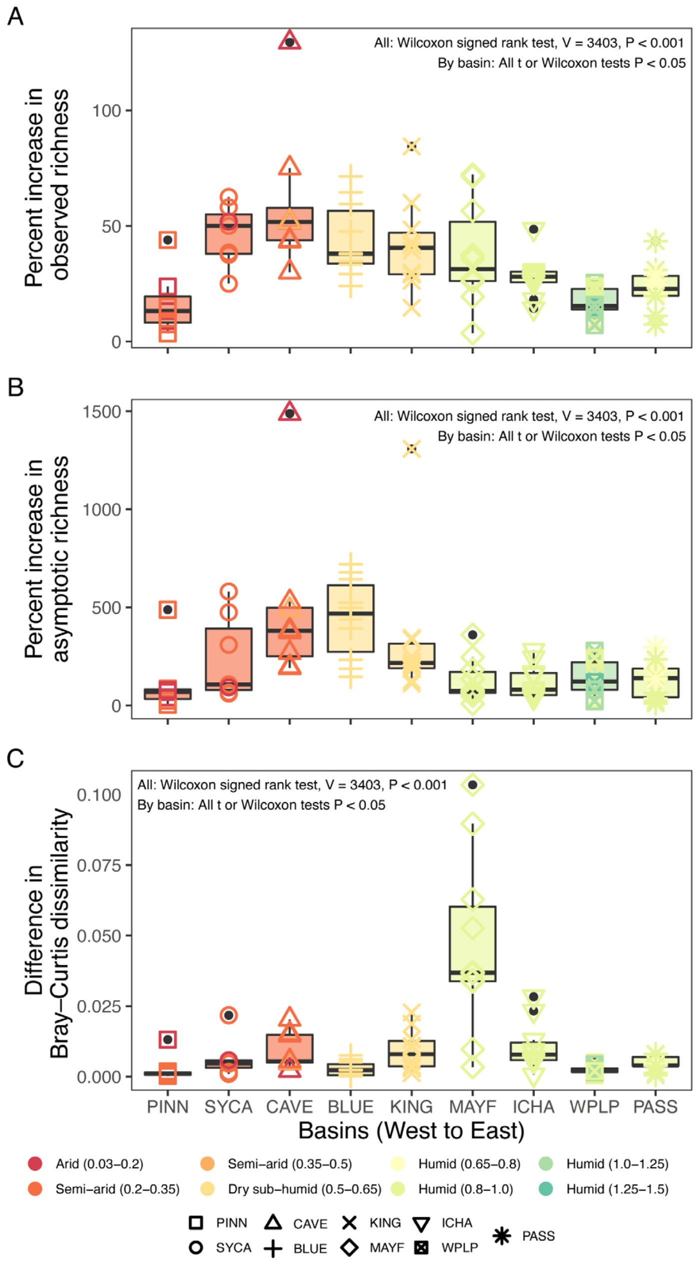

Fig. 3. Significant increases in observed richness (**A**), asymptotic richness (**B**), and Bray–Curtis dissimilarities (**C**) between RWB and combined ("plus one" method) samples for each basin. Boxplots are colorized by the average mean aridity index of all reaches within basins, whereas data points are colorized by individual mean reach aridity index values. All data—treated together and by basin—are significantly different from zero (all $p < 0.05$).

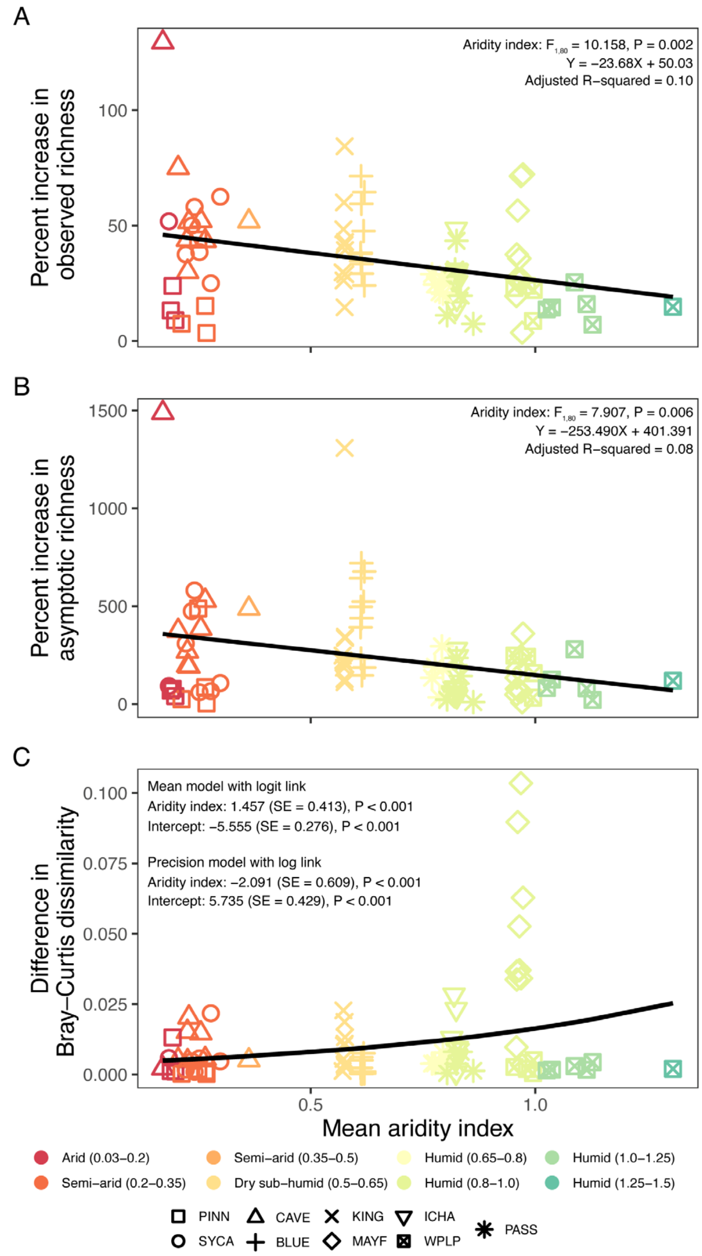

Fig. 4. Relationships between mean basin aridity index and percent increase in observed richness (**A**), percent increase in asymptotic richness (**B**), and difference in Bray–Curtis dissimilarity between combined ("plus one" method) and RWB sample types (**C**). Trends in both percent increase in observed richness and asymptotic richness were insensitive to removal of high-influence data (i.e., without high-influence data: *Percent increase in observed richness*: Aridity index: $F_{1,73} = 30.328$, $p < 0.001$, $Y = -32.285X + 55.493$, adjusted $R^2 = 0.28$; *Percent increase in asymptotic richness*: Aridity index: $F_{1,78} = 6.434$, $p = 0.013$, $Y = -160.582X + 312.633$, adjusted $R^2 = 0.06$).

in assemblage composition both increased and decreased with aridity depending on how edge and benthic samples were combined. Collectively, our results demonstrate that targeted sampling of stream edge microhabitats can have significant effects on diversity and compositional metrics for aquatic macroinvertebrate assemblages and that those effects can vary with aridity.

Sensitivity analyses revealed that estimates of asymptotic richness and assemblage dissimilarities were sensitive to the method used for combining RWB and edge samples. Of note were some cases where the estimated asymptotic richness of the combined sample was lower than that of the corresponding

RWB sample from the same study reach (Supporting Information Fig. S3). This result is likely caused by the “sum” method decreasing the relative proportions of rare taxa in combined samples and thus estimated values of asymptotic richness by using counts for edge taxa that are inflated by targeted sampling. Assemblage dissimilarities between sample types also increased by a factor of 35 using the “sum” method (Supporting Information Fig. S5A) and a factor of 25 using the “presence/absence” method (Supporting Information Fig. S5B). Both of these alternate combination methods place much higher weight on rare edge taxa in the Bray–Curtis dissimilarity formula when compared to the “plus one” method, highlighting the need for careful consideration of combination and transformation techniques in multivariate statistics (McCune et al. 2002). Regardless of these sensitivities, the overwhelming majority of our analyses showed increases in central tendencies of asymptotic richness and assemblage dissimilarities regardless of which of the three combination methods was used.

Similarly, relationships between differences in richness, assemblage composition, and aridity were sensitive to the method used for combining RWB and edge samples. Using the “sum” method, we found no relationship between increases in asymptotic richness and aridity (Supporting Information Fig. S5). The lack of a relationship can be explained by the effects of the “sum” combination method on the asymptotic richness that was likely to be uneven across the aridity gradient (i.e., strong effects in basins with very distinct edge assemblages and less so elsewhere). Using the “sum” and “presence/absence” methods, we found that assemblage dissimilarity increased rather than decreased with aridity (Supporting Information Fig. S6). These results are consistent with our hypothesis that differences between RWB and combined samples should be greater in arid study basins, where edge habitats are common and local species are accustomed to flow intermittency (Blanchette and Pearson 2012; Bogan et al. 2015). However, the “plus one” method resulted in decreasing assemblage dissimilarity between RWB and combined samples with increasing aridity, highlighting that the exact manner in which and how rare taxa are added to data sets can have complex effects and should be carefully considered (van Sickle et al. 2007).

The sensitivity of our results highlights some vulnerabilities of our approach. First, we found that assemblages were represented more comprehensively by combining RWB and edge sampling. In this case, the need to combine quantitative and qualitative methods is unavoidable because of the complex nature of edge habitats, and in presenting results from three different sample combination methods, we aimed to bracket the range of possibilities for analytical outcomes. Second, in combining RWB and edge samples, the sampled area of the combined sample was always larger than that of the RWB sample alone, and thus, we should expect to find more taxa based on basic species/area relationships (MacArthur and Wilson 1963). While we concede that it is possible that increasing benthic sampling efforts could detect some taxa found in edge samples, it is unlikely to detect taxa that are edge specialists (Bogan et al. 2015). Third, the addition of edge sampling prompts the question of how many additional microhabitats should or could be included in sampling efforts. While ideally, all microhabitats would be sampled, we recognize that there are tradeoffs between effort and information gained (Chessman et al. 2007), and thus, we recommend further systematic examinations of the added value of sampling specific microhabitats. Sampling tradeoffs may be particularly relevant in systems where there are reasons to expect a distinct or unique fauna in different microhabitats, such as hyporheic species in non-perennial streams (Wood et al. 2010; Datry 2012; Stubbington 2012) and edge species in arid regions (Chessman et al. 2007). Taken together, we acknowledge the aforementioned vulnerabilities and do not aim to prescribe a new standard sampling method, but rather to provoke thought about which sampling methods capture sufficient data needed for research questions of interest.

Another consideration in interpreting our results is that aridity conditions are not static and they do not change uniformly across the continent. For example, in some of our western study basins, aridity has increased in the last two decades (Williams et al. 2020), and the widely used long-term aridity index, which we used in our study, has not yet incorporated those increasingly dry years (Zomer et al. 2022). However, we do not believe that the recent drought has strongly impacted our results for two reasons. First, there is often a lag of many years between a period of enhanced aridity and declines in baseflow (or complete drying) in groundwater-dependent streams in the western USA (e.g., Meyers et al. 2021). Second, the magnitude of differences in aridity between our western and eastern basis is already so large (with western basins being up to eight times more arid than eastern basins) that small increases in aridity in some western basins over the last 20 years are unlikely to affect the continental-scale patterns we observed.

Despite the associated challenges, the detection of edge sampling effects that vary across conditions of aridity could have important implications for studies of stream macroinvertebrate assemblages. Systematic omission of edge faunas could lead to inaccurate estimates of reach-scale diversity. Moreover, the effects of missing edge taxa that vary with climate could lead to the misrepresentation of large-scale diversity trends. For example, if benthic samples from arid regions consistently underestimate reach-scale diversity, then data syntheses that use these data could overestimate the impacts of drying or aridity on aquatic biodiversity unless analyses focus on examining trends within each separate dataset (Datry et al. 2014*a*). Given that most aquatic biodiversity studies occur in perennial streams, misrepresenting the relatively low amount of data from non-perennial systems could be particularly problematic in meta-analyses, given the relative weight of inference placed on fewer observations. In demonstrating that our estimates of reach-scale invertebrate diversity and composition varied significantly by sampling method and aridity, we hope that researchers will carefully consider (1) their sampling methods in field studies that span climatic zones and (2) the comparability of existing data sets when conducting data syntheses (Cao and Hawkins 2011). These considerations are especially important in the context of non-perennial streams (Larned et al. 2010; Acuña et al. 2014; Datry et al. 2014*b*) and streams in arid regions (Chessman et al. 2007).

References

Acuña, V., and others. 2014. Why should we care about temporary waterways? Science **343**: 1080–1081. doi:10.1126/science.1246666

Blanchette, M. L., and R. G. Pearson. 2012. Macroinvertebrate assemblages in rivers of the Australian dry tropics are highly variable. Freshw. Sci. **31**: 865–881. doi:10.1899/11-068.1

Blocksom, K. A., B. C. Autrey, M. Passmore, and L. Reynolds. 2008. A comparison of single and multiple habitat protocols for collecting macroinvertebrates in wadeable streams. J. Am. Water Resour. Assoc. **44**: 577–593. doi:10.1111/j.1752-1688.2008.00183.x

Bogan, M. T., K. S. Boersma, and D. A. Lytle. 2015. Resistance and resilience of invertebrate communities to seasonal and supraseasonal drought in arid-land headwater streams. Freshw. Biol. **60**: 2547–2558. doi:10.1111/fwb.12522

Bonada, N., M. Rieradevall, and N. Prat. 2007. Macroinvertebrate community structure and biological traits related to flow permanence in a Mediterranean river network. Hydrobiologia **589**: 91–106. doi:10.1007/s10750-007-0723-5

Boulton, A. J., C. G. Peterson, N. B. Grimm, and S. G. Fisher. 1992. Stability of an aquatic macroinvertebrate community in a multiyear hydrologic disturbance regime. Ecology **73**: 2192–2207. doi:10.2307/1941467

Callisto, M., R. Mugnai, D. M. P. Castro, and M. S. Linares. 2021. Sampling methods for aquatic insects, p. 523–543. *In* J. C. Santos and G. W. Fernandes [eds.], Measuring arthropod biodiversity. Springer International Publishing. doi:10.1007/978-3-030-53226-0_20

Cao, Y., and C. P. Hawkins. 2011. The comparability of bioassessments: A review of conceptual and methodological issues. J. North Am. Benthol. Soc. **30**: 680–701. doi:10.1899/10-067.1

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. Scand. J. Stat. **11**: 265–270. doi:10.2307/4615964

Chao, A. 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics **43**: 783. doi:10.2307/2531532

Chessman, B., S. Williams, and C. Besley. 2007. Bioassessment of streams with macroinvertebrates: Effect of sampled habitat and taxonomic resolution. J. North Am. Benthol. Soc. **26**: 546–565. doi:10.1899/06-074.1

Crabot, J., J. Heino, B. Launay, and T. Datry. 2020. Drying determines the temporal dynamics of stream invertebrate structural and functional beta diversity. Ecography **43**: 620–635. doi:10.1111/ecog.04835

Cribari-Neto, F., and A. Zeileis. 2010. Beta regression in R. J. Stat. Softw. **34**: 1–24. doi:10.18637/jss.v034.i02

Datry, T. 2012. Benthic and hyporheic invertebrate assemblages along a flow intermittence gradient: Effects of duration of dry events. Freshw. Biol. **57**: 563–574. doi:10.1111/j.1365-2427.2011.02725.x

Datry, T., S. T. Larned, K. M. Fritz, M. T. Bogan, P. J. Wood, E. I. Meyer, and A. N. Santos. 2014*a*. Broad-scale patterns of invertebrate richness and community composition in temporary rivers: Effects of flow intermittence. Ecography **37**: 94–104. doi:10.1111/j.1600-0587.2013.00287.x

Datry, T., S. T. Larned, and K. Tockner. 2014*b*. Intermittent rivers: A challenge for freshwater ecology. Bioscience **64**: 229–235. doi:10.1093/biosci/bit027

Dewson, Z. S., A. B. W. James, and R. G. Death. 2007. A review of the consequences of decreased flow for instream habitat and macroinvertebrates. J. North Am. Benthol. Soc. **26**: 401–415. doi:10.1899/06-110.1

Eppehimer, D. E., H. Hamdhani, K. D. Hollien, and M. T. Bogan. 2020. Evaluating the potential of treated effluent as novel habitats for aquatic invertebrates in arid regions. Hydrobiologia **847**: 3381–3396. doi:10.1007/s10750-020-04343-6

Ferrari, S. L. P., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. J. Appl. Stat. **31**: 799–815. doi:10.1080/0266476042000214501

Florida Department of Environmental Protection. 1996. Standard operating procedures for biological assessment.

Gerth, W. J., and A. T. Herlihy. 2006. Effect of sampling different habitat types in regional macroinvertebrate bioassessment surveys. J. North Am. Benthol. Soc. **25**: 501–512. doi:10.1899/0887-3593(2006)25[501:EOSDHT]2.0.CO;2

Hsieh, T. C., K. H. Ma, and A. Chao. 2016. iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol. Evol. **7**: 1451–1456. doi:10.1111/2041-210X.12613

Jackson, J., V. Resh, D. Batzer, R. Merritt, and K. Cummins. 2019. Sampling aquatic insects: Collection devices, statistical considerations, and rearing procedures. *In* R. Merritt, K. Cummins, and M. B. Berg [eds.], An introduction to the aquatic insects of North America. Kendall Hunt Publishing Company.

Larned, S. T., T. Datry, D. B. Arscott, and K. Tockner. 2010. Emerging concepts in temporary-river ecology. Freshw. Biol. **55**: 717–738. doi:10.1111/j.1365-2427.2009.02322.x

Ligeiro, R., R. M. Hughes, P. R. Kaufmann, J. Heino, A. S. Melo, and M. Callisto. 2020. Choice of field and laboratory methods affects the detection of anthropogenic disturbances using stream macroinvertebrate assemblages. Ecol. Indic. **115**: 106382. doi:10.1016/j.ecolind.2020.106382

MacArthur, R. H., and E. Wilson. 1963. An equilibrium theory of insular zoogeography. Int. J. Org. Evol. **17**: 373–387. doi:10.1017/CBO9781107415324.004

Massachusetts Department of Environmental Protection. 1995. Massachusetts DEP preliminary biological monitoring and assessment protocols for wadable rivers and streams.

McCune, B., J. B. Grace, and D. L. Urban. 2002. Analysis of ecological communities. Mjm Software Design.

Merritt, R., K. Cummins, and M. B. Berg. 2019. An introduction to the aquatic insects of North America. Kendall Hunt Publishing.

Meyers, Z. P., M. D. Frisbee, L. K. Rademacher, and N. S. Stewart-Maddox. 2021. Old groundwater buffers the effects of a major drought in groundwater-dependent ecosystems of the eastern Sierra Nevada (CA). Environ. Res. Lett. **16**: 044044. doi:10.1088/1748-9326/abde5f

Moore, K. M. S. 1987. Ecology of aquatic habitats associated with stream margins. Oregon State University.

Ode, P. R., A. E. Fetscher, and L. B. Busse. 2016. Standard operating procedures for the collection of field data for bioassessments of California wadeable streams: Benthic macroinvertebrates, algae, and physical habitat.

Parsons, M., and R. Norris. 1996. The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. Freshw. Biol. **36**: 419–434. doi:10.1046/j.1365-2427.1996.00088.x

R Core Team. 2020. R: A language and environment for statistical computing.

Rehn, A. C., P. R. Ode, and C. P. Hawkins. 2007. Comparisons of targeted-riffle and reach-wide benthic macroinvertebrate samples: Implications for data sharing in stream-condition assessments. J. North Am. Benthol. Soc. **26**: 332–348. doi:10.1899/0887-3593(2007)26[332:COTARB]2.0.CO;2

Resh, V. H., and J. K. Jackson. 1993. Rapid assessment approaches to biomonitoring using benthic macroinvertebrates, p. 195–234. *In* D. M. Rosenberg and V. H. Resh [eds.], Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall.

Resh, V., and E. McElravy. 1993. Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates, p. 159–194. *In* D. Rosenberg and V. Resh [eds.], Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall.

Rosenberg, D. M., and V. H. Resh [eds.]. 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall.

Simas, A. B., W. Barreto-Souza, and A. V. Rocha. 2010. Improved estimators for a general class of beta regression models. Comput. Stat. Data Anal. **54**: 348–366. doi:10.1016/j.csda.2009.08.017

Stanley, E. H., S. G. Fisher, and N. B. Grimm. 1997. Ecosystem expansion and contraction in streams. Bioscience **47**: 427–435. doi:10.2307/1313058

Stewart, K. W., and B. P. Stark. 2002. Nymphs of North American stonefly genera. The Caddis Press.

Stubbington, R. 2012. The hyporheic zone as an invertebrate refuge: A review of variability in space, time, taxa and behaviour. Mar. Freshw. Res. **63**: 293. doi:10.1071/MF11196

Stubbington, R., M. T. Bogan, N. Bonada, A. J. Boulton, T. Datry, C. Leigh, and R. Vander Vorste. 2017. The biota of intermittent Rivers and ephemeral streams: Aquatic invertebrates, p. 217–243. *In* T. Datry, N. Bonada, and A. Boulton [eds.], Intermittent rivers and ephemeral streams. Academic Press. doi:10.1016/B978-0-12-803835-2.00007-3

Tennessen, K. J. 2019. Dragonfly nymphs of North America. Springer International Publishing.

Thorp, J. H., and A. P. Covich [eds.]. 2001. Ecology and classification of North American freshwater invertebrates. Academic Press.

van Sickle, J., D. P. Larsen, and C. P. Hawkins. 2007. Exclusion of rare taxa affects performance of the *O/E* index in bioassessments. J. North Am. Benthol. Soc. **26**: 319–331. doi:10.1899/0887-3593(2007)26[319:EORTAP]2.0.CO;2

Vander Vorste, R., and others. 2021. Climatic aridity increases temporal nestedness of invertebrate communities in naturally drying rivers. Ecography **44**: 860–869. doi:10.1111/ecog.05349

Wiggins, G. B. 2015. Larvae of the North American caddisfly genera. Univ. of Toronto Press.

Williams, A. P., and others. 2020. Large contribution from anthropogenic warming to an emerging North American megadrought. Science **368**: 314–318. doi:10.1126/science.aaz9600

Wood, P. J., A. J. Boulton, S. Little, and R. Stubbington. 2010. Is the hyporheic zone a refugium for aquatic macroinvertebrates during severe low flow conditions? Fundam. Appl. Limnol. **176**: 377–390. doi:10.1127/1863-9135/2010/0176-0377

Yeardley, R., Jr., S. Jacobs, K. Fritz, and W. Thoeny. 2020. Comparison of three macroinvertebrate sampling methods for use in assessment of water quality changes in flashy urban streams. J. Environ. Prot. **11**: 585–609. doi:10.4236/jep.2020.118035

Zomer, R. J., J. Xu, and A. Trabucco. 2022. Version 3 of the global aridity index and potential evapotranspiration database. Sci. Data **9**: 409. doi:10.1038/s41597-022-01493-1

Acknowledgments

We thank members of the Bogan Lab at The University of Arizona for reviewing this manuscript during its preparation; this research was supported by USA National Science Foundation Macrosystems Biology and NEON-Enabled Science program (awards: DEB-2207680, DEB-1802766, DEB-1802811, DEB-1802714, DEB-2150626).

Submitted 17 November 2022
Revised 15 January 2024
Accepted 26 January 2024

Associate editor: Michael Beman