

DNA barcoding vs morphology: Assessing different methods for invertebrate identification and determination of industrial effluent impacts

Laura M. Dutheil

laura.dutheil@hotmail.ca

University of Guelph (Integrative Biology)

Karl Cottenie

University of Guelph (Integrative Biology)

Helga Sonnenberg

Ecoreg Solutions

Dirk Steinke

University of Guelph (Integrative Biology)

Sarah J. Adamowicz

University of Guelph (Integrative Biology)

Research Article

Keywords: benthic invertebrates, EEM, assessment, biomonitoring

Posted Date: March 9th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8253547/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

In Canada, all active metal and diamond mines and pulp and paper mills discharging effluent are required to monitor potential effluent related impacts on benthic invertebrate communities through regulated Environmental Effects Monitoring (EEM) programs. Currently, EEMs include sampling benthic invertebrates and identifying them morphologically to the family level. We sought to directly compare traditional taxonomic identification methods with DNA barcoding of each specimen as they apply to EEM studies. Five industry sites were sampled, each with a sample area receiving effluent and a reference area not receiving effluent for a total of 67 benthic invertebrate sample stations across all five sites. Specimens were identified morphologically and with DNA barcoding and statistical differences of four endpoints between paired freshwater areas were assessed. Additionally, OTU (operational taxonomic unit) level analyses were compared to the family-level results to determine if the OTU level could be more informative of differences in community composition between effluent receiving environments and their reference areas. Family-level morphology vs. family-level barcode-based identifications delivered the same conclusions for paired-area comparisons over 70% of the time. DNA barcoding results at the OTU level differed when determining impacts between paired areas 19–30% of the time across all four endpoints compared to family-level barcoding and family-level morphology, respectively. Overall, DNA barcoding could be used in EEM studies as it demonstrated similar results as morphological identifications at the family level and provides greater taxonomic resolution and sensitivity for detecting paired area differences.

Introduction

To help assess the adequacy of effluent regulations and to manage and mitigate impacts, the Canadian Government, through the Fisheries Act, enacted a national Environmental Effects Monitoring (EEM) program, which specifies monitoring requirements of effluent and the associated receiving environments it is discharged into (Environment and Climate Change Canada (ECCC), 2008). Within this article, reference to the EEM program encompasses both the Pulp and Paper Effluent Regulations (Legislative Services Branch (LSB), 2018), as well as the Metal and Diamond Mining Effluent Regulations (LSB, 2024). EEM studies are performed every three years for all active mines and mills and consist of several monitoring components, including the amount of effluent discharged, concentrations of deleterious substances, and monitoring for potential impacts on fish and fish habitat (Government of Canada (GC), 2002). The biotic components of an EEM study focus on the health of fish populations (using survival, growth, reproduction, and energy storage endpoints) and on fish habitat by evaluating multiple benthic invertebrate community endpoints (GC, 2002). An effect on any of the EEM components is determined by a significant difference in endpoints between the area receiving effluent (exposure area) and a nearby area unimpacted by effluent (reference area). For benthic invertebrates, there are four 'effect' endpoints; total benthic invertebrate density, taxonomic richness, taxonomic evenness (using Simpson's evenness index), and community composition (using Bray-Curtis dissimilarity) (ECCC, 2012). The benthic invertebrate surveys were designed to assess habitat quality and the availability of benthic invertebrates as fish food in exposed receiving environments (lakes and streams) (ECCC, 2012). When the EEM

program was first implemented, many practitioners used the Lowest Practical Level (LPL), which was often genus level morphological identifications. Currently, the required level for invertebrate taxonomic identification for the calculation of endpoints is at the family level resolution (ECCC, 2012).

Benthic invertebrates in EEM studies are identified by manually sampling and sorting specimens from detritus followed by family-level classification using morphological characteristics. While morphological identification is more accessible for researchers and industry as it requires no specialized equipment beyond the availability of a microscope, it also has several shortcomings. The various sizes, life stages, and complex morphologies of invertebrates often hamper morphological identification, and as a result, taxonomy datasets have a taxonomic resolution that is not consistent across all organisms (Hewlett, 2000; Lenat & Resh, 2001; Pawlowksi et al., 2018). This inconsistency is further complicated by the experience level of the taxonomist conducting the identifications. Other constraints of morphological taxonomic classification include specimen quality, the availability of identification keys and resources, and availability of highly experienced taxonomic experts (Bush et al., 2019).

With the advent of newer methods for identification that use specimen DNA, morphology often begins to fall short when comparing the quality of taxonomic classification and time efficiency (Gibson et al., 2015; Bush et al., 2019; Pawlowski et al., 2018). Molecular tools like DNA barcoding (Hebert et al. 2003) can bridge many of the limitations listed above and help support morphological identifications. Relying on the extracted DNA of specimens, these methods can provide accurate and high-resolution identifications, often to the species level or a species proxy, known as an operational taxonomic unit (OTU) (Sweeney et al., 2011; Gibson et al., 2015; Pawlowski et al., 2018). DNA barcoding is also helpful in linking life stages, as diagnostic characteristics are often present in adults only, making morphological larval identifications difficult (Sweeney et al., 2011). Most importantly, barcoding has shown to yield similar results to morphological methods when used in biomonitoring, often demonstrating greater accuracy (Sweeney et al., 2011; Gibson et al., 2015; Pawlowski et al., 2018).

While DNA metabarcoding is receiving increasing attention in the literature for use in biomonitoring studies (Fernández et al., 2018; Uchida et al., 2020; Brantschen et al., 2021), EEMs rely on individual-based endpoints that cannot currently be fully achieved through metabarcoding. There is thus a need to compare morphological methods with DNA barcoding, as barcoding provides a molecular, specimen-by-specimen approach that aligns with EEM methods, procedures, and outcomes.

DNA barcoding is a molecular method for specimen identification that uses a specimen-by-specimen workflow and requires a similar process of sample sorting as morphological methods. Therefore, barcoding is subject to some of the same limitations as morphological methods, as the sample processing can be time-consuming, and both methods often rely on the subsampling of a larger sample, which can lead to variable results in analyses (Clarke et al., 2006; Vlek et al., 2006). Due to the lengthy nature of specimen sorting, portions of the total sample are taken, sorted, and adjusted to represent the entire sample. Depending on the size of the subsample, this practice can result in differing endpoint results, thereby hindering the ability to properly identify water bodies as impacted or unimpacted by a

given contaminant (Doberstein et al., 2000; Vlek et al., 2006; Petkovska & Urbanič, 2010). Despite this shortcoming, DNA barcoding sets itself apart from morphology by its increased taxonomic resolution and by its greater ability to encompass and assess the broader invertebrate community (Sweeney et al., 2011; Carew et al., 2013). Gibson et al. (2015) noted that while DNA barcoding was able to replicate all studied metrics found morphologically, Sweeney et al. (2011) observed that barcoding resulted in a higher number of taxa and a greater degree of accuracy for the calculated metrics.

There is continued discussion as to which taxonomic level is most appropriate when sampling invertebrates for biomonitoring purposes. While some studies, including EEMs, only require family-level identifications, many others have found this identification level unsatisfactory (Lenat & Resh, 2001; King & Richardson, 2002; Molineri et al., 2020). Specifically, family and even genus-level identifications of invertebrates were shown to be inadequate when classifying habitat quality compared to species-level identification (Lenat & Resh, 2001, King & Richardson, 2002; Carew et al., 2011; Molineri et al., 2020). There is concern that sampling at the family level will mask much of the variability found at the species level, as several families and genera are known to have species of varying pollution tolerances (Resh & Unzicker, 1975; Lenat & Resh, 2001).

Freshwater invertebrates are an important part of a healthy ecosystem and ensuring that they are accurately monitored at the appropriate identification level is vital. As an important foundation of freshwater food webs, benthic invertebrates provide nutrients primarily to fish, but also to terrestrial mammals, reptiles, amphibians, and birds (Macadam & Stockan, 2015). Invertebrates aid in the breakdown of organic matter in the water column, cycling of nutrients, and water filtration (Losey & Vaughan, 2006; Macadam & Stockan, 2015). Ensuring accurate biomonitoring of freshwater benthic invertebrates is a first step in protecting these ecosystems and the services they provide.

To evaluate whether the DNA barcoding method can yield similar results to the currently used morphological methods of identifying invertebrates, we conducted sampling at four mines and one pulp and paper mill in Canada following EEM protocols and identified invertebrates using traditional morphological identifications and DNA barcoding. We hypothesized that family-level identifications obtained using DNA barcoding would yield similar results to the morphological results at the family level as molecular approaches have been shown to replicate the results found by morphology (Martins et al., 2021; Shackleton et al., 2021). Further, with the addition of higher taxonomic resolution through molecular methods, we hypothesized that OTU-level impacts will vary from family-level and provide more accurate classifications as higher taxonomic resolution has shown to provide greater insight into the quality of impacted streams (Sweeney et al., 2011; Stein et al., 2014; Macher et al., 2016). This study is the first to perform a specimen-based comparison of morphological and molecular methods for EEMs in Canada and is important for providing biological and methodological insights into the use of barcoding methods relative to those derived from traditional morphological taxonomy.

Methods

Aquatic invertebrates were sampled by EcoReg Solutions from five sites including four mines and one paper mill in Canada in the fall of 2020 and 2021. Each site consists of an area receiving effluent (exposure) and a corresponding reference area not receiving effluent, with each consisting of five to six benthic invertebrate community sample stations. Part of the EEM study design also requires the careful selection of exposure and reference areas as they must reflect similar habitats and watersheds to reduce natural variability between sample areas. The sites in this study will be denoted with numbers (Site 1, or S1, S2... S5) to protect the identities of private partner organizations, with one site referring to all its reference and exposure areas. Each benthic station consisted of a composite of three sediment replicates with a Petite Ponar® dredge for a total sampled area of 0.0696 m². After collection, composite benthic samples from each replicate station were fixed with ethanol and stored at room temperature. Prior to sorting, each benthic sample was strained over a 500 µm steel strainer to remove silt and smaller organisms that are not used in standard EEM protocols. The collected invertebrates were sorted and identified to the lowest practical level based on morphology and imaged either with a Leica M205-A Z-stack microscope or by a Keyence VHX-7000 directly in a microplate done at the CBG (Centre for Biodiversity Genomics, University of Guelph). The CBG also performed the sequencing of specimens.

DNA extraction from individual invertebrates was conducted using a magnetic bead-based protocol (deWaard et al., 2008). Mollusc DNA was extracted using a CTAB (cetyltrimethylammonium bromide) protocol (Steinke et al., 2016). A standard primer pair, C_LepFolF + C_LepFolR (see Table 1 for details) was used to amplify the COI barcode region for all invertebrates except Mollusca, where several alternative primer cocktails were used (see Table 1 for details). The following thermocycling protocol was used for all primer combinations: initial denaturation for two minutes at 94°C, then five cycles of denaturation for 40 seconds at 94°C followed by annealing for 40 seconds at 45°C and extension for one minute at 72°C, 35 cycles of denaturation for 40 seconds at 94°C followed by annealing for 40 seconds at 51°C, extension for one minute at 72°C, and final extension for five minutes at 72°C. Sequencing of mollusc DNA was done using an ABI 3730xl DNA Analyzer (Applied Biosystems), while the remaining invertebrates were sequenced using a Pacific Biosciences Sequel II platform (Hebert et al., 2018). The resulting sequences were all uploaded to the Barcode of Life Data System (BOLD - Ratnasingham & Hebert, 2007) and were downloaded and filtered based on quality checks including discarding too-short sequences, detecting and correcting sequences with possible insertion/deletion errors or stop codons, and checking for sequence outliers in the dataset. The project containing the dataset used here is available on BOLD through the following doi: dx.doi.org/10.5883/DS-ECORG. The specimens in the barcoding dataset were assigned their taxonomy using BOLDigger (Buchner & Leese, 2020), sorting results by percent match and selecting the top hits. The dataset was then filtered for the presence of family-level identification.

Table 1

Primer details, including sequences and publications, for all of the primers used for PCR amplification of the barcode region of the COI gene in this study.

Original Target Taxa	Target Taxa for this Study	Primer Code	Sequences	Publication
Lepidoptera	Arthropoda, Annelida	C_LepFolF	LepF1: ATTCAACCAATCATAAAGATATTGG LCO1490: GGTCAACAAATCATAAAGATATTGG	Hebert et al. (2004) Folmer et al. (1994)
Lepidoptera	Arthropoda, Annelida	C_LepFolR	LepR1: TAAACTTCTGGATGTCCAAAAAATCA HCO2198: TAAACTTCAGGGTGACCAAAAAATCA	Hebert et al. (2004) Folmer et al. (1994)
Bivalves	Molluscs (Bivalves)	BivF4_t1	TGTAAAACGACGGCCAGTGKTCWAC WAATCATAARGATATTGG	Layton et al. (2014)
Bivalves	Molluscs (Bivalves)	BivR1_t1	CAGGAAACAGCTATGACTAMACCTC WGGRTGVCCRAARAACCA	Layton et al. (2014)
Gastropod	Gastropod	C_GasF1_t1	GasF1_t1: TGTAACGACGGCCAGTTTTCAACA AACCATAARGATATTGG GasF2_t1: TGTAACGACGGCCAGTATTCTACA AACCACAAAGACATCGG GasF3_t1: TGTAACGACGGCCAGTTTTCWACW AATCATAAAGATATTGG	Steinke et al. (2016)
Gastropod	Gastropod	GasR1_t1	CAGGAAACAGCTATGACACTTCWGG RTGHCCRAARAATCARAA	Stein et al. (2013)

All further analyses were performed in R (version 4.3.3, R Core Team, 2021), and code used for quality, cleaning, and analyses can be found at Dutheil (2025a) and Dutheil (2025b). The four EEM 'effect' endpoints were calculated for each of the two methods of taxonomic identification (morphology and

barcoding) and at two taxonomic levels (family and species-like units represented by BOLD's OTUs, Barcode Index Numbers (BINs, Ratnasingham & Hebert, 2013)). Endpoint values were calculated for each benthic station, meaning that one site would have two or three sampled areas (a reference and one to two exposures). In total, this dataset had 67 benthic stations across the 5 study sites.

Taxonomic richness was calculated as the total number of unique taxonomic groups (whether families or BINs) for each benthic subsample. Density (m^{-2}) was calculated using the number of individuals divided by the sampled area from the ponar grab ($0.0232 \text{ m}^2 \times 3$ grabs for a total sampled area of 0.0696 m^2) per benthic subsample. Simpson's Evenness Index follows the equation described by ECCC (ECCC, 2022), which takes the inverse Simpson's Evenness divided by the total number of taxa in a benthic subsample. In R, the diversity function from the vegan package (version 2.6–6.1, Oksanen et al., 2024) was used to calculate the inverse Simpson index, which was subsequently divided by the taxonomic richness. Bray-Curtis Dissimilarity also follows the ECCC method for EEM calculation (ECCC, 2022), which calculates the dissimilarity index based on species composition. The function vegdist from the vegan package was used in the calculation of the EEM Bray-Curtis index (Oksanen et al., 2024). The mean, standard error, standard deviation, median, minimum, and maximum values of all benthic stations within an area for each endpoint were also included. Effect sizes were calculated for each exposure area, denoted as: (exposure mean – reference mean) / reference standard deviation.

Significant differences between endpoint values of each exposure area and its corresponding reference area were calculated using either an ANOVA or, if various assumptions of the test could not be met, a Mann-Whitney test. First, the assumptions of an ANOVA test were checked using Shapiro's test (stats package version 4.3.3, R Core Team, 2021) for normal distribution and Levene's test (car package version 3.1-2, Fox et al., 2024) for homogeneity of variance. If the assumptions were met, an ANOVA test was done, and if not, then a non-parametric (Mann-Whitney) test was done (stats package, R Core Team, 2021), with one test per exposure area as some industrial locations have two exposure areas paired with a single reference area. For EEMs, a p-value less than 0.1 represents a significant difference between areas. To compare each identification method, the number of exposure areas deemed significantly different or not significantly different from their paired reference areas was tallied and used to assess each method's ability to determine statistical significance, indicating a potential impact.

Results

In total, 7,180 specimens were sorted from all five sites, including reference and effluent exposure areas, which were used for the morphology and DNA barcoding identification methods. From this total, and after filtering out specimens with no identifications at the family level, the morphology dataset comprised 7,153 specimens. It included 6,108 Arthropoda, 818 Mollusca, and 227 Annelida. Of the 7,180 specimens plated and sent for sequencing, 5,713 records (80% of the morphology dataset) had sequences, with an additional 555 flagged records including 436 contaminated sequences and 119 sequences with stop codons. After filtering for high-quality sequences and identifications at the family level, the barcoding dataset consisted of 4,674 specimens, with 4,204 Arthropoda, 379 Mollusca, and 91

Annelida. This dataset represents 82% of the records with sequences (5,713) and 65% of the total number of specimens collected and filtered (7,160). Post filtering, approximately 99% of specimens in the barcoding dataset had a percent match between 98–100% with the assigned taxonomy from BOLDigger, with the lowest percent match at 89.24%. The BIN dataset, which uses the same sequences from the barcoding dataset, has a total of 4,711 specimens, as we did not filter out specimens without family-level identification, and includes 389 BINs, 90 of which are unique within BOLD. Thus, there are two datasets resulting from the barcoding process, but they are each used at their respective taxonomic levels in this study (family-level and BIN-level).

In one of the locations (Site 1 or S1), three of the five sample stations contained no invertebrates in each of the two exposure areas for this location, meaning that six out of 10 benthic stations were empty. These stations were considered empty and included in the calculation of all the endpoints. However, benthic stations that did collect some specimens but whose DNA did not get successfully sequenced or was filtered out (resulting in no sequences for an entire benthic station) were not included in the dataset. Such cases were only present in S1, with two stations missing from exposure area 1 (EXP1) and one from exposure area 2 (EXP2). S1 is distinct from the other locations, as both exposure areas collected very few specimens, totalling three (EXP1) and four (EXP2) invertebrates, compared to 590 invertebrates in the corresponding reference area. None of the other locations showed such a large discrepancy between paired areas.

For the remaining datasets, paired areas allowed for comparing reference and exposure areas and for calculating effect sizes. Of the three datasets and four endpoints, only one set (family-level barcoding, Bray-Curtis dissimilarity) passed the assumptions for normal distribution and homogeneity of variance. All 11 other sets did not pass, so we applied the Mann-Whitney non-parametric tests on all datasets for consistency.

P-values of the non-parametric tests for the family-level morphology, family-level barcoding, and BIN-level barcoding are shown in Table 2 and raw values of the endpoints for each dataset are available in Appendix 1. In total, every effluent exposure area exhibited a significant difference between paired areas for at least one endpoint and one identification method. In a few cases, the significant difference determined by Simpson's evenness index was in favour of the exposure area, meaning that the exposure area had a higher value for this index. A key finding was that many of the determined significant differences for an area were similar between family-level morphology and family-level barcoding, with the BIN-level barcoding demonstrating similar but slightly differing conclusions. To illustrate the comparison between morphological and molecular methods for family-level identifications, the consensus between each method in determining a significant difference between exposure and reference areas is illustrated in Fig. 1. The figure highlights the percentage at which both methods determined the same effect (significant, $p < 0.1$ or non-significant, $p > 0.1$) across all locations and for each endpoint. All four endpoints demonstrated a consensus between the methods that is greater than 60%, for an average agreement between methods of 72%. Specifically, the taxonomic richness and Bray-Curtis endpoints had 100% consensus between the morphological and barcoding methods, meaning

they came to the same conclusion regarding statistical differences between paired reference and exposure areas at all locations.

Table 2

Mann-Whitney test p-values for all three datasets (BIN-level barcoding, family-level morphology, and family-level barcoding) for all four EEM endpoints when comparing exposure areas to reference areas. Areas are denoted as the exposure area for an effluent location, with some locations having up to two exposure areas paired with a single reference area. Values in bold designate a significant difference between the exposure area and its paired reference area. Values with * denote a p-value that is significantly different between paired areas but in the direction of the exposure area, meaning that greater values are found at the exposure area, and not the reference.

	Areas	Morphology (Family)	Barcoding (Family)	Barcoding (BIN)
Density (m ²)	S1-EXP1	0.0112	0.0325	0.0325
	S1-EXP2	0.0109	0.0179	0.0179
	S2-EXP1	1.0000	0.9166	0.9166
	S2-EXP2	0.1508	0.1732	0.1732
	S3-EXP	0.3939	0.6991	0.6991
	S4-EXP	0.6905	0.1508	0.1508
	S5-EXP1	0.0317	0.2222	0.2222
	S5-EXP2	0.0079	0.0079	0.0079
Taxonomic Richness	S1-EXP1	0.0107	0.0282	0.0325
	S1-EXP2	0.0107	0.0160	0.0179
	S2-EXP1	0.0117	0.0400	0.1732
	S2-EXP2	0.1693	0.7496	0.8340
	S3-EXP	1.0000	0.5669	0.8723
	S4-EXP	0.0290	0.0192	0.0212
	S5-EXP1	0.0109	0.0073	0.0119
	S5-EXP2	0.0847	0.0731	0.0079
Bray-Curtis Dissimilarity	S1-EXP1	NA	NA	NA
	S1-EXP2	NA	NA	NA
	S2-EXP1	0.3095	0.5556	0.0318
	S2-EXP2	0.3095	0.6905	0.0159
	S3-EXP	0.0152	0.0649	0.0022
	S4-EXP	0.0317	0.0556	0.0079

Note: NA values for S1-EXP1 and S1-EXP2 for some of the endpoints are due to a very low specimen abundance in the exposure areas for that site and calculations could not be made.

	Areas	Morphology (Family)	Barcoding (Family)	Barcoding (BIN)
	S5-EXP1	0.0952	0.0952	0.0079
	S5-EXP2	0.0079	0.0159	0.0079
Simpson's Evenness Index	S1-EXP1	NA	NA	NA
	S1-EXP2	NA	NA	NA
	S2-EXP1	0.0317*	0.2857	1.0000
	S2-EXP2	0.0317	0.3095	0.2222
	S3-EXP	0.8182	0.5887	0.4848
	S4-EXP	0.0317*	0.0318*	0.2222
	S5-EXP1	0.0079*	0.0119*	0.5476
	S5-EXP2	0.0317*	0.0318*	0.0952*
Note: NA values for S1-EXP1 and S1-EXP2 for some of the endpoints are due to a very low specimen abundance in the exposure areas for that site and calculations could not be made.				

The values for all four endpoints were also compared across all the different identification methods. Family-level morphological and barcoding identifications were compared for the density, richness, Bray-Curtis dissimilarity, and Simpson's evenness endpoints (Figure 2). Similar to Figure 1, there is a strong agreement between the values of the endpoints for both methods, with some exceptions shown for richness and density endpoints. To compare the BIN barcode values, family barcode, and family morphology datasets, density was removed as these values would be identical between both barcoding datasets, and taxonomic richness was removed as it would be greater for the BIN level as it represents a higher level of resolution. Thus, the three datasets were compared for the diversity endpoints Bray-Curtis dissimilarity and Simpson's evenness (Figure 3). Here, the disagreement of statistical significance of the BIN endpoints compared to both family-level datasets is pronounced.

Finally, effect sizes for the morphology and two barcoding datasets illustrate the directionality for all four endpoints (Figure 4). Directionality is shown by negative or positive values, as the formula of (exposure mean – reference mean) / reference standard deviation denotes higher values if the exposure mean > reference mean and negative values if exposure mean < reference mean.

Discussion

Our study determined whether morphological and molecular identifications were comparable in terms of assessing environmental impacts when used in biomonitoring studies such as an EEM. We evaluated if the endpoints calculated using both family-level morphology-based identifications and family-level molecular-based identifications would yield the same or similar results in the assessment of potential

effects from pulp mill or mine effluent on the downstream receiving area. Ultimately, our study shows that family-level morphology and family-level barcoding methods for identifying benthic invertebrates produce similar and, in many cases, identical results. Additionally, we explored whether BIN-level assessments of paired areas could provide greater detail on determining differences between reference and exposure areas. Compared to the morphological and molecular family-level assessments, BIN-level assessments did vary and consistently determined a significant difference between areas with the Bray-Curtis endpoint where family-level assessments did not.

When comparing the statistical significance between reference and exposure areas, both morphology and barcoding agree on whether an area exposed to effluent demonstrates effects over 70% of the time across all four endpoints (density, taxonomic richness, Bray-Curtis dissimilarity, and Simpson's evenness index). As shown in Table 2, when the morphological dataset identifies significant differences between paired areas, the molecular dataset often aligns. The opposite is also true as when one method determines there is no significant difference between areas, the other often does the same. This consensus between methods is further represented in Fig. 1, where the percentage of agreement between each method for a given endpoint across all exposure areas is shown. The taxonomic richness and Bray-Curtis endpoints have the highest consensus, at 100%, while density and Simpson's evenness follow behind at 87.5% and 66%, respectively. The high concordance of effect endpoint results shows the effectiveness of DNA barcoding to identify specimens and support biomonitoring studies and taxonomists using benthic invertebrates.

Furthermore, when comparing the range of values obtained for each of the four endpoints for family-level morphology and family-level barcoding (Fig. 2), there is much similarity between the two methods. This is especially true for the two diversity endpoints, Bray-Curtis dissimilarity and Simpson's evenness, as shown by the highly correlated boxplots for each location. There is less concordance between methods for the density and family richness endpoints, likely because these endpoints rely heavily on the size of the dataset for calculations, and the barcoding dataset represents 65% of the morphological one. Such a difference in datasets would impact both endpoints, explaining the differences seen visually, but as discussed above, both methods still yielded consistent results when determining impacts between paired areas (Fig. 1). Furthermore, there is concordance between both methods in terms of effect size values and directionality (Fig. 4) in which the family-level barcoding and family-level morphology values are very similar to each other.

Similar studies comparing morphological and barcoding methods at the same taxonomic level are scarce, as the comparison is often family or genus-level morphology compared to species or OTU level from barcoding (Sweeney et al., 2011; Hajibabaei et al., 2012; Macher et al., 2016). This is likely due to the appealing feature that molecular methods can provide a greater degree of taxonomic resolution and potentially greater sensitivity when classifying impacts for bioassessment. However, in a related study, Shackleton et al. (2021) compared bulk tissue metabarcoding to morphology at the same taxonomic level when determining river health indices. They found that the results of both methods yielded comparable characterizations of stream health when calculated at the family and genus levels. When

classifying streams for water quality, the molecular scores had high congruence with the morphological scores, and when they deviated, they classified streams to an adjacent score. Despite using metabarcoding methods, Shackleton et al. (2021) demonstrate that tissue-based DNA sampling can yield similar results to morphological results when classifying streams. These findings are reflected in this study, as DNA barcoding produced similar classifications of impacted and non-impacted streams when compared to morphology.

In addition to comparing morphological and molecular methods for specimen identification at the family level, this study also took advantage of the molecular methods' ability to obtain identifications at higher taxonomic resolution such as at the OTU or, in this case, BIN level. The goal was to determine whether BIN-level identification of invertebrates would provide greater sensitivity when classifying reference and exposure areas as significantly different from one another, as family-level identification can often mask more varied responses at the genus, species, or BIN level (King & Richardson, 2002; Carew et al., 2011; Macher et al., 2016). When comparing the significance of the results from the Mann-Whitney tests (Table 2), BIN-level results disagree with those obtained from the family-level morphological dataset over 30% of the time on average and 19% of the time from the family-level barcoding dataset, or in eight and five instances across all four endpoints, respectively. For six and three of these differences, the BIN dataset classified a difference as being non-significant whereas the family-level datasets classified it as a significant difference. Of these, four and two of the instances occurred for the Simpson's evenness endpoint, where most of the significant differences observed were in favour of the exposure area having greater values. The remaining two areas diverging from each family-level dataset were for the Bray-Curtis endpoint and represented instances where the BIN dataset assigned a significant difference between areas where the other datasets did not. It is likely that these divergences from the morphology results are due to the greater level of variation present at lower taxonomic levels compared to family-level assessments. Of all the endpoints, BIN-level changes at the Bray-Curtis endpoint are significant, as BINs represent a greater ability to distinguish communities between paired areas, an ability that is only present because of the greater degree of taxonomic resolution. Bray-Curtis is the only endpoint that considers the identities of specimens, and so it makes sense that this endpoint is the one where BIN-level assessments outline a significant difference between paired areas where family-level assessments do not.

However, it should be noted that the Bray-Curtis endpoint is the most sensitive to natural habitat differences compared to the other endpoints, which could factor into the differences seen when identifying specimens with higher resolution. This endpoint has been problematic in EEM evaluations and Environment Canada issued an evaluation on the way it was calculated and developed a new method (Borcard and Legendre, 2013). However, of the four effect endpoints, this endpoint still has the highest propensity to lead to spurious conclusions regarding potential impacts due to two main reasons, 1) the control-impact design used in the vast majority of EEM studies suffers from pseudo replication (Huebert et al., 2011) and, 2) the benthic invertebrate community composition of two reference areas with the same habitat is likely to have a different community composition (Bailey et al., 2004). Thus, it is not surprising that the Bray-Curtis endpoint calculated with BIN level assignment resulted in instances

demonstrating a statistical difference (Table 2). For this reason, any significant difference regarding Bray-Curtis dissimilarity should be further evaluated to determine if the actual benthic invertebrate community present is comprised of species that represent a negative impact (ECCC 2010).

The values obtained for the Bray-Curtis dissimilarity and Simpson's evenness endpoints when compared across BIN, family-level morphology, and family-level barcoding reveal similar divergences of the BIN results. While Fig. 2 illustrates where morphology and barcoding yield very similar values at the family level, Fig. 3 adds the BIN values which are more varied from the other two datasets. This variation is especially clear for the Bray-Curtis endpoint, as nearly all the ranges of the BIN values are outside of the ranges for the morphology and barcoding values (within an area), often representing greater dissimilarity of results. This variation supports the idea that BIN-level assessments provide greater resolution on benthic communities and can help distinguish areas that may be indistinguishable at higher taxonomic levels. Additionally, a similar variance is reflected in the BIN-level effect sizes (Fig. 4) when compared to family-level barcoding and family-level morphology, while there are several instances where the effect sizes are all similar, there are also cases where the BIN-level effect size varies from the other two methods.

Overall, the BIN-level results do not agree as closely with the other two datasets as do the family-level morphology and family-level barcoding to each other. Several other studies have compared family or genus-level morphological identifications to OTU or species-level barcoding identifications and have found that the barcoding assessments of stream health have resulted in differing results regarding the characterization of stream quality than what was found using morphology (Sweeney et al., 2011; Gill et al., 2014; Jackson et al., 2014; Stein et al., 2014). Sweeney et al. (2011) compared 17 metrics at the family (morphology), genus (morphology), genus/species (morphology), and OTU (barcoding) levels, including taxonomic richness for several categories such as for Chironomidae (non-biting midges), Ephemeroptera (mayflies) and others, as well as several biotic indices including Shannon's diversity and Simpson's evenness. They found that the results of these metrics changed significantly with increasing taxonomic resolution, which resulted in an increased ability to detect differences between areas, going from 36% of metrics being significantly different at the family level to 76% at the OTU barcode level (Sweeney et al., 2011). It has also been suggested that biomonitoring using morphology alone could be underestimating the biodiversity of benthic invertebrates by not identifying specimens to finer taxonomic levels through barcoding (Jackson et al., 2014). Gill et al. (2014) also compared morphology to OTU-level DNA barcoding and found that taxonomic richness and community similarities between areas were significantly different, leading to an increased ability for DNA barcoding to detect variability in studied areas. Thus, there is agreement that DNA barcoding and measures of stream health at the OTU level offer not only different results from measures using morphology but can provide greater sensitivity to differences between areas. It is important to note that the current study evaluated whether the Bray-Curtis dissimilarity endpoint was statistically different, and not what the species differences were between reference and exposure areas.

Our study is not without limitations, as we experienced moderate sequencing success for the barcoding methods, which represent 65% of the morphological dataset. A variety of field and laboratory factors influenced the moderate sequencing success including sample storage and time spent until sequencing. Sample storage and preservation are important for molecular work, as temperature, time, and ethanol concentration can all influence DNA preservation. It is generally recommended to store samples in -80°C or -20°C freezers when not in use for long periods as DNA degrades faster at warmer temperatures (Post et al., 1993; Vink et al., 2005; Fowler et al., 2024). In addition, time can also play a role in DNA degradation, where the longer time a sample spends in storage results in a decrease in DNA yield (Post et al., 1993). Furthermore, the samples in this study were stored with 80% ethanol to preserve the physical specimens, but higher ethanol concentrations (95%) are shown to have an improved ability for DNA preservation (Marquina et al., 2021), possibly further contributing to the moderate success of the barcoding approach. With improved and consistent DNA storage protocols, such as controlled temperatures and shorter time from sampling to processing, the barcoding approach could have a larger representation of the morphological dataset, allowing for a more complete comparison between specimen identification methods. The molluscs specifically have low sequencing success, and improvements in this regard could help bridge some gaps between the datasets.

Of the 389 unique BINs, approximately 40% of them are singletons, appearing only once in the dataset, which suggests incomplete sampling of the benthic invertebrate communities across five sites. The high proportion of singletons can influence results at the BIN level, particularly for Bray-Curtis, since singletons appearing only in one area (reference or exposure) can increase the dissimilarity values. BIN richness is also impacted as the singletons inflate the overall richness of a site, but paired comparisons in this study are less biased because singletons are present in both reference and exposure areas. The effect of singletons on density and Simpson's Index is minimal as these metrics are less sensitive to rare taxa. In future work, increased sampling or the filtering of singletons from a larger dataset than shown here could help reduce this bias.

A more encompassing molecular dataset may eliminate much of the variation seen between the endpoints relying on quantities of individuals, such as density and taxonomic richness. Some recovery of sequences could be possible, as filtering for the highest percent matches at the family level for the barcoding dataset, rather than the highest percent match overall, could recover up to 1% of the sequences, partially bridging the gap between the family-level and BIN-level barcoding datasets. It is also important to note that, while molecular methods tend to provide more accurate taxonomic identifications, they rely heavily on the completeness of reference libraries, which still require significant expansion, as seen with the 90 new BINs found in this dataset (Curry et al., 2018; Weigand et al., 2019). However, significant efforts have been made to increase the reference library for macroinvertebrates in Canada (Bush et al., 2019). This study also only includes five effluent locations and is not representative of effects that may be observed outside of the geographical range of this study. Future work with a greater number of industrial and reference locations and geographic spread would help expand upon the findings of this study.

Overall, barcoding methods could be an alternative to morphological-based identifications of invertebrates in EEM studies, as the calculation of the four 'effect' endpoints and the comparison of paired areas reveal very similar results at the family level. Barcoding methods offer the ability to consistently identify benthic invertebrates to the species level resulting in a higher taxonomic resolution than what can reasonably be achieved through morphology (Sweeney et al., 2011; Orlofske & Baird, 2013). Barcoding has also resulted in similar classifications of impacted areas when comparing at the same taxonomic level as morphology (Shackleton et al., 2021). While one of the appeals of molecular methods is the decrease in cost and time to analysis (Gibson et al., 2015; Hulley et al., 2018), this isn't always true for barcoding methods. The cost to barcode each benthic invertebrate was greater than traditional taxonomy in this study, as both methods require organisms to be sorted from the detritus, which is followed by either identifications by an experienced taxonomist or the barcoding protocol. Stein et al. (2014) compared DNA barcoding and traditional morphology and found barcoding to be 1.7–3.4 times more expensive using Sanger sequencing. However, they did find that the overall time from sampling to results was significantly reduced compared to morphology. In this study, SEQUEL sequencing was mainly used instead of Sanger due to lower associated costs. Future research could include a more thorough analysis of time and cost to results using barcoding methods. In addition, Stein et al. (2014) also compared next-generation sequencing (NGS), such as with bulk-tissue metabarcoding, to the two other methods, and found that this method was on par with morphology in terms of cost and substantially reduced the time required for results. NGS approaches using bulk tissues could provide a solution to the costly process of barcoding and the lengthy process of sorting and morphological identification, as it has been shown to provide a greater degree of sensitivity to differences between areas (Gibson et al., 2015), similar to OTU-level barcoding methods (Gill et al., 2014; Jackson et al., 2014). Additionally, bulk-tissue metabarcoding has proven to yield comparable results to traditional morphological methods for calculating river health biotic indices when used at the same taxonomic level (Shackleton et al., 2021), further indicating the potential for this method to be used in EEM studies of benthic invertebrates. There is potential for eDNA metabarcoding work as well, though there is little research comparing morphological, barcoding, and metabarcoding results in the context of regulated biomonitoring studies.

BIN or OTU-level analyses of biomonitoring studies have the potential to provide a higher resolution for describing impacts on benthic invertebrate communities. In this study, we demonstrated that while family-level impacts on areas were similar between morphology and barcoding, BIN-level impacts differed, especially from morphology. BIN level (or OTU, or species) assessments of invertebrate communities may be a more accurate representation of the impacts measured in biomonitoring, particularly if the study design differs from the pseudo replication design conducted for most EEM studies. It is common to observe varying disturbance tolerances of invertebrates within a unique genus or family (Resh & Unzicker, 1975; Lenat & Resh, 2001), indicating that not all species within these taxa respond the same way to a disturbance such as industrial effluent. The concern is that family-level identification of invertebrates masks more varied responses at the species or BIN level, resulting in an inaccurate representation of changes in a community facing disturbance. Here, we demonstrate that

while there is some agreement with BIN-level results to family-level results, there are differences as well. This effect was especially true for the endpoint that considers specimen identities, Bray-Curtis dissimilarity, where having greater taxonomic resolution is most impactful. It is important that biomonitoring studies, represent an accurate description of invertebrate communities to ensure that appropriate measures are taken when an impact is determined so that further determination of causes can be undertaken and monitored.

Conclusions

Family-level identifications with the morphological and DNA barcoding methods prove to draw similar conclusions regarding both the results of the four EEM endpoints and the determination of whether an exposure area is significantly different from its paired reference area. Despite the lower sequencing success rate, the DNA barcoding dataset replicated results found by the morphological dataset over 70% of the time. With the inclusion of molecular methods in these EEM studies, we were able to obtain species-proxy levels of taxonomic resolution through BINs and determine that there is more variance in the endpoints at the BIN level than is present at the family level for both methods. In certain cases, conclusions drawn by the family-level morphology and barcoding differed from the conclusions of the BIN-level dataset. BIN-level taxonomic resolution alleviates the various issues that morphological taxonomist face, which limits identifications to the genus or species level. Molecular barcoding techniques that provide consistent and accurate identifications of species will enhance impact assessment. Overall, DNA barcoding demonstrated similar results to morphological-based identifications at the family level of identification, justifying its use as a supplement or substitute for morphological methods. While barcoding decreases the time from sample processing to results, the use of sequencing platforms such as SEQUEL (rather than Sanger methods), and other methods such as bulk-tissue metabarcoding using next-generation sequencing will further reduce barcoding costs.

Declarations

All authors have read, understood, and have complied as applicable with the statement on “Ethical responsibilities of Authors” as found in the Instructions for Authors. The authors declare no competing interests.

Funding

for this study was provided by Ecoreg Solutions, Ontario Genomics, and Natural Sciences and Engineering Research Council Discovery Grant. The authors would like to thank our colleagues at Ecoreg Solutions, and in particular Ian Thompson, Bill Morton and many others for their contributions to the sampling, reviewing, and handling of data for this project. We are also grateful for the support of our Industry Partners.

Author Contribution

SJM, HS, and LMD conceived the original idea, and LD wrote the manuscript with input from all authors (SJM, KC, DS, HS). Sampling and data collection were led by HS, and data were analyzed by LD with support from KC, SJM, and DS.

Acknowledgement

The authors would like to thank our colleagues at Ecoreg Solutions, in particular Ian Thompson, Bill Morton and many others, for their contributions to the sampling, reviewing, and handling of data for this project. We are also grateful for the support of our Industry Partners.

Data Availability

All data supporting the findings of this study are available within the paper: R pipelines are available in public repositories that were cited in the paper (Dutheil, 2025a and 2025b) and the data generated by BOLD and used for this study are shared through a public doi ([dx.doi.org/10.5883/DS-ECORG](https://doi.org/10.5883/DS-ECORG)) included in the body of the paper.

References

1. Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. Bioassessment of Freshwater Ecosystems: Using the Reference Condition Approach. Kluwer Academic Publishers, Massachusetts, USA.
2. Borcard, D and P. Legendre. 2013. Review of the pros and cons of available indices applicable to the Environmental Effects Monitoring (EEM) to evaluate changes in benthic invertebrate community structure in response to effluent exposure. Project # K2A80-12-0010
3. Brantschen, J., Blackman, R. C., Walser, J.-C., & Altermatt, F. (2021). Environmental DNA gives comparable results to morphology-based indices of macroinvertebrates in a large-scale ecological assessment. *PLOS ONE*, 16(9), e0257510. <https://doi.org/10.1371/journal.pone.0257510>
4. Buchner, D., & Leese, F. (2020). BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics*, 4, e53535. <https://doi.org/10.3897/mbmg.4.53535>
5. Bush, A., Compson, Z. G., Monk, W. A., Porter, T. M., Steeves, R., Emilson, E., Gagne, N., Hajibabaei, M., Roy, M., & Baird, D. J. (2019). Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates. *Frontiers in Ecology and Evolution*, 7. <https://doi.org/10.3389/fevo.2019.00434>
6. Carew, M. E., Miller, A. D., & Hoffmann, A. A. (2011). Phylogenetic signals and ecotoxicological responses: Potential implications for aquatic biomonitoring. *Ecotoxicology*, 20(3), 595–606. <https://doi.org/10.1007/s10646-011-0615-3>

7. Carew, M. E., Pettigrove, V. J., Metzeling, L., & Hoffmann, A. A. (2013). Environmental monitoring using next generation sequencing: Rapid identification of macroinvertebrate bioindicator species. *Frontiers in Zoology*, 10(1), 45. <https://doi.org/10.1186/1742-9994-10-45>
8. Clarke, R. T., Lorenz, A., Sandin, L., Schmidt-Kloiber, A., Strackbein, J., Kneebone, N. T., & Haase, P. (2006). Effects of sampling and sub-sampling variation using the STAR-AQEM sampling protocol on the precision of macroinvertebrate metrics. In M. T. Furse, D. Hering, K. Brabec, A. Buffagni, L. Sandin, & P. F. M. Verdonschot (Eds.), *The Ecological Status of European Rivers: Evaluation and Intercalibration of Assessment Methods* (pp. 441–459). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5493-8_31
9. Curry, C. J., Gibson, J. F., Shokralla, S., Hajibabaei, M., & Baird, D. J. (2018). Identifying North American freshwater invertebrates using DNA barcodes: Are existing COI sequence libraries fit for purpose? *Freshwater Science*, 37(1), 178–189. <https://doi.org/10.1086/696613>
10. deWaard, J. R., Ivanova, N. V., Hajibabaei, M., & Hebert, P. D. N. (2008). Assembling DNA Barcodes. In C. C. Martin & C. C. Martin (Eds.), *Environmental Genomics* (pp. 275–294). Humana Press. https://doi.org/10.1007/978-1-59745-548-0_15
11. Doberstein, C. P., Karr, J. R., & Conquest, L. L. (2000). The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams. *Freshwater Biology*, 44(2), 355–371. <https://doi.org/10.1046/j.1365-2427.2000.00575.x>
12. Dudka, S., & Adriano, D. C. (1997). Environmental Impacts of Metal Ore Mining and Processing: A Review. *Journal of Environmental Quality*, 26(3), 590–602. <https://doi.org/10.2134/jeq1997.00472425002600030003x>
13. Dutheil, L. (2025a). *Laura-Dutheil/Endpoint-Analysis* [R]. <https://github.com/Laura-Dutheil/Endpoint-Analysis> (Original work published 2025)
14. Dutheil, L. (2025b). *Quality-Checking-Sequences* [R]. <https://github.com/Laura-Dutheil/Quality-Checking-Sequences> (Original work published 2025)
15. Environment and Climate Change. (2022, July 31). 4. *Effects on Fish Habitat: Benthic Invertebrate Community Survey* [Regulations]. <https://www.canada.ca/en/environment-climate-change/services/managing-pollution/environmental-effects-monitoring/metal-mining-technical-guidance/metal-mining-technical-guidance-environmental-effects-monitoring/chapter-4.html>
16. Environment and Climate Change Canada. (2008, March 13). *Environmental effects monitoring*. <https://www.canada.ca/en/environment-climate-change/services/managing-pollution/environmental-effects-monitoring.html>
17. Environment and Climate Change Canada. (2012). *Metal Mining Technical Guidance for Environmental Effects Monitoring*. [https://www.ec.gc.ca/esee-eem/AEC7C481-D66F-4B9B-BA08-A5DC960CDE5E/COM-1434---Tec-Guide-for-Metal-Mining-Env-Effects-Monitoring_En_02\[1\].pdf](https://www.ec.gc.ca/esee-eem/AEC7C481-D66F-4B9B-BA08-A5DC960CDE5E/COM-1434---Tec-Guide-for-Metal-Mining-Env-Effects-Monitoring_En_02[1].pdf)
18. Environment and Climate Change Canada. (2010). *Pulp and Paper Environmental Effects Monitoring (EEM) Technical Guidance Document*. https://www.ec.gc.ca/esee-eem/3E389BD4-E48E-4301-A740-171C7A887EE9/PP_full_versionENGLISH%5B1%5D-FINAL-2.0.pdf

19. Fernández, S., Rodríguez, S., Martínez, J. L., Borrell, Y. J., Ardura, A., & García-Vázquez, E. (2018). Evaluating freshwater macroinvertebrates from eDNA metabarcoding: A river Nalón case study. *PLOS ONE*, 13(8), e0201741. <https://doi.org/10.1371/journal.pone.0201741>
20. Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), 294–299.
21. Fowler, E. V., Starkie, M. L., Blacket, M. J., Mayer, D. G., & Schutze, M. K. (2024). Effect of temperature and humidity on insect DNA integrity evaluated by real-time PCR. *Journal of Economic Entomology*, toae193. <https://doi.org/10.1093/jee/toae193>
22. Fox, J., Weisberg, S., Price, B., Adler, D., Bates, D., Baud-Bovy, G., Bolker, B., Ellison, S., Firth, D., Friendly, M., Gorjanc, G., Graves, S., Heiberger, R., Krivitsky, P., Laboissiere, R., Maechler, M., Monette, G., Murdoch, D., Nilsson, H., ... R-Core. (2024). *car: Companion to Applied Regression* (Version 3.1-3) [Computer software]. <https://cran.r-project.org/web/packages/car/index.html>
23. Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. *PLOS ONE*, 10(10), e0138432. <https://doi.org/10.1371/journal.pone.0138432>
24. Gill, B. A., Harrington, R. A., Kondratieff, B. C., Zamudio, K. R., LeRoy Poff, N., & Chris Funk, W. (2014). Morphological taxonomy, DNA barcoding, and species diversity in southern Rocky Mountain headwater streams. *Freshwater Science*, 33(1), 288–301. <https://doi.org/10.1086/674526>
25. Government of Canada. (2002). *Metal and Diamond Mining Effluent Regulations*. <https://laws-lois.justice.gc.ca/PDF/SOR-2002-222.pdf>
26. Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology*, 12(1), 28. <https://doi.org/10.1186/1472-6785-12-28>
27. Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., Janzen, D. H., Hallwachs, W., Naik, S., Sones, J. E., & Zakharov, E. V. (2018). A Sequel to Sanger: Amplicon sequencing that scales. *BMC Genomics*, 19, 219. <https://doi.org/10.1186/s12864-018-4611-3>
28. Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings. Biological Sciences*, 270(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218>
29. Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences*, 101(41), 14812–14817. <https://doi.org/10.1073/pnas.0406166101>
30. Huebert, D.B., M. McKernan and C. Paquette. 2011. Canada's Environmental Effects Monitoring Program: Areas for Improvement. *Integrated Environmental Assessment and Management*, 7:143–

31. Hewlett, R. (2000). Implications of taxonomic resolution and sample habitat for stream classification at a broad geographic scale. *Journal of the North American Benthological Society*, 19(2), 352–361. <https://doi.org/10.2307/1468077>
32. Hulley, E. N., Taylor, N. D. J., Zarnke, A. M., Somers, C. M., Manzon, R. G., Wilson, J. Y., & Boreham, D. R. (2018). DNA barcoding vs. morphological identification of larval fish and embryos in Lake Huron: Advantages to a molecular approach. *Journal of Great Lakes Research*, 44(5), 1110–1116. <https://doi.org/10.1016/j.jglr.2018.07.013>
33. Jackson, J. K., Battle, J. M., White, B. P., Pilgrim, E. M., Stein, E. D., Miller, P. E., & Sweeney, B. W. (2014). Cryptic biodiversity in streams: A comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science*, 33(1), 312–324. <https://doi.org/10.1086/675225>
34. King, R. S., & Richardson, C. J. (2002). Evaluating Subsampling Approaches and Macroinvertebrate Taxonomic Resolution for Wetland Bioassessment. *Journal of the North American Benthological Society*, 21(1), 150–171. <https://doi.org/10.2307/1468306>
35. Layton, K. K. S., Martel, A. L., & Hebert, P. D. (2014). Patterns of DNA Barcode Variation in Canadian Marine Molluscs. *PLOS ONE*, 9(4), e95003. <https://doi.org/10.1371/journal.pone.0095003>
36. Legislative Services Branch. (2018, September 26). *Consolidated federal laws of canada, Pulp and Paper Effluent Regulations*. <https://laws-lois.justice.gc.ca/eng/regulations/SOR-92-269/?showtoc=&instrumentnumber=SOR-92-269>
37. Legislative Services Branch. (2024, June 10). *Consolidated federal laws of Canada, Metal and Diamond Mining Effluent Regulations*. <https://laws-lois.justice.gc.ca/eng/regulations/SOR-2002-222/index.html>
38. Lenat, D. R., & Resh, V. H. (2001). Taxonomy and stream ecology—The benefits of genus- and species-level identifications. *Journal of the North American Benthological Society*, 20(2), 287–298. <https://doi.org/10.2307/1468323>
39. Losey, J. E., & Vaughan, M. (2006). The Economic Value of Ecological Services Provided by Insects. *BioScience*, 56(4), 311–323. [https://doi.org/10.1641/0006-3568\(2006\)56\[311:TEVOES\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)56[311:TEVOES]2.0.CO;2)
40. Macadam, C. R., & Stockan, J. A. (2015). More than just fish food: Ecosystem services provided by freshwater insects. *Ecological Entomology*, 40(S1), 113–123. <https://doi.org/10.1111/een.12245>
41. Macher, J. N., Salis, R. K., Blakemore, K. S., Tollrian, R., Matthaei, C. D., & Leese, F. (2016). Multiple-stressor effects on stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, 61, 159–169. <https://doi.org/10.1016/j.ecolind.2015.08.024>
42. Maret, T. R., Cain, D. J., MacCoy, D. E., & Short, T. M. (2003). Response of benthic invertebrate assemblages to metal exposure and bioaccumulation associated with hard-rock mining in northwestern streams, USA. *Journal of the North American Benthological Society*, 22(4), 598–620. <https://doi.org/10.2307/1468356>

43. Marquina, D., Buczek, M., Ronquist, F., & Łukasik, P. (2021). The effect of ethanol concentration on the morphological and molecular preservation of insects for biodiversity studies. *PeerJ*, 9, e10799. <https://doi.org/10.7717/peerj.10799>
44. Martins, F. M. S., Feio, M. J., Porto, M., Filipe, A. F., Bonin, A., Serra, S. R. Q., Alves, P. C., Taberlet, P., & Beja, P. (2021). Assessing changes in stream macroinvertebrate communities across ecological gradients using morphological versus DNA metabarcoding approaches. *Science of The Total Environment*, 797, 149030. <https://doi.org/10.1016/j.scitotenv.2021.149030>
45. Mebane, C. A., Eakins, R. J., Fraser, B. G., & Adams, W. J. (2015). Recovery of a mining-damaged stream ecosystem. *Elementa: Science of the Anthropocene*, 3, 000042. <https://doi.org/10.12952/journal.elementa.000042>
46. Mocq, J., & Hare, L. (2018). Influence of Acid Mine Drainage, and Its Remediation, on Lakewater Quality and Benthic Invertebrate Communities. *Water, Air, & Soil Pollution*, 229(2), 28. <https://doi.org/10.1007/s11270-017-3671-3>
47. Molineri, C., Tejerina, E. G., Torrejón, S. E., Pero, E. J. I., & Hankel, G. E. (2020). Indicative value of different taxonomic levels of Chironomidae for assessing the water quality. *Ecological Indicators*, 108, 105703. <https://doi.org/10.1016/j.ecolind.2019.105703>
48. Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., Caceres, M. D., Durand, S., ... Weedon, J. (2024). *vegan: Community Ecology Package* (Version 2.6-8) [Computer software]. <https://cran.r-project.org/web/packages/vegan/index.html>
49. Orlofske, J. M., & Baird, D. J. (2013). The tiny mayfly in the room: Implications of size-dependent invertebrate taxonomic identification for biomonitoring data properties. *Aquatic Ecology*, 47(4), 481–494. <https://doi.org/10.1007/s10452-013-9460-1>
50. Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéoz-Perret-Gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Feio, M. J., Filipe, A. F., Fornaroli, R., Graf, W., Herder, J., van der Hoorn, B., Iwan Jones, J., Sagova-Mareckova, M., Moritz, C., ... Kahlert, M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of The Total Environment*, 637–638, 1295–1310. <https://doi.org/10.1016/j.scitotenv.2018.05.002>
51. Petkovska, V., & Urbanič, G. (2010). Effect of fixed-fraction subsampling on macroinvertebrate bioassessment of rivers. *Environmental Monitoring and Assessment*, 169(1), 179–201. <https://doi.org/10.1007/s10661-009-1161-9>
52. Post, R. J., Flook, P. K., & Millest, A. L. (1993). Methods for the preservation of insects for DNA studies. *Biochemical Systematics and Ecology*, 21(1), 85–92. [https://doi.org/10.1016/0305-1978\(93\)90012-G](https://doi.org/10.1016/0305-1978(93)90012-G)
53. R Core Team. (2021). *R: A language and environment for statistical computing*. [Computer software]. R Foundation for Statistical Computing. <https://www.r-project.org/>

54. Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes*, 7(3), 355–364.
<https://doi.org/10.1111/j.1471-8286.2007.01678.x>
55. Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE*, 8(7), e66213.
<https://doi.org/10.1371/journal.pone.0066213>
56. Resh, V. H., & Unzicker, J. D. (1975). Water Quality Monitoring and Aquatic Organisms: The Importance of Species Identification. *Journal (Water Pollution Control Federation)*, 47(1), 9–19.
57. Shackleton, M. E., Dafforn, K. A., Murphy, N. P., Greenfield, P., Cassidy, M., & Besley, C. H. (2021). How does molecular taxonomy for deriving river health indices correlate with traditional morphological taxonomy? *Ecological Indicators*, 125, 107537. <https://doi.org/10.1016/j.ecolind.2021.107537>
58. Stein, E. D., Martinez, M. C., Stiles, S., Miller, P. E., & Zakharov, E. V. (2014). Is DNA Barcoding Actually Cheaper and Faster than Traditional Morphological Methods: Results from a Survey of Freshwater Bioassessment Efforts in the United States? *PLOS ONE*, 9(4), e95525.
<https://doi.org/10.1371/journal.pone.0095525>
59. Steinke, D., Prosser, S. W. J., & Hebert, P. D. N. (2016). DNA Barcoding of Marine Metazoans. In S. J. Bourlat (Ed.), *Marine Genomics: Methods and Protocols* (pp. 155–168). Springer.
https://doi.org/10.1007/978-1-4939-3774-5_10
60. Sweeney, B. W., Battle, J. M., Jackson, J. K., & Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? *Journal of the North American Benthological Society*, 30(1), 195–216. <https://doi.org/10.1899/10-016.1>
61. Uchida, N., Kubota, K., Aita, S., & Kazama, S. (2020). Aquatic insect community structure revealed by eDNA metabarcoding derives indices for environmental assessment. *PeerJ*, 8, e9176.
<https://doi.org/10.7717/peerj.9176>
62. Vink, C. J., Thomas, S. M., Paquin, P., Hayashi, C. Y., & Hedin, M. (2005). The effects of preservatives and temperatures on arachnid DNA. *Invertebrate Systematics*, 19(2), 99–104.
<https://doi.org/10.1071/IS04039>
63. Vlek, H. E., Šporka, F., & Krno, I. (2006). Influence of macroinvertebrate sample size on bioassessment of streams. In M. T. Furse, D. Hering, K. Brabec, A. Buffagni, L. Sandin, & P. F. M. Verdonschot (Eds.), *The Ecological Status of European Rivers: Evaluation and Intercalibration of Assessment Methods* (pp. 523–542). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5493-8_35
64. Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zaťovičová, Z., Ferreira, S., ... Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of The Total Environment*, 678, 499–524. <https://doi.org/10.1016/j.scitotenv.2019.04.247>

Figures

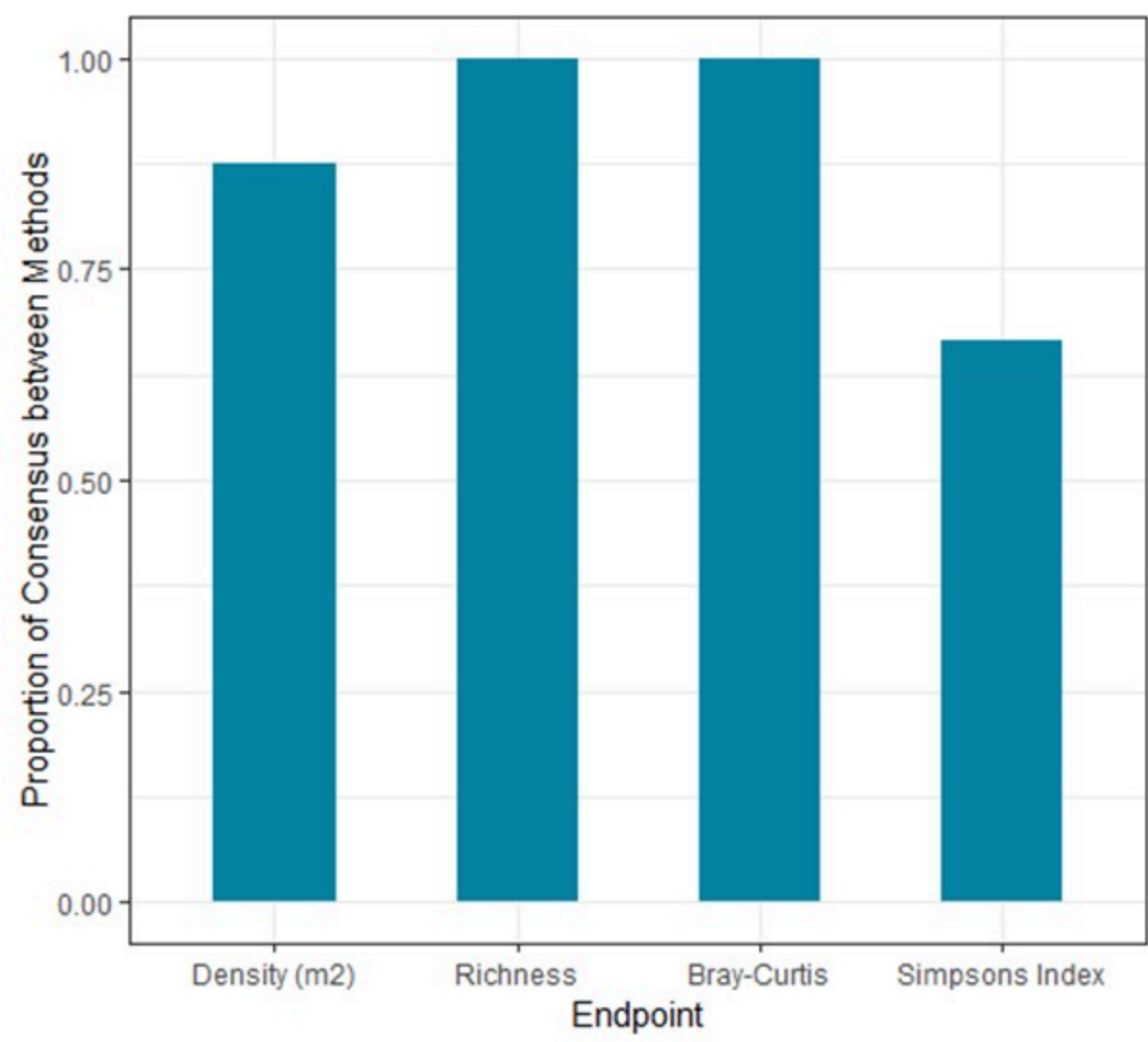


Figure 1

Proportion of consensus between family-level morphological and family-level barcoding determinations of impact at a location across all four endpoints. When both methods make the same conclusion regarding a paired area (either significant difference or no significant difference), this is considered a consensus between methods. Represented here is the percent of consensus occurrences for each endpoint.

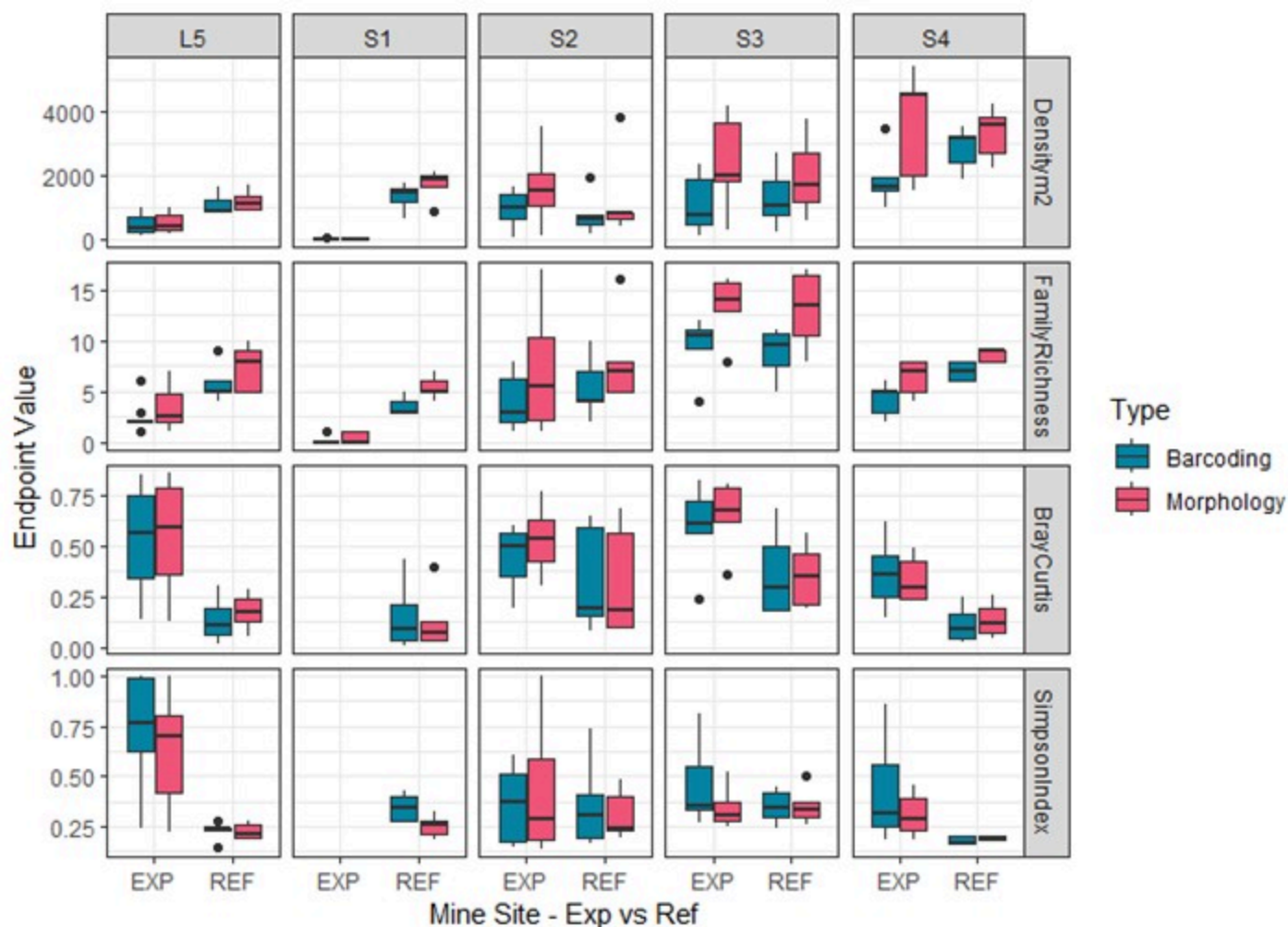


Figure 2

Comparison of family-level barcoding and family-level morphology results for all four endpoints and all five sites. When a site has two exposure areas (such as with S1, S2, and S5), the values from both exposure areas were summarized together to represent a single exposure area for visual simplicity. Note that six benthic stations of the two exposure areas in S1 found no invertebrates, resulting in empty values for the Bray-Curtis and Simpson's evenness endpoints.

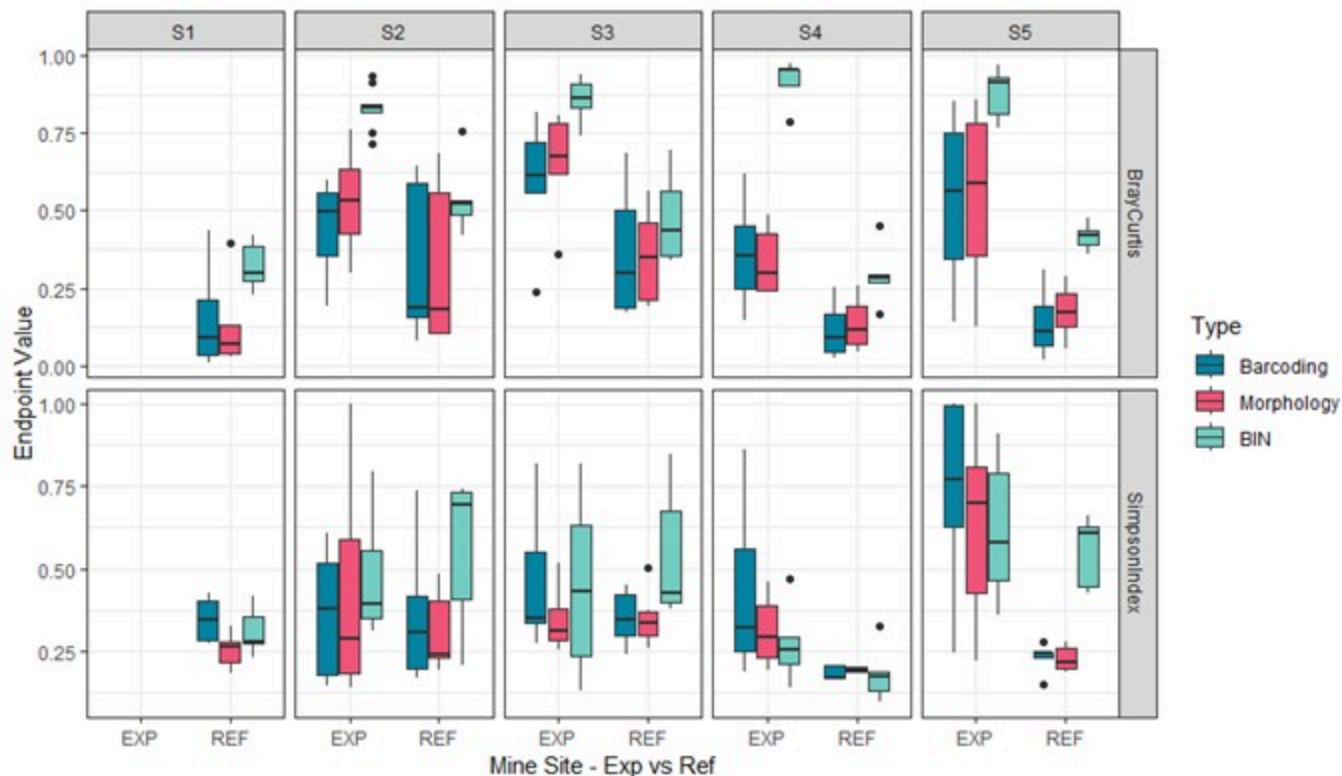


Figure 3

Comparison of the family-level barcoding, family-level morphology, and BIN-level barcoding results for the diversity endpoints Bray-Curtis dissimilarity and Simpson's evenness index for all five sites. Note: no data for the S1 exposure sites is due to very low specimen abundance at that site's exposure areas and calculations could not be made.

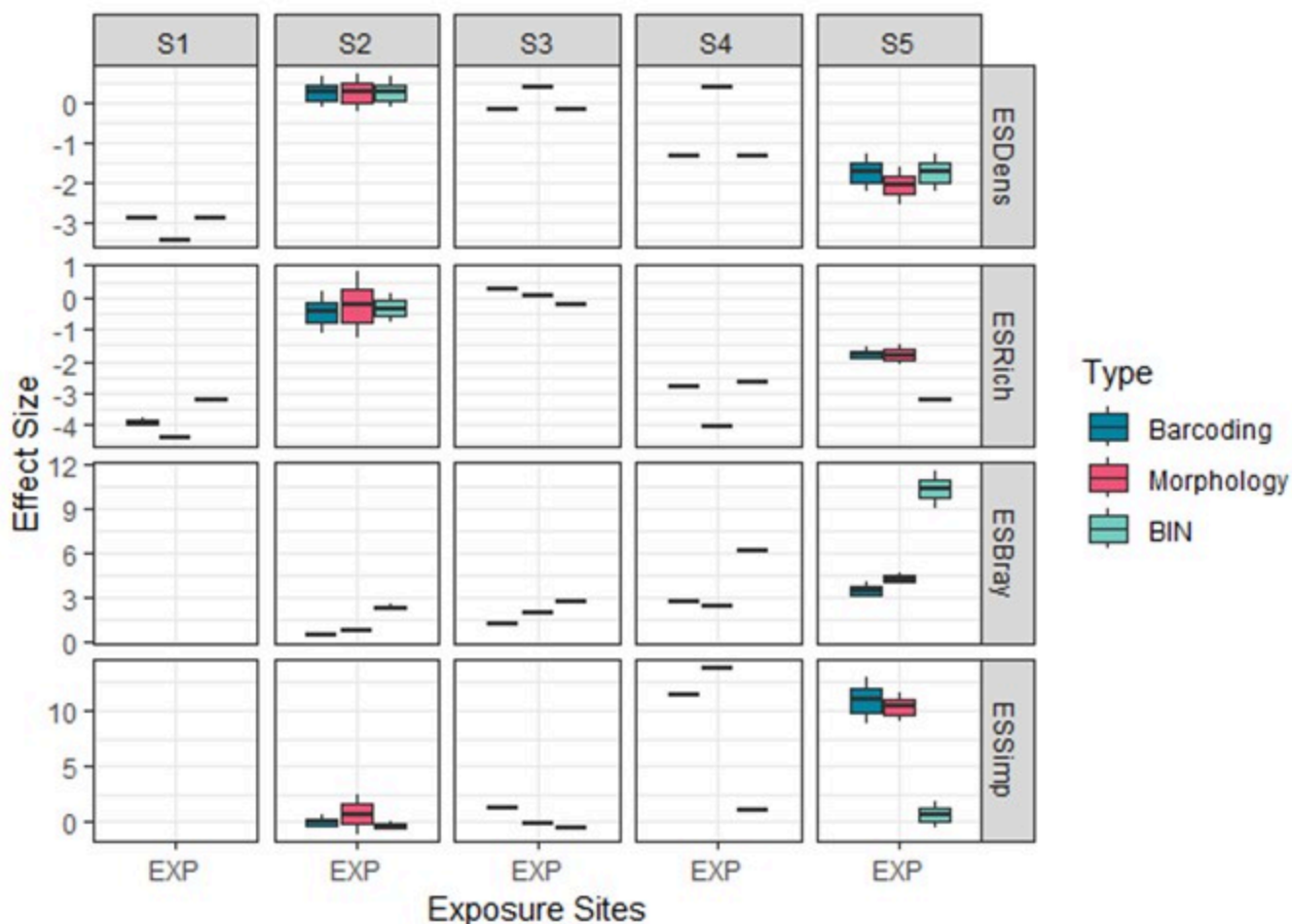


Figure 4

Effect sizes (exposure mean – reference mean / reference standard deviation) for the three datasets (family barcoding, family morphology, and BIN barcoding) and for all four endpoints. An effect size is calculated for each exposure area, and in the case of two exposure areas for one location, the values of replicates from both areas were used to derive box plots.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.docx](#)