



DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds

Downloaded from: <https://research.chalmers.se>, 2025-12-05 03:11 UTC

Citation for the original published paper (version of record):

Sun, Z., Majaneva, M., Sokolova, E. et al (2019). DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds. *Ecology and Evolution*, 9(17): 9712-9722. <http://dx.doi.org/10.1002/ece3.5503>

N.B. When citing this work, cite the original published paper.

ORIGINAL RESEARCH

DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds

Zhenhua Sun¹  | Markus Majaneva² | Ekaterina Sokolova¹ | Sebastien Rauch¹ | Sondre Meland^{3,4} | Torbjørn Ekrem²

¹Architecture and Civil Engineering, Water Environment Technology, Chalmers University of Technology, Gothenburg, Sweden

²Department of Natural History, Norwegian University of Science and Technology, NTNU University Museum, Trondheim, Norway

³Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

⁴Norwegian Institute for Water Research (NIVA), Oslo, Norway

Correspondence

Zhenhua Sun, Architecture and Civil Engineering, Water Environment Technology, Chalmers University of Technology, 412 96 Gothenburg, Sweden.
Email: zhenhua.sun@chalmers.se

Funding information

Norwegian Public Roads Administration; Ferry-free E39 Programme

Abstract:

1. Stormwater ponds are used to compensate for the adverse effects that road runoff might have on the natural environment. Depending on their design and placement, stormwater ponds can act as both refugia and traps for local biodiversity. To evaluate the impact of stormwater ponds on biodiversity, it is critical to use effective and precise methods for identification of life associated with the water body. DNA metabarcoding has recently become a promising tool for identification and assessment of freshwater biodiversity.
2. Using both morphology and DNA metabarcoding, we analyze species richness and biological composition of samples from 12 stormwater ponds and investigate the impact of pond size and pollution levels in the sediments and water column on the macroinvertebrate community structure.
3. DNA metabarcoding captured and identified more than twice the number of taxa compared to morphological identification. The (dis)similarity of macroinvertebrate community composition in different ponds showed that the ponds appear better separated in the results obtained by DNA metabarcoding, but that the explained variation is higher for the results obtained by morphologically identification, since it provides abundance data.
4. The reliance on morphological methods has limited our perception of the aquatic biodiversity in response to anthropogenic stressors, thereby providing inaccurate information for appropriate design and management of stormwater ponds; these drawbacks can be overcome by DNA metabarcoding.
5. *Synthesis and applications.* The results indicate that DNA metabarcoding is a useful tool in identifying species, especially Diptera, which are difficult to determine. Application of DNA metabarcoding greatly increases the number of species identified at each sampling site, thereby providing a more accurate information regarding the way the ponds function and how they are affected by management.

KEYWORDS

DNA metabarcoding, high throughput sequencing, macroinvertebrates, pollutants, road runoff, stormwater ponds

1 | INTRODUCTION

Roads may cause harmful effects on aquatic ecosystems, due to alternations in hydrological processes (Karlson & Mörtberg, 2015) and pollutants from vehicles, asphalt, and operation and maintenance activities (Meland et al., 2011; Meland, Salbu, & Rosseland, 2010). The awareness that roads may significantly impair the aquatic environment has resulted in the construction of an increasing number of stormwater ponds along roads. Several studies have reported on the biodiversity services the stormwater ponds provide as anthropogenic habitats, enhancing urban biodiversity (e.g., Hassall, 2014; Hill et al., 2017; Le Viol, Chiron, Julliard, & Kerbiriou, 2012; Stephansen, Nielsen, Hvitved-Jacobsen, Pedersen, & Vollertsen, 2016). However, some studies, for example, Helfield and Diamond (1997), question such uses of stormwater ponds, mainly since their primary function of improving water quality leads to the presence and bioavailability of pollutants. In order to clarify the biodiversity value of stormwater ponds and improve the management of stormwater systems to support biodiversity, effective methods should be applied to determine effects of pollutants on the biota.

Current research on aquatic biodiversity in stormwater ponds is typically based on identification of organisms by morphological properties (Briers, 2014; Le Viol, Mocq, Julliard, & Kerbiriou, 2009; Tixier, Rochfort, Grapentine, Marsalek, & Lafont, 2012). However, identification/sorting of invertebrate using morphological properties to species level is difficult and expensive (Baloğlu, Clews, & Meier, 2018; Wang, Srivathsan, Foo, Yamane, & Meier, 2018). It has been shown that the cost of, for example, midge, identification using morphological properties is quite high due to the requirement of dissection and mounting of specimens onto microscopic slides (Baloğlu et al., 2018). As a result, the rarely used species-level data lead to the undesirable information of responses to stressors, because different species have different sensitivity to environmental parameters (Baloğlu et al., 2018; Beermann, Zizka, Elbrecht, Baranov, & Leese, 2018; Macher et al., 2015).

Recently, DNA-based identification methods have gained popularity in overcoming the limitations of morphological identification. DNA metabarcoding allows rapid and cost-effective assessment and monitoring of biodiversity using massive parallel sequencing of bulk samples (Cristescu, 2014; Leese et al., 2018). Several studies have applied and compared DNA metabarcoding with morphological identification in assessments of macroinvertebrate diversity (Chan et al., 2014; Elbrecht & Leese, 2017; Renaud, Savage, & Adamowicz, 2012). Beermann et al. (2018) indicated that although, for example, in the case of chironomids, most operational taxonomic units (OTUs) (>85%) do not match with high similarity (>98%) to formally described species due to incomplete reference libraries, DNA-based approaches are still potentially useful when studying environmental impacts, especially in the context of multiple stressors.

A major challenge in (PCR-based) metabarcoding is that it does not provide quantitative abundance data, especially in the case of diverse community samples (Elbrecht & Leese, 2015). Several studies have detected significant linear relationships between the number

of morphologically identified specimens and the number of sequencing reads but with poor fit and high scatter (Carew, Pettigrove, Metzeling, & Hoffmann, 2013; Clarke, Beard, Swadling, & Deagle, 2017; Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017). This means that DNA metabarcoding data can be used only as a semi-quantitative abundance estimator. In addition, if the specimen in the reference library has been wrongly identified or not updated according to the latest taxonomic revision, a match will be assigned to the wrong name (Pawlowski et al., 2018).

This study is focused on testing the following hypothesis: DNA metabarcoding can provide better results and explanation than morphological identification in identifying organisms to species level and studying environmental impacts. Macroinvertebrates in stormwater ponds were identified by both morphology and DNA metabarcoding, allowing us to (a) compare the (dis)similarity of macroinvertebrate community composition in different ponds using the data obtained by morphological identification and DNA metabarcoding; (b) compare the differences in species richness of groups in common using the data obtained by morphological identification and DNA metabarcoding; and (c) compare the power of morphological identification and DNA metabarcoding in predicting the impacts of environmental variables on the macroinvertebrate community composition.

2 | MATERIALS AND METHODS

Twelve stormwater ponds were investigated. The ponds are located along the major highways E6 and E18 in the counties of Oslo, Akershus, and Østfold in southern Norway (Figure S1).

2.1 | Biological sampling

Biological samples were collected four times (April, June, August, and October) in 2013. The samples were collected with a kick net and traps close to the inlet and twice on either side of the main basin within each pond. A kick net with a 30 × 30 cm opening and a mesh size of 0.45 mm was used. Kick sampling with five sweeps was applied, when there were small stones at the bottom; five sweeps were taken through the water and aquatic vegetation at approximately 50 cm depth, when the bottom material was not stony. Traps made of 1.5 L transparent plastic bottles were also used to collect samples (Sun et al., 2018); the traps were left for 1–4 days, according to time of the year. The samples were preserved in 70% ethanol.

Morphological identification was performed by sorting organisms in the laboratory and identifying them to species level when possible using morphological characteristics. Nilsson (1996, 1997) was used to identify benthic macroinvertebrates; Nilssen (1974) was used to identify Chaoboridae.

The morphologically identified specimens were collected in one glass vial per sample. The samples were preserved in 70% ethanol and stored at +5°C. DNA was extracted by first pouring the specimens onto Petri dishes, carefully pipetting away excess ethanol,

and drying the samples overnight until the specimens were completely dry. The dried specimens were collected in sterile mortar and ground with a pestle in liquid nitrogen. The obtained tissue powder was transferred to microcentrifuge tubes and stored at -20°C . From each sample, 10–20 mg subsamples were taken in triplicate to DNeasy 96 plates where DNA was extracted using DNeasy 96 Blood & Tissue Kit (Qiagen GmbH). The samples were incubated with proteinase K and buffer ATL at 56°C overnight, and the subsequent steps were done according to the manufacturer's protocol, except DNA was eluted in $2 \times 200 \mu\text{l}$ of DNase-free water.

Two fragments of the mitochondrial COI gene were amplified three times from each three DNA extraction subsamples. The approximately 230 bp long fragment at the 5' end of the barcoding region was amplified, using the LCO1490 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) forward primer (GGTCAACAAATCATAAAGATATTGG) and the 230_R (Gibson et al., 2015) reverse primer (CTTATRTTTRTTATTCGIGGAAIGC). The approximately 420 bp long fragment at the 3' end of the barcoding region was amplified, using the forward primer BF2 (GCHCCHGAYATRGCHTTYCC) and the reverse primer BR2 (TCDGGRTGNCCRAARAAYCA) described by Elbrecht and Leese (2017). The amplifications were carried out with Illumina forward (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG) and reverse (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG) adapters.

The PCR and bioinformatic processing were conducted as described in Majaneva, Diserud, Eagle, Hajibabaei, and Ekrem (2018). One exception is that for the reads amplified with BF2 and BR2 primers, all reads shorter than 430 and longer than 490 bases were removed. The exact commands are provided in the Appendix S1.

2.2 | Environmental variables

Water and sediment samples were collected once in April 2013. Water samples were collected close to the inlet of the ponds. Sampling was performed using acid washed 125 ml polyethylene (PE) bottles, including acid washed bottles for metal analysis. Top-layer sediments were collected close to the inlet using a spade and stored in 1 L glass bottles. Since the highest contaminant concentrations are typically found in the inlet (Allen, Haynes, & Arthur, 2017), the samples serve as a proxy for detecting contamination within the ponds.

All sediment and water quality analyses were performed by Rambøll Analytics Laboratories in Finland. Fourteen water quality variables were analyzed in this study, including metals (aluminum (Al), barium (Ba), potassium (K), calcium (Ca), copper (Cu), magnesium (Mg), manganese (Mn), silicon (Si), iron (Fe), and titanium (Ti)), chloride (Cl^-), total nitrogen (TN), total phosphorus (TP), and sulfate (SO_4^{2-}). Eleven sediment variables were analyzed, including total organic carbon (TOC), total hydrocarbons, 16 polycyclic aromatic hydrocarbons (US EPA 16 PAHs), aluminum (Al), calcium (Ca), chromium (Cr), copper (Cu), iron (Fe), nickel (Ni), lead (Pb), and zinc (Zn). In the subsequent statistical analysis, pyrene was used as a proxy for PAH pollution, because the concentrations of most other PAH

compounds were below the limit of quantification (LOQ) in more than 15% of the total number of samples.

Pond size, which ranges from 89 to $1,400 \text{ m}^2$, was obtained from the Norwegian Public Roads Administration (NPRA) (Table S1).

2.3 | Statistical analyses

The number of species from groups in common between DNA metabarcoding and morphological identification was compared using bar charts generated with R software (R Core Team, 2013).

Principal coordinates analysis (PCoA) and distance-based redundancy analysis (db-RDA) were performed on the datasets obtained using both DNA metabarcoding and morphological identification. PCoA was used to compare the (dis)similarity of whole macroinvertebrate community composition in different ponds. db-RDA was used to explore the impacts of environmental variables on the macroinvertebrate community composition of groups in common of the two methods, aiming to compare the power of morphological identification with that of DNA metabarcoding. Groups that were included in the db-RDA were Gastropoda, Odonata, Ephemeroptera, Plecoptera, Trichoptera, Lepidoptera, Megaloptera, Hemiptera, Diptera, Coleoptera, Oligochaeta, and Hirudinea. Using morphology, some macroinvertebrates were only identified to genus or (sub)family level. DNA metabarcoding can only provide reliable data for presence/absence, while morphological identification can provide abundance data. Therefore, both abundance and presence/absence of each taxon were used in the morphological dataset, while presence/absence of each species was used in the DNA metabarcoding dataset, in which 1 represents presence and 0 represents absence. The abundance data were $\log(x + 1)$ transformed before PCoA and db-RDA to increase the influence of subordinate species; otherwise, the results would be completely dependent on the dominant species (Šmilauer & Lepš, 2014). PCoA was used to calculate a distance matrix and represent the (dis)similarity measures in the Euclidean space; the (dis)similarity can then be assessed using linear models (Legendre & Anderson, 1999). The distances between points in the plot are close to original dissimilarities. Two similarity measures were used: Sørensen binary distance (Sørensen, 1948) for presence–absence data and Bray–Curtis distance (Bray & Curtis, 1957) for quantitative data (abundance); the equations are defined in the Appendix S1. In order to provide a better overview of correlation between different methods, Pearson's correlation analysis was applied using PCoA scores extracted from the first four ordination axes, and a matrix of the correlation coefficients was generated using R software (R Core Team, 2013).

Principal component analysis (PCA) was performed on both the sediment and water quality data to reduce the skewness and improve the normality of the data for the db-RDA, and the data were $\log(x + 1)$ transformed prior to the PCA. PCA sample scores extracted from axes 1 and 2 were used to represent pollution levels in the water column and sediments, respectively (Figure S2). PCA, PCoA, and db-RDA were performed using CANOCO 5 software (Canoco5, 2012). Monte Carlo permutation tests (999 permutations, $p < .05$) were used to determine the statistical significance. The species that

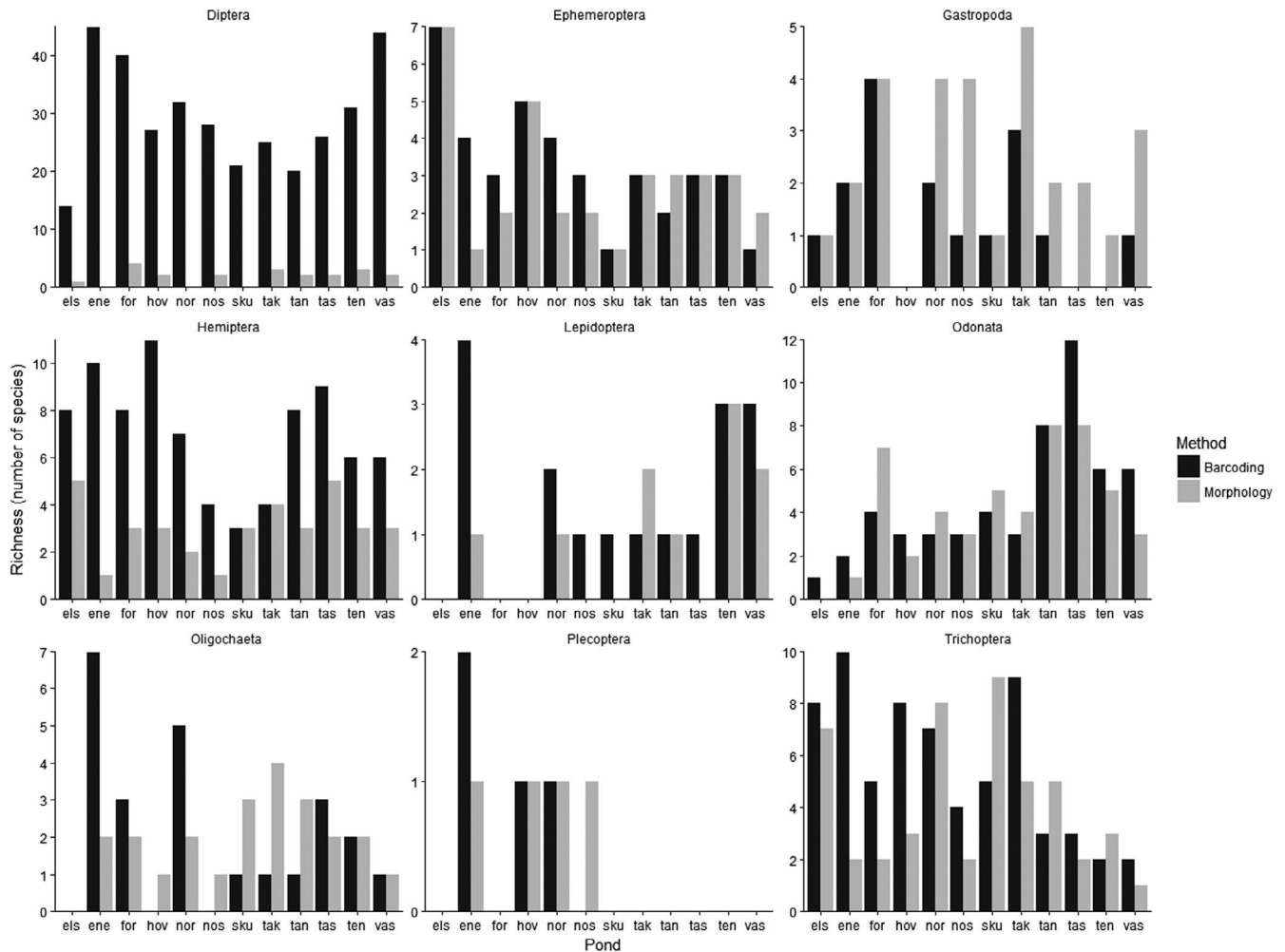


FIGURE 1 Comparison of the number of species detected by DNA metabarcoding and morphology in the taxonomic groups in common. The abbreviations along the x-axis represent names of ponds: els-Elstadmoen, ene-Enebekk, for-Fornebu, hov-Hovinmoen, nor-Nordby, nos-Nøstvedt, sku-Skullerud, tak-Taraldrud crossing, tan-Taraldrud north, tas-Taraldrud south, ten-Tenor, and vas-Vassum

showed correlation with ordination axes smaller than -0.6 or larger than 0.6 were displayed in the plots in order to avoid the ordination plots becoming too cluttered.

3 | RESULTS

3.1 | Comparison of species richness

Using morphological identification, a total of 156 macroinvertebrates were identified, of which 88 (56%) were identified to the species level and 68 (44%) were identified either to the genus or (sub)family level. Four macroinvertebrates are red-listed species in Norway, that is, *Planorbis planorbis* (deficient category), *Coenagrion lunulatum* (vulnerable category), *Orthetrum cancellatum* (vulnerable category), and *Chaoborus pallidus* (threatened category) (Henriksen & Hilmo, 2015). Among these four species, only one was identified using DNA metabarcoding.

Using DNA metabarcoding, collectively for all samples, a total of 10.7 and 12.9 million reads were generated from the F230 and BFR2 fragments, respectively. After quality control and merging the taxonomic

information from the two fragments, 599 OTUs were identified to invertebrate taxa, most of which were species, but some were genus or family level, because the read matched several species with a high similarity value. In the final dataset, the species that had five or less reads in one sample were excluded. As a result, 342 OTUs from benthic macroinvertebrates were included in the statistical analysis. DNA metabarcoding captured and identified more than twice the number of taxa compared to identification by morphology across the 12 study ponds. A total of 196 (57%) OTUs could be unambiguously associated with a species name in the barcode reference library, while 146 (43%) were left without species-level identification due to the low percentage match for local species or due to matching multiple species with same percentage.

Both identification methods detected Oligochaeta, Diptera, Ephemeroptera, Gastropoda, Lepidoptera, Hemiptera, Odonata, Plecoptera, and Trichoptera; the number of species from these groups was compared. Identification of species by DNA metabarcoding resulted in considerably more Diptera species than identification by morphology in all ponds (Figure 1), and 66% of all Diptera belonged to the large and ecologically diverse family Chironomidae. There is no comprehensive literature that can be used to identify

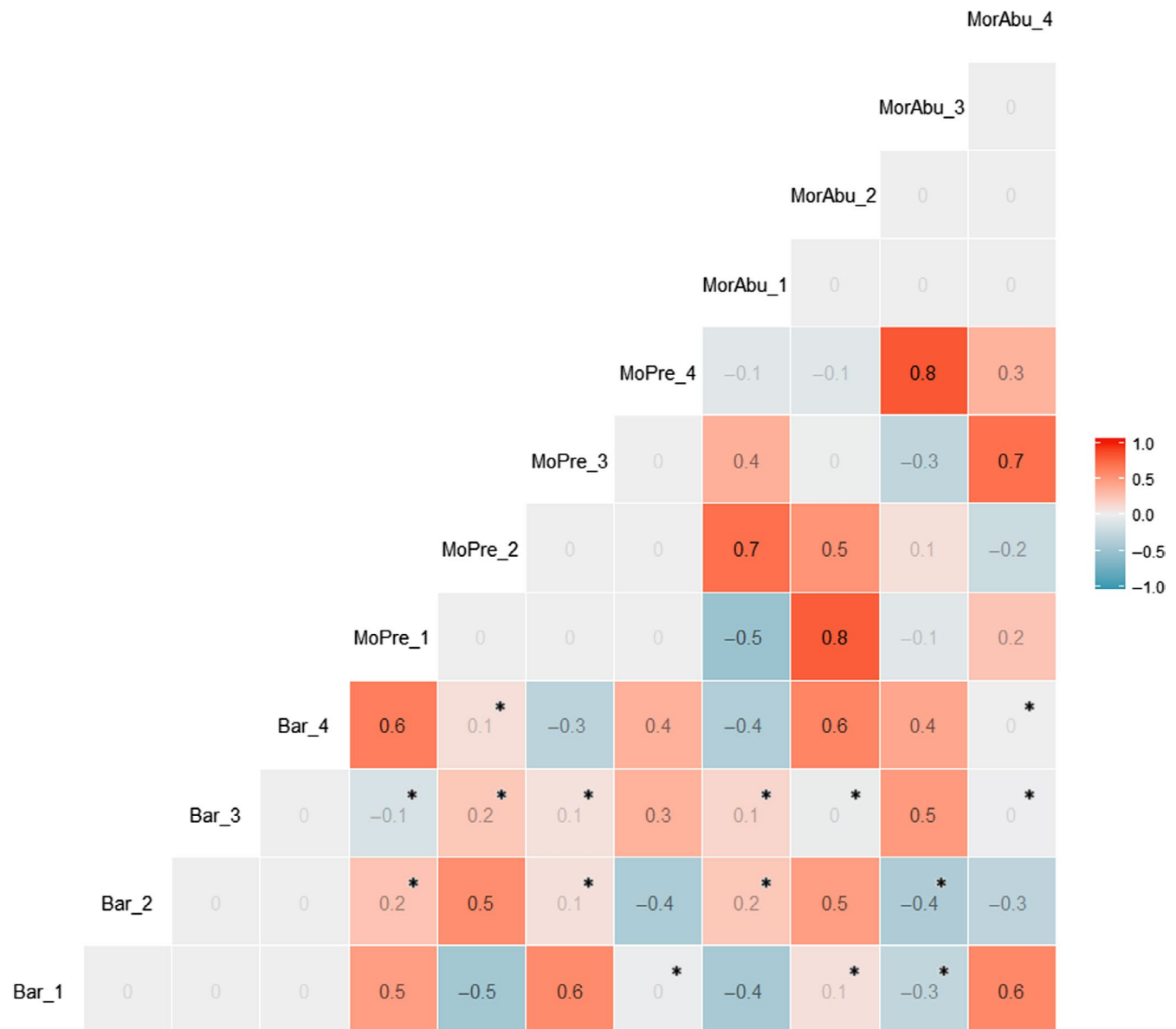


FIGURE 3 The Pearson correlation matrix of principal coordinates analysis (PCoA) scores extracted from the first four axes in three datasets, including DNA metabarcoding, morphological identification based on presence/absence, and morphological identification based on abundance. “Bar_1,” “Bar_2,” “Bar_3,” and “Bar_4” represent PCoA scores extracted from axes 1, 2, 3, and 4 from DNA metabarcoding dataset, respectively; “MoPre_1,” “MoPre_2,” “MoPre_3,” and “MoPre_4” represent PCoA scores extracted from axes 1, 2, 3 and 4 from morphological identification based on presence/absence dataset, respectively; “MorAbu_1,” “MorAbu_2,” “MorAbu_3,” and “MorAbu_4” represent PCoA scores extracted from axes 1, 2, 3, and 4 from morphological identification based on abundance, respectively. “*” represents the nonsignificant correlations between DNA metabarcoding and morphological identifications

pond size, explained 32.2% of variances (first axis: $F = 1.3$, $p = .02$; all axes: $F = 1.3$, $p = .01$) in the macroinvertebrate community composition, and the first two axes explained 14% and 9%, respectively (Figure 4a). For the morphological dataset based on presence/absence, the three variables explained 34.4% of variances (first axis: $F = 1.3$, $p = .081$; all axes: $F = 1.4$, $p = .004$), and the first two axes explained 14% and 13%, respectively; for the morphological dataset based on abundance, these three environmental variables explained 36.6% of variances (first axis: $F = 1.8$, $p = .07$; all axes: $F = 1.5$, $p = .023$), and the first two axes explained 18% and 14%, respectively. However, the first axis based on datasets derived from

morphological identifications was not significantly correlated with the macroinvertebrate community composition (Figure 4b,c). Similar to the results of PCoA, the abundance dataset explained most of the variation, followed by presence/absence dataset based on morphological identification. The DNA metabarcoding dataset explained the least variation, a consequence of increased complexity when the number of species increases, because species responds differently. In addition, it is normal that the explained variation in first axes is reduced when the number of response variables is increased. As an example, when DNA metabarcoding was applied to a species-rich group, for example, Diptera, it provided better results, which

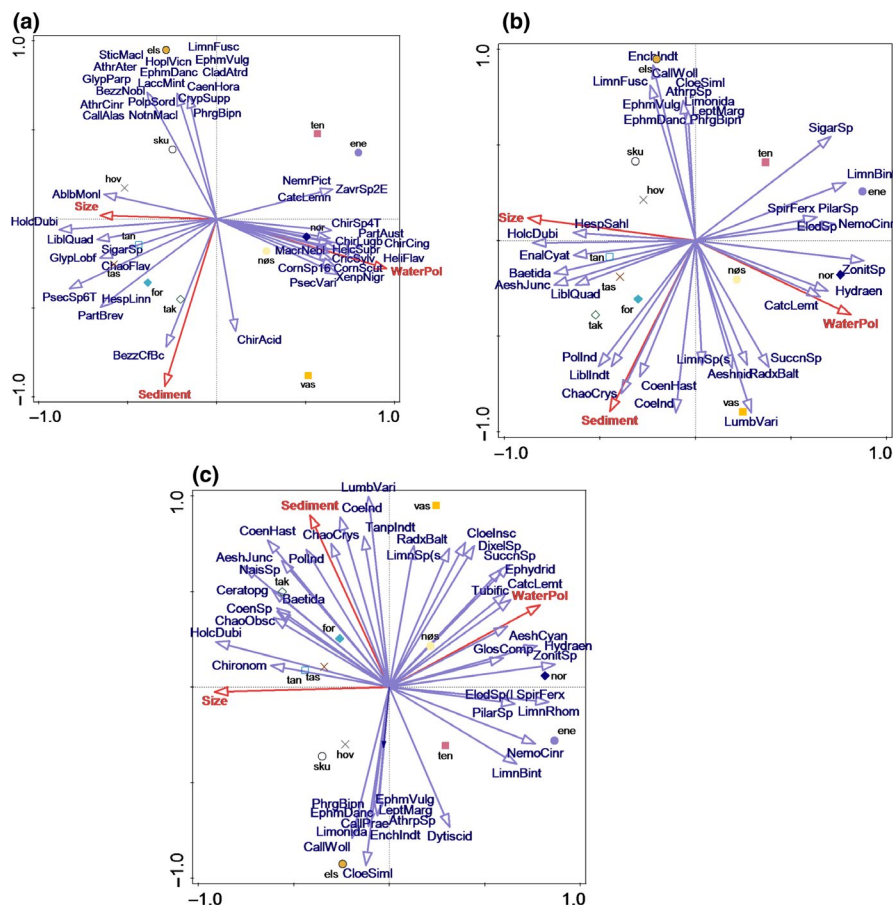


FIGURE 4 The relationship between macroinvertebrate community composition and pollution levels in the water column and sediments, as well as pond size. (a) Results obtained from db-RDA for the DNA metabarcoding dataset. (b) Results obtained from db-RDA for the morphological dataset based on presence/absence. (c) Results obtained from db-RDA for the morphological dataset based on abundance. The purple arrows represent different species; the red arrows represent environmental variables, in which WaterPol and Sediment represent pollution levels in the water column and sediments, respectively

explained the highest variation in the macroinvertebrate community composition (Figure S3).

DNA metabarcoding-based results showed that the taxa were divided into several obvious clusters (Figure 4a), while the morphology-based results showed that the taxa were distributed all over the space (Figure 4b,c). For the DNA metabarcoding dataset, 26 dipteran species were exhibited in the plot; 12 of these were positively correlated with the pollution level in the sediments, while 14 were negatively correlated. Similarly, 13 dipteran species were positively correlated with pollution level in the water column, while 13 were negatively correlated. Different from DNA metabarcoding-based results, morphology-based results based on abundance showed that most of the displayed dipteran taxa (seven out of nine) were positively correlated with pollution level in the sediment. However, five of the dipteran taxa that were positively correlated with pollution level in the sediment were only identified to the higher taxonomic levels: one was identified to genus level (*Dixella*), one was identified to subfamily level (Tanypodinae), and three were identified to family level (Ceratopogonidae, Chironomidae, and Ephydriidae). Only two were identified to the species level, *Chaoborus crystallinus* and *Chaoborus obscuripes*, and they were recognized by both DNA metabarcoding and morphology, and had the same responses to environmental variables according to both methods (Figure 4), that is, they were positively correlated with pond size and pollution level in the sediments, but negatively correlated with pollution level in the water column.

Morphological identification including abundance showed that among the taxa that were negatively correlated with the pollution level in sediments, 43% belonged to Ephemeroptera, Plecoptera, and Trichoptera, and most of them were identified to the species level. Several species of Trichoptera and Ephemeroptera were also found by DNA metabarcoding, and they exhibited the same responses to the environmental variables as morphological identification. For instance, *Phryganea bipunctata*, *Ephemera vulgata*, *Ephemera danica*, and *Limnephilus fuscineis* were all positively correlated with pond size and negatively correlated with the pollution levels in the sediments and water column.

Odonata (dragonflies (Anisoptera) and damselflies (Zygoptera)) and Hemiptera were the groups shown in the plots for both datasets. Common species, such as *Libellula quadrimaculata*, exhibited the same responses to the environmental variables for the two methods, in which they were positively correlated with the pond size and pollution level in the sediment, but negatively correlated with the pollution level in the water column. One common hemipteran species, *Sigara* sp. exhibited different responses to the environmental variables between different methods.

4 | DISCUSSION

Existing freshwater bioassessment programs often use higher-level taxonomic groups, for example, genus or family, due to the practical

reasons (e.g., Carlson, Donadi, & Sandin, 2018; Pandey et al., 2017). Consequently, there is still a great deal of uncertainty in determining species-specific responses of pond ecosystem to various environmental factors. This study reveals the combined effects of pollution levels in sediments and water column as well as pond size on macroinvertebrates identified using both DNA metabarcoding and morphology to provide more precise information for bioassessment and management.

DNA metabarcoding worked especially well for Diptera, because it identified considerably more Diptera species than morphological identification in all ponds. Morphological identification of Diptera larvae, especially Chironomidae, is often very difficult and complicated even for specialists, due to the lack of diagnostic characters (Baloğlu et al., 2018; Dobson, 2013; Silva & Wiedenbrug, 2014). In addition, improper specimen handling leading to the damage of important morphological features, such as scales and bristles, makes morphological identification a difficult task (Chan et al., 2014). The results of the current study are in agreement with the findings of, for example, Ekrem, Stur, and Hebert (2010), Silva and Wiedenbrug (2014) and Lin, Stur, and Ekrem (2018), who found that DNA barcoding is a promising tool for identification of Diptera and Chironomidae. DNA barcoding is also effective in detecting previously unknown diversity (Lin, Stur, & Ekrem, 2015), as shown in our study.

However, for some groups, for example, Oligochaeta and Odonata, DNA metabarcoding identified less species in some samples compared with the morphological identification. The accuracy of DNA metabarcoding identification depends on the "barcode gap," the existence of a gap between interspecific and intraspecific genetic divergence. Thus, for species complexes where maximum intraspecific distance exceeds the distance to the nearest neighbor, species-level identification becomes ambiguous (Jinbo, Kato, & Ito, 2011; Sun et al., 2016). There are various reasons resulting in the lack of large interspecific divergence, such as hybridization between closely related species, short time since speciation, and incomplete lineage sorting (Lin et al., 2015). High intraspecific divergence can also affect identification success if they represent cryptic species (i.e., multiple genetic clusters in the reference library have the same name because they cannot be separated morphologically) or are due to infection with endoparasites (e.g., *Wolbachia*) that affect sexual compatibility and fertility resulting in divergent mitochondrial lineages within a species (e.g., Smith et al., 2012). Furthermore, if reference sequences are lacking in established reference databases, DNA metabarcoding identification will also fail (Chan et al., 2014). The missing Oligochaeta, Odonata, and Gastropoda species in our DNA metabarcoding results can be attributed to several reasons. First of all, it can be that our primers did not match well with the species; but this is not plausible for all taxa, as several of the missing species in some samples were detected by DNA metabarcoding in other samples. Secondly, the oligochaete reads could have been missed due to our read abundance threshold of five reads per the sample in order to be included in the analyses; this could mean that oligochaetes were present in the samples, but in very low abundance. Thirdly, the species that were identified morphologically may have not been included in the reference library, and therefore excluded from the

DNA metabarcoding results, as we have taken into account only the reads that had a higher than 97% match with the reference library. Last but not least, the samples may not have been preserved well, as Oligochaeta normally need more specialized preservation to keep their DNA available (Rodriguez & Reynoldson, 2011). Thus, DNA metabarcoding gives more species-level information in general, but results are only as good as the reference library (Hajibabaei, Baird, Fahner, Beiko, & Golding, 2016).

Regarding the (dis)similarity of macroinvertebrate community composition, the taxa analyzed by morphological identification using presence/absence spread all over the space, while the species identified using DNA metabarcoding were clearly divided into several clusters. Results from morphological identification using abundance provided the highest explained variation of macroinvertebrates. However, DNA metabarcoding detected a much higher number of taxa. Therefore, the dimensionality of the DNA metabarcoding data is considerably more complex compared to the morphological data, and an increase in the number of species is normally followed by a decrease in the explained variation.

The correlation matrix showed that PCoA scores extracted from the first axis of DNA metabarcoding were strongly correlated with PCoA scores extracted from the first three axes of morphological identification based on presence/absence, indicating community composition patterns were similar between the DNA metabarcoding and morphological identification based on presence/absence. Despite a correspondence in community composition patterns between datasets from DNA metabarcoding and morphological identification based on presence/absence, there was a lack of significant correlation between the DNA metabarcoding and morphological identification based on abundance, and this could be attributed to the difference in data formats, in which DNA metabarcoding was semi-quantitative dataset and morphological identification based on abundance was quantitative dataset.

Regarding the results of db-RDA, it should be mentioned that the water pollution level in our study just reflected a snapshot of the water quality and that concentrations in stormwater ponds vary a lot due to the episodic nature of road runoff. Therefore, the sediment pollution is more reliable as an indicator of pollutant levels in the ponds, and probably much more reliable when comparing the pollutant levels between different ponds. Casey et al. (2007) also demonstrated that the loadings of sediments with pollutants are relatively stable over time.

Our results for the morphological methods are in line with those of Tchakonté et al. (2015), Arimoro, Odume, Uhunoma, and Edegbene (2015), and Hilsenhoff (1988), who found that most dipteran taxa are known as pollution tolerant, normally colonizing polluted waters. However, Diptera in these studies were identified to genus, subfamily, or family level only. Raunio, Paavola, and Muotka (2007) suggested that when chironomid genera are identified to the species level, they should be classified as pollution intolerant, because congeneric species may be different in their tolerance, this being particularly relevant for the species-rich genera. Different from the results for morphological identification, the results for DNA metabarcoding

indicates that species in the highly species-rich group Diptera show a broader variation in their tolerance to pollution than just classifying them as pollutant tolerant taxa.

The common species from Odonata group, *Libellula quadrimaculata*, that exhibited the same response to environmental variables in DNA metabarcoding and morphological identification belongs to the family Libellulidae, which contains many tolerant species (Villalobos-Jiménez, Dunn, & Hassall, 2016). Dragonflies and damselflies are normally considered sensitive to different stressors, such as pollutants (Ferreras-Romero, Márquez-Rodríguez, & Ruiz-García, 2009); however, although some sensitive taxa belong to Odonata, the sensitivity/tolerance to a stressor depends on the particular species rather than a larger taxon (Villalobos-Jiménez et al., 2016). For example, *Aeshna juncea* in our results exhibited positive correlation with pollution level in the sediments and pond size, while negative correlation with the pollution level in the water column. Girgin, Kazanci, and Dügel (2010) have proven that *Aeshna juncea* is highly tolerant of manganese and nickel.

DNA metabarcoding is proposed as a tool to reduce the time and cost necessary for morphological identification: the cost of DNA metabarcoding using high throughput sequencing techniques (HTS) is comparable to or even lower than traditional morphology-based methods (Hulley et al., 2018; Yu et al., 2012). The use of DNA metabarcoding could also result in substantial reduction in time spent on identification. It normally takes much longer time to process samples and produce data for morphology-based methods, while the corresponding time for DNA metabarcoding based on HTS is several days due to the higher throughput and no need for sorting individual samples (Tnah et al., 2019).

DNA metabarcoding is a powerful tool for identification of invertebrates (Beermann et al., 2018; Chan et al., 2014; Elbrecht et al., 2017). In our study, comparisons of morphological identification and DNA metabarcoding clearly illustrate that DNA metabarcoding was particularly beneficial for the identification of dipteran species. Since different species within higher taxonomic groups may respond differently to the pollution levels and other environmental variables, such differences are likely to be underestimated or go unnoticed if the taxonomic resolution is low. The results of this study have shown that reliance on morphological methods has limited our perception of the aquatic biodiversity in response to anthropogenic stressors. Even though most families of Diptera have been considered moderately to highly tolerant, most dipteran species vary with respect to pollution tolerance. Application of DNA metabarcoding would greatly increase the number of species identified at a sampling site, thereby providing a more accurate estimate of how aquatic biodiversity responds to anthropogenic stressors. Although the morphological dataset that included abundance provided the best-explained variation of macroinvertebrates and of environmental variables on the biological community composition, we cannot say that morphological identification including abundance performs significantly better than DNA metabarcoding. This is because DNA metabarcoding gave a significant increase in the number of species, substantially increasing the complexity and dimensionality of the data. Therefore, it is imperative that future

research and monitoring apply DNA metabarcoding to ensure that bioassessments do not underestimate the differences in species-specific responses to environmental factors.

ACKNOWLEDGMENTS

This study was funded by the Norwegian Public Roads Administration as part of the Research and Development Programme Nordic Road Water (NORWAT, www.vegvesen.no/norwat) and the Ferry-free E39 Programme. John E. Brittain, Svein Jakob Saltveit and Henning Pavels from the Natural History Museum, University of Oslo for collecting and identifying the biological samples. Ole Wiggo Røstad from the Norwegian University of Life Sciences verified the identification of Dytiscidae. Trond Bremnes from the Natural History Museum, University of Oslo verified Trichoptera and checked random samples of Heteroptera and Odonata. The sequencing was performed in the Genomics Core Facility (GCF), Norwegian University of Science and Technology (NTNU). GCF is funded by the Faculty of Medicine and Health Sciences at NTNU and Central Norway Regional Health Authority.

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

All authors contributed to the design of the study. M.M. conducted the molecular laboratory work; M.M. and Z.S. performed data analysis; Z.S. drafted the article, and M.M., E.S., S.R., S.M., and T.E. contributed revisions. All authors gave final approval for publication.

DATA AVAILABILITY STATEMENT

Whole dataset generated and analyzed during the current study is available in the ENA SRA repository with study name PRJEB30841.

ORCID

Zhenhua Sun  <https://orcid.org/0000-0002-0313-849X>

OPEN PRACTICES



This article has earned an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at <https://www.ebi.ac.uk/ena/data/view/PRJEB30841>.

REFERENCES

- Allen, D., Haynes, H., & Arthur, S. (2017). Contamination of detained sediment in sustainable urban drainage systems. *Water*, 9, 355.

- Arimoro, F. O., Odume, O. N., Uhunoma, S. I., & Edegbene, A. O. (2015). Anthropogenic impact on water chemistry and benthic macroinvertebrate associated changes in a southern Nigeria stream. *Environmental Monitoring and Assessment*, 187, 14. <https://doi.org/10.1007/s10661-014-4251-2>
- Baloğlu, B., Clews, E., & Meier, R. (2018). NGS barcoding reveals high resistance of a hyperdiverse chironomid (Diptera) swamp fauna against invasion from adjacent freshwater reservoirs. *Frontiers in Zoology*, 15, 31. <https://doi.org/10.1186/s12983-018-0276-7>
- Beermann, A. J., Zizka, V. M. A., Elbrecht, V., Baranov, V., & Leese, F. (2018). DNA metabarcoding reveals the complex and hidden responses of chironomids to multiple stressors. *Environmental Sciences Europe*, 30, 26. <https://doi.org/10.1186/s12302-018-0157-x>
- Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs*, 27, 325–349. <https://doi.org/10.2307/1942268>
- Briers, R. A. (2014). Invertebrate communities and environmental conditions in a series of urban drainage ponds in eastern Scotland: Implications for biodiversity and conservation value of SUDS. *CLEAN – Soil, Air, Water*, 42, 193–200.
- Canoco 5. (2012). *Canoco for visualization of multivariate data (version 5.0)*. Ithaca, NY: Microcomputer Power.
- 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, 45. <https://doi.org/10.1186/1742-9994-10-45>
- Carlson, P. E., Donadi, S., & Sandin, L. (2018). Responses of macroinvertebrate communities to small dam removals: Implications for bioassessment and restoration. *Journal of Applied Ecology*, 55, 1896–1907. <https://doi.org/10.1111/1365-2664.13102>
- Casey, R. E., Simon, J. A., Atueyi, S., Snodgrass, J. W., Karouna-Renier, N., & Sparling, D. W. (2007). Temporal trends of trace metals in sediment and invertebrates from stormwater management ponds. *Water, Air, and Soil Pollution*, 178, 69–77. <https://doi.org/10.1007/s11270-006-9132-z>
- Chan, A., Chiang, L.-P., Hapuarachchi, H. C., Tan, C.-H., Pang, S.-C., Lee, R., ... Lam-Phua, S.-G. (2014). DNA barcoding: Complementing morphological identification of mosquito species in Singapore. *Parasites & Vectors*, 7, 569. <https://doi.org/10.1186/s13071-014-0569-4>
- Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, 7, 873–883. <https://doi.org/10.1002/ece3.2667>
- Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: Towards an integrative approach to the study of global biodiversity. *Trends in Ecology & Evolution*, 29, 566–571. <https://doi.org/10.1016/j.tree.2014.08.001>
- Dobson, M. (2013). Family-level keys to freshwater fly (diptera) larvae: A brief review and a key to European families avoiding use of mouthpart characters. *Freshwater Reviews*, 6, 1–32. <https://doi.org/10.1608/FRJ-6.1.450>
- Ekrem, T., Stur, E., & Hebert, P. D. N. (2010). Females do count: Documenting Chironomidae (Diptera) species diversity using DNA barcoding. *Organisms Diversity & Evolution*, 10, 397–408. <https://doi.org/10.1007/s13127-010-0034-y>
- Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass–sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, 10, e0130324. <https://doi.org/10.1371/journal.pone.0130324>
- Elbrecht, V., & Leese, F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science*, 5, 11.
- Elbrecht, V., Vámos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8, 1265–1275. <https://doi.org/10.1111/2041-210X.12789>
- Ferreras-Romero, M., Márquez-Rodríguez, J., & Ruiz-García, A. (2009). Implications of anthropogenic disturbance factors on the Odonata assemblage in a Mediterranean fluvial system. *International Journal of Odonatology*, 12, 413–428. <https://doi.org/10.1080/13887890.2009.9748354>
- 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, 294–299.
- 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, e0138432. <https://doi.org/10.1371/journal.pone.0138432>
- Girgin, S., Kazanci, N., & Dügel, M. (2010). Relationship between aquatic insects and heavy metals in an urban stream using multivariate techniques. *International Journal of Environmental Science & Technology*, 7, 653–664. <https://doi.org/10.1007/BF03326175>
- Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: How DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371, 20150330. <https://doi.org/10.1098/rstb.2015.0330>
- Hassall, C. (2014). The ecology and biodiversity of urban ponds. *Wiley Interdisciplinary Reviews: Water*, 1, 187–206. <https://doi.org/10.1002/wat2.1014>
- Helfield, M. J., & Diamond, L. M. (1997). Use of constructed wetlands for urban stream restoration: A critical analysis. *Environmental Management*, 21, 329–341. <https://doi.org/10.1007/s002679900033>
- Henriksen, S., & Hilmo, O. (2015). *Norwegian Red List of Species 2015 – methods and results*. Norway: Norwegian Biodiversity Information Centre.
- Hill, M. J., Biggs, J., Thornhill, I., Briers, R. A., Gledhill, D. G., White, J. C., ... Hassall, C. (2017). Urban ponds as an aquatic biodiversity resource in modified landscapes. *Global Change Biology*, 23, 986–999. <https://doi.org/10.1111/gcb.13401>
- Hilsenhoff, W. L. (1988). Rapid field assessment of organic pollution with a family-level biotic index. *Journal of the North American Benthological Society*, 7, 65–68. <https://doi.org/10.2307/1467832>
- 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, 1110–1116. <https://doi.org/10.1016/j.jglr.2018.07.013>
- Jinbo, U., Kato, T., & Ito, M. (2011). Current progress in DNA barcoding and future implications for entomology. *Entomological Science*, 14, 107–124. <https://doi.org/10.1111/j.1479-8298.2011.00449.x>
- Karlson, M., & Mörtberg, U. (2015). A spatial ecological assessment of fragmentation and disturbance effects of the Swedish road network. *Landscape and Urban Planning*, 134, 53–65. <https://doi.org/10.1016/j.landurbplan.2014.10.009>
- Le Viol, I., Chiron, F., Julliard, R., & Kerbiriou, C. (2012). More amphibians than expected in highway stormwater ponds. *Ecological Engineering*, 47, 146–154. <https://doi.org/10.1016/j.ecoleng.2012.06.031>
- Le Viol, I., Mocq, J., Julliard, R., & Kerbiriou, C. (2009). The contribution of motorway stormwater retention ponds to the biodiversity of aquatic macroinvertebrates. *Biological Conservation*, 142, 3163–3171. <https://doi.org/10.1016/j.biocon.2009.08.018>
- Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., ... Weigand, A. M. (2018) Chapter Two - Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: A perspective derived from the DNAqua-Net COST action. In D. A. Bohan, A. J. Dumbrell, G. Woodward & M.

- Jackson (Eds.), *Advances in ecological research* (pp. 63–99). Cambridge, MA: Academic Press.
- Legendre, P., & Anderson, M. J. (1999). Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, 69, 1–24. [https://doi.org/10.1890/0012-9615\(1999\)069\[0001:DBRATM\]2.0.CO;2](https://doi.org/10.1890/0012-9615(1999)069[0001:DBRATM]2.0.CO;2)
- Lin, X., Stur, E., & Ekrem, T. (2015). Exploring genetic divergence in a species-rich insect genus using 2790 DNA barcodes. *PLoS ONE*, 10, e0138993. <https://doi.org/10.1371/journal.pone.0138993>
- Lin, X., Stur, E., & Ekrem, T. (2018). DNA barcodes and morphology reveal unrecognized species in Chironomidae (Diptera). *Insect Systematics & Evolution*, 49, 329–398. <https://doi.org/10.1163/1876312X-00002172>
- Macher, J. N., Salis, R. K., Blakemore, K. S., Tollrian, R., Matthaei, C. D., & Leese, F. (2015). 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>
- Majaneva, M., Diserud, O. H., Eagle, S. H. C., Hajibabaei, M., & Ekrem, T. (2018). Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, 2, e26664. <https://doi.org/10.3897/mbmg.2.26664>
- Meland, S., Farmen, E., Heier, L. S., Rosseland, B. O., Salbu, B., Song, Y., & Tollefsen, K. E. (2011). Hepatic gene expression profile in brown trout (*Salmo trutta*) exposed to traffic related contaminants. *Science of the Total Environment*, 409, 1430–1443. <https://doi.org/10.1016/j.scitotenv.2011.01.013>
- Meland, S., Salbu, B., & Rosseland, B. O. (2010). Ecotoxicological impact of highway runoff using brown trout (*Salmo trutta* L.) as an indicator model. *Journal of Environmental Monitoring*, 12, 654–664. <https://doi.org/10.1039/B919420G>
- Nilssen, J. P. (1974). On the ecology and distribution of the Norwegian larvae of Chaoborus (Diptera, Chaoboridae). *Norsk ent. Tidsskr*, 21, 37–44.
- Nilsson, A. N. (1996, 1997) *Aquatic insects of north Europe: a taxonomic handbook*, vol. 1 and 2. Stenstrup, Denmark: Apollo Books.
- Pandey, L. K., Bergey, E. A., Lyu, J., Park, J., Choi, S., Lee, H., ... Han, T. (2017). The use of diatoms in ecotoxicology and bioassessment: Insights, advances and challenges. *Water Research*, 118, 39–58. <https://doi.org/10.1016/j.watres.2017.01.062>
- Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéoz-Perret-Gentil, L., Beja, P., Boggero, A., ... 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>
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Core Team.
- Raunio, J., Paavola, R., & Muotka, T. (2007). Effects of emergence phenology, taxa tolerances and taxonomic resolution on the use of the Chironomid Pupal Exuvial Technique in river biomonitoring. *Freshwater Biology*, 52, 165–176. <https://doi.org/10.1111/j.1365-2427.2006.01670.x>
- Renaud, A. K., Savage, J., & Adamowicz, S. J. (2012). DNA barcoding of Northern Nearctic Muscidae (Diptera) reveals high correspondence between morphological and molecular species limits. *BMC Ecology*, 12, 24–24. <https://doi.org/10.1186/1472-6785-12-24>
- Rodriguez, P., & Reynoldson, T. (2011). *The pollution biology of aquatic oligochaetes*.
- Silva, F., & Wiedenbrug, S. (2014). Integrating DNA barcodes and morphology for species delimitation in the Corynoneura group (Diptera: Chironomidae: Orthoclaadiinae). *Bulletin of Entomological Research*, 104, 65–78.
- Šmilauer, P., & Lepš, J. (2014). *Multivariate analysis of ecological data using Canoco 5*, 2nd edn. New York: Cambridge University Press.
- Smith, M. A., Bertrand, C., Crosby, K., Eveleigh, E. S., Fernandez-Triana, J., Fisher, B. L., ... Zhou, X. (2012). Wolbachia and DNA barcoding insects: Patterns, potential, and problems. *PLoS ONE*, 7, e36514. <https://doi.org/10.1371/journal.pone.0036514>
- Sørensen, T. (1948). *A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons*. Copenhagen: Biologiske Skrifter.
- Stephansen, D. A., Nielsen, A. H., Hvitved-Jacobsen, T., Pedersen, M. L., & Vollertsen, J. (2016). Invertebrates in stormwater wet detention ponds – Sediment accumulation and bioaccumulation of heavy metals have no effect on biodiversity and community structure. *Science of the Total Environment*, 566–567, 1579–1587. <https://doi.org/10.1016/j.scitotenv.2016.06.050>
- Sun, S., Li, Q. I., Kong, L., Yu, H., Zheng, X., Yu, R., ... Lin, J. (2016). DNA barcoding reveal patterns of species diversity among northwestern Pacific molluscs. *Scientific Reports*, 6, 33367. <https://doi.org/10.1038/srep33367>
- Sun, Z., Brittain, J. E., Sokolova, E., Thygesen, H., Saltveit, S. J., Rauch, S., & Meland, S. (2018). Aquatic biodiversity in sedimentation ponds receiving road runoff – What are the key drivers? *Science of the Total Environment*, 610, 1527–1535. <https://doi.org/10.1016/j.scitotenv.2017.06.080>
- Tchakonté, S., Ajeagah, G. A., Camara, A. I., Diomandé, D., Nyamsi Tchatcho, N. L., & Ngassam, P. (2015). Impact of urbanization on aquatic insect assemblages in the coastal zone of Cameroon: The use of biotraits and indicator taxa to assess environmental pollution. *Hydrobiologia*, 755, 123–144. <https://doi.org/10.1007/s10750-015-2221-5>
- Tixier, G., Rochfort, Q., Grapentine, L., Marsalek, J., & Lafont, M. (2012). Spatial and seasonal toxicity in a stormwater management facility: Evidence obtained by adapting an integrated sediment quality assessment approach. *Water Research*, 46, 6671–6682. <https://doi.org/10.1016/j.watres.2011.12.031>
- Tnah, L. H., Lee, S. L., Tan, A. L., Lee, C. T., Ng, K. K. S., Ng, C. H., & Nurul Farhanah, Z. (2019). DNA barcode database of common herbal plants in the tropics: A resource for herbal product authentication. *Food Control*, 95, 318–326. <https://doi.org/10.1016/j.foodcont.2018.08.022>
- Villalobos-Jiménez, G., Dunn, A., & Hassall, C. (2016). Dragonflies and damselflies (Odonata) in urban ecosystems: A review. *European Journal of Entomology*, 113, 217–232. <https://doi.org/10.14411/eje.2016.027>
- Wang, W. Y., Srivathsan, A., Foo, M., Yamane, S. K., & Meier, R. (2018). Sorting specimen-rich invertebrate samples with cost-effective NGS barcodes: Validating a reverse workflow for specimen processing. *Molecular Ecology Resources*, 18, 490–501. <https://doi.org/10.1111/1755-0998.12751>
- Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623. <https://doi.org/10.1111/j.2041-210X.2012.00198.x>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Sun Z, Majaneva M, Sokolova E, Rauch S, Meland S, Ekrem T. DNA metabarcoding adds valuable information for management of biodiversity in roadside stormwater ponds. *Ecol Evol*. 2019;9:9712–9722. <https://doi.org/10.1002/ece3.5503>