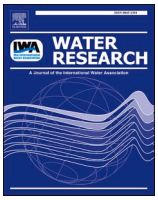
Water Research 272 (2025) 122987

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres



Fit for purpose? Evaluating benthic invertebrate DNA metabarcoding for ecological status class assessment in streams under the Water Framework Directive

Till-Hendrik Macher [a,b,*], Arne J. Beermann [a,c], Jens Arle [d], Julia Foerster [e], Matthias Greyer [f], Demetrio Mora [g,h,i], Jan Koschorreck [d], Peter Rolauffs [j], Anne Rother [f], Susanne Schüler [e], Jonas Zimmermann [g], Daniel Hering [j,c], Florian Leese [a,c,*]

[a] University of Duisburg-Essen, Aquatic Ecosystem Research, Universitaetsstr. 5, 45141, Essen, Germany
[b] University of Trier, Biogeography, Universitaetsring 15, 54296, Trier, Germany
[c] University of Duisburg-Essen, Centre for Water and Environmental Research (ZWU), Universitaetsstr. 3, 45141, Essen, Germany
[d] German Environment Agency, Wörlitzer Platz 1, 06844, Dessau-Roßlau, Germany
[e] State Agency for Nature, Environment and Consumer Protection North-Rhine Westphalia, 40208, Düsseldorf, Germany
[f] Saxon State Company for Environment and Agriculture, Altwahnsdorf 12, 01445 Radebeul, Germany
[g] Botanischer Garten und Botanisches Museum Berlin, Freie Universität Berlin, Königin–Luise–Straße 6–8, 14195, Berlin, Germany
[h] Referat U2 – Mikrobielle Ökologie, Bundesanstalt für Gewässerkunde, Am Mainzer Tor 1, 56068, Koblenz, Germany
[i] Luxembourg Institute of Science and Technology, 41 rue du Brill, 4422, Belvaux, Luxembourg
[j] University of Duisburg-Essen, Aquatic Ecology, Universitaetsstr. 5, 45141, Essen, Germany

ARTICLE INFO

Keywords:
Bioassessment
Biomonitoring
Regulation
Rivers

ABSTRACT

The ecological state of aquatic ecosystems is systematically monitored using various bioindicators in many countries worldwide. In the European Union, freshwater biomonitoring is the central component of the EU Water Framework Directive (WFD, 2000/60/EC) and currently based on morpho-taxonomic methods. DNA metabarcoding is a novel approach to assess the ecological state fast and efficiently based on organismal DNA signatures and thereby support and upscale biomonitoring. However, compliance of metabarcoding with existing morpho-taxonomic methods must be ensured prior to official implementation. Thus, this study, co-designed by research institutions and environmental agencies, explored necessary key parameters and performed method intercalibration for the implementation of metabarcoding into WFD assessments of running waters. We focussed on benthic invertebrates as the most commonly used bioindicators. We analysed 170 invertebrate samples collected as part of the German federal state WFD routine stream biomonitoring, first via microscopic determination and then using metabarcoding. Our goals were to quantify overlap in i) taxonomic composition and ii) ecological status derived with both methods. For this purpose, we established data harmonisation measures to integrate invertebrate metabarcoding data into the official national WFD classification modules considering abundance and presence/absence data. Our results revealed a high (ca. 70 %) overlap of bioindicator taxa found with both methods. Metabarcoding identified significantly more small invertebrate taxa and detected similar proportions of the important bioindicator 'EPT' taxa (mayflies, stoneflies, caddisflies). Despite deviations in some detected bioindicator taxa, the derived ecological status classes were highly correlated between methods, particularly after intercalibration ($R^2 = 0.74$, Spearman rho = 0.86). Regardless of whether we used abundance or presence/absence data, the resulting stream type classifications showed strong agreement. Thus, our study not only demonstrates the consistency of the methods for the stream types analysed but is also the first to operationalise a path to integration of metabarcoding data into the WFD assessment modules based on formal intercalibration guidelines.

* Corresponding authors.
E-mail addresses: macher@uni-trier.de (T.-H. Macher), florian.leese@uni-due.de (F. Leese).

https://doi.org/10.1016/j.watres.2024.122987
Received 16 September 2024; Received in revised form 12 December 2024; Accepted 14 December 2024
Available online 15 December 2024

1. Introduction

Environmental management in many countries worldwide involves biological monitoring of aquatic systems as part of regulatory programs to assess their ecological status. These monitoring activities typically focus on selected indicator taxa that reflect ecological integrity. In the European Union (EU), freshwater biomonitoring mainly serves the purposes of the EU Water Framework Directive (WFD, 2000/60/EC), which aims to achieve and maintain good chemical and ecological status or potential of EU water bodies by the year 2027 and requires accompanying monitoring activities to track progress in status improvement. The monitoring efforts under the WFD encompass more than 120,000 water bodies of rivers, lakes, transitional, and coastal waters (EEA 2018). Multiple bioindicator organism groups, referred to as "Biological Quality Elements" (BQEs), are examined including “phytoplankton”, “aquatic flora” (“macrophytes” and “phytobenthos”), “benthic invertebrates”, and “fishes”. By investigating these BQEs, a comprehensive understanding of the ecological status can be achieved. The obtained ecological status classes (ESC) range from "high" (1) over "good" (2), "moderate” (3), "poor” (4) to "bad” (5). The overall status of a water body is determined by the worst status of all BQEs investigated. Prior to the WFD, many member states (MSs) already had their own methods for assessing ecological status and many of these were modified to comply with the WFD requirements (Birk et al., 2012).

Accordingly, many different assessment methods are being used under the WFD, which reflects the natural variability of Europe’s waters but also the different monitoring traditions of the MSs. The methods used by MSs were intercalibrated to enhance comparability of the crucial class boundaries “high/good” and “good/moderate” (Birk et al. 2012; Poikane et al. 2014). Almost all methods rely on visual identification of specimens and derive or estimate abundances, the latter being a legal requirement of the WFD (Annex V). In practice, however, traditional monitoring methods prove challenging because identifying and counting organisms relies on taxonomically trained personnel, is time-consuming and thus costly. However, as budgets for monitoring programs face limitations and taxonomic expertise dwindles, there is a growing demand to enhance the efficiency of the monitoring process, whilst sustaining comparability, quality, and robustness (Hering et al. 2018).

DNA-based approaches, including DNA and environmental DNA (eDNA) metabarcoding, allow for rapid taxonomic assessments in monitoring programs (Hering et al. 2018; Pont et al. 2021). By analysing DNA fragments present in biological bulk samples, such as net or scrape samples, or water and sediment samples, it is possible to detect taxa representing different BQEs simultaneously (Taberlet et al. 2012). Furthermore, DNA metabarcoding offers advantages such as increased efficiency through high-throughput methods and workflow automation potentially enhancing cost-efficiencies, taxonomic coverage, and also spatio-temporal monitoring intensity compared to traditional methods (Elbrecht and Leese 2017; Vasselon et al. 2017; Macher et al. 2021b; Buchner et al., 2024). The general potential of DNA metabarcoding has been demonstrated for the WFD biological quality elements invertebrates (Elbrecht et al. 2017b; Kuntke et al. 2020; Meyer et al. 2020; Brantschen et al. 2021), fish (Muri et al. 2020; Macher et al. 2021c), and diatoms (Zimmermann et al. 2015; Rivera et al. 2018; Vasselon et al. 2019). However, DNA metabarcoding also has limitations, which must be considered for statutory biomonitoring.

One limitation is that DNA metabarcoding cannot, or only poorly, assess species abundance and biomass, since metabarcoding is influenced by PCR, primer and biomass bias (Elbrecht et al. 2017a; Krehenwinkel et al. 2017; Nichols et al. 2018; Muri et al. 2020). Another limitation is that DNA metabarcoding does not provide information about the age or life stage of specimens, which is a requirement for some WFD-compliant fish assessments. While information about the age of an organism can be retrieved under certain condition (Zhao et al., 2023), life stage will most likely not be routinely assessable with DNA-based monitoring methods in the foreseeable future.

Nonetheless, incorporation of DNA-based methods into WFD regulatory monitoring programs is under intense discussion given the maturity, reliability and cost-effectiveness of DNA metabarcoding (Hering et al. 2018; Vasselon et al. 2019; Pont et al. 2021; Buchner et al., 2024). Furthermore, the implementation of DNA metabarcoding could help to develop an even more comprehensive and efficient monitoring program for various BQEs especially in terms of spatial and temporal coverage, ultimately aiding in the conservation and management of aquatic environments also in view of other monitoring frameworks such as the Kunming-Montreal Global Biodiversity Framework, the European Habitats Directive, the Nature Restoration law and many further. While DNA metabarcoding has been applied in several studies and the general compliance has been shown (Kuntke et al. 2020; Brantschen et al. 2021; Múrria et al. 2024), methodological standardisation and intercalibration with the classical approaches has yet to be conducted prior to an adoption into WFD by 2027 (Hering et al. 2018). This requires detailed analyses of underlying statistical properties of the modules used to calculate ESCs and adjustments of analysis metrics.

This study was designed by researchers and practitioners from environmental agencies to evaluate practical feasibility and critical points of implementing DNA metabarcoding as a complementary method for water quality assessment in accordance with the WFD in Germany. To evaluate the reliability and robustness of DNA metabarcoding, we compared the DNA-based results obtained for the BQE “benthic invertebrates” to the taxa lists generated with morpho-taxonomic methods. First, we examined the community composition inferred with both methods to evaluate the resolution and richness of each approach. Second, we established intercalibration measures to facilitate the integration of DNA metabarcoding-derived data into the official WFD macroinvertebrate status class assessment approach “Perlodes”. Based on these results we assessed how differences in community composition of both methods translated into the inferred ESCs to evaluate the robustness and potential of DNA metabarcoding for WFD monitoring. For this, we operationalised the workflow for a revised national classification using the European Commission’s Guidance Document No. 30 on fitting new classification methods (EC 2015). The study ends with recommendations for practical implementation of DNA-based methods into the WFD and other regulations.

2. Methods

2.1. Sampling and morpho-taxonomic identification

The complete workflow is depicted in Fig. 1 and further details on the methods are provided in Appendix A. The benthic invertebrate sampling was conducted during WFD monitoring campaigns of 2020 and 2021 by expert field teams from North-Rhine Westphalia (Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen; LANUV), Saxony (Staatliche Betriebsgesellschaft für Umwelt und Landwirtschaft; BfUL), and Bavaria (Bayerisches Landesamt für Umwelt; LfU). In total, 170 sampling sites were selected based on river type and the previously reported ESC. We focussed on the most common stream types in Germany: type 5 (“coarse material-rich, siliceous low mountain streams”), 9 (“siliceous, fine- to coarse-material-rich low mountain streams.”), 14 (“sand-dominated lowland streams”), and 15 (“sand- and clay-dominated lowland rivers”). Samples were collected according to national/European standards (EN 17136), i.e., multi-habitat sampling according to Perlodes (invertebrates) sampling performed between March and August. Samples were preserved with 96 % denatured EtOH in a ratio of 1:3 sample to EtOH. Within 24 h the EtOH was exchanged to prevent dilution of the EtOH by water released from the conserved specimens. Traditional identification was conducted by experts from the LANUV, BfUL, and LfU. After traditional identification, samples were stored under dark conditions and were subsequently shipped to the University of Duisburg-Essen for the DNA metabarcoding analyses (refer

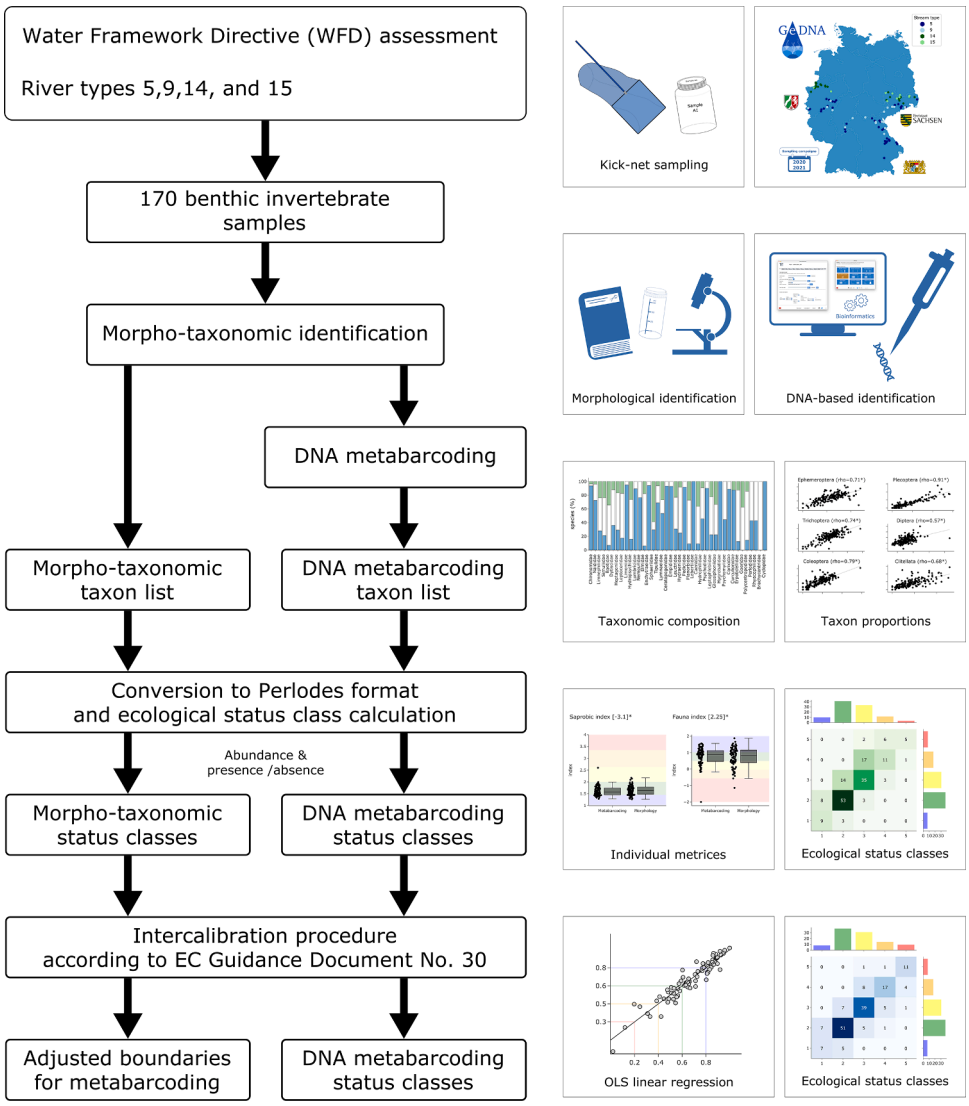


Fig. 1. Workflow of the benthic invertebrate comparison between DNA metabarcoding and morpho-taxonomic identification.

to Appendix A.1 for more details).

2.2. Lysis and DNA extraction

Initially, samples were homogenised in a kitchen blender (Mini Blender & Blender Smoothie, Homgeek). At this point, 12 negative controls (each containing 100 mL 96 % EtOH) per 84 samples were included. The lysis included a protocol, based on Proteinase K and bead-beating (Appendix A.2). All subsequent steps were carried out on a Biomek FXP Automated Workstation (Beckman Coulter), as described in Macher et al. 2021b. In short, DNA was extracted using a bead-based protocol in 96-well plate format. In total, two technical extraction replicates were conducted for each sample.

2.3. DNA amplification and sequencing

A two-step PCR approach was applied for amplifying the extracted DNA with the primer pair fwhF2 and fwhR2n, which targets a 205 bp fragment of the cytochrome c oxidase I subunit (COI; Vamos et al. 2017) and has been demonstrated as efficient primer pair for the detection of invertebrates (Buchner et al. 2024). In the 1st-step PCR each PCR plate was tagged with a unique combination of inline tags in the 1st-step primer. To remove primers and inhibitors the 1st-step product was purified using a bead-based clean-up protocol. In the 2nd-step PCR sequencing adapters and unique dual-twin indexes were added. Following the 2nd-step PCR, samples were normalised and size-selected (to remove primer dimers), again using a magnetic bead protocol. Then, the normalised samples were pooled into one library and concentrated. In total, two benthic invertebrate libraries (2020 and 2021 campaign) were sequenced on two separate HiSeq platform runs (2 × 150 bp) at Macrogen Europe (Appendix A.3).

2.4. Bioinformatic processing

Raw reads were processed with the APSCALE-GUI pipeline v1.2.0 (Macher et al. 2022). For the sequence processing, all settings were kept as default (Appendix A.4). The dataset was first filtered according to target fragment length thresholds (fwh2: 195–205 bp) and then sequences were clustered into 97 % OTUs (Operational Taxonomic Unit), which is the default clustering threshold for the fwh primer (Vamos et al. 2017). The software BOLDigger v2.1.1 (Buchner and Leese 2020) was used to assign taxonomy. The obtained taxonomy table was filtered according to the "BOLDigger" method and flags were added (Appendix Table C.1).

The taxonomy and read table were then converted to TaXon tables for downstream analyses in TaxonTableTools v1.5.1 (TTT, Macher et al.

2021a). Initially, technical replicates were merged, and only OTUs that were present in both PCR replicates were kept. To account for potential contamination, a strict read filter was applied where the sum of reads per OTU that were present in negative controls were subtracted from the reads per OTU of each sample. The dataset was subsequently filtered by taxonomic group. Here, only OTUs assigned to the phyla Annelida, Arthropoda, Cnidaria, Mollusca, Nematoda, Platyhelminthes, and Porifera with a similarity of ≥85 % were kept.

The morphology-based taxon lists obtained from the federal state agencies were converted to TaXon table format using a custom python script (Appendix B). Initially, all taxa were extracted from the taxon lists and the taxonomic backbone was downloaded via the GBIF API (GBIF.org (2023), GBIF Home Page. Available from: https:// w.gbif.org [03 July 2023]), including the phylum, class, order, family, genus, and species level. Thus, a backbone taxonomy reference table was created, which was manually curated for missing or erroneous information (i.e., lacking taxonomic level, spelling mistakes). Subsequently, the backbone taxonomy table was used to convert the traditional taxon lists into TaXon table format (Appendix Table C.3). The traditional taxon lists were filtered by taxonomic group according to the DNA metabarcoding dataset. Lastly, both tables were filtered to retain only samples with data available from both methods (Appendix Table C.2 and C.4). These tables were used for all subsequent analyses.

For all subsequent analyses, “species” refers to OTUs assigned to species level. “Unique taxa” refers to the best possible taxonomic assignment, which can include higher ranks if the similarity was low, or the assignment ambiguous. Unique taxa are essential for ESC calculation, where mixed taxonomic levels are used for practical reasons. Each assignment is counted only once, with OTUs having identical taxonomic assignments merged.

The datasets were converted to the Perlodes file format, which is employed for the WFD status class evaluation in Germany (https://gewaesser-bewertung-berechnung.de). To account for synonyms and spelling variations, dedicated conversion tables were established for Perlodes (Appendix Table C.5). The Perlodes conversion table includes taxa and their corresponding taxon ID numbers from the official German operational taxon list, as well as their corresponding entries in the GBIF taxonomy format. This allows for the conversion of special cases that cannot be derived from the information provided in the TaXon table format, such as undifferentiated higher taxonomic levels (e.g., "Naididae" converted to "Naididae/Tubificidae Gen. sp."). Then, both the DNA-based and traditional benthic invertebrate datasets were converted to the Perlodes format, using the WFD conversion module integrated in TaxonTableTools. All tables were exported in two versions, one retaining the abundance information (read or specimen counts) and the second as presence/absence format. For all datasets the ESC was determined using the Perlodes-Online module (v5.0.9) and results were exported for further analyses.

Lastly, all species that were exclusively detected with morphological identification were further investigated. Therefore, data was collected from BOLDsystems and GBIF (Appendix A.3). Based on the available information, we assessed potential bias on the absence of the specific species in the DNA metabarcoding data (Appendix Table C.6)

2.5. Statistical analyses

Initially, the number of shared and exclusive taxa (family, genus, species, and unique taxa) between the DNA-based and traditional benthic invertebrate dataset were calculated and visualised as Venn diagrams, using TaxonTableTools (Fig. 2A). Then, the number of initial taxa and the number of taxa remaining after the conversion to the Perlodes format for both the DNA-based and traditional datasets were calculated (Fig. 2B). The conversion to the WFD assessment formats used in Perlodes was required to harmonise the taxonomy used for the calculation of ESCs. The ESC range covers a range of classes, with "1" representing "high", followed by "2" for "good", "3" for "moderate", "4" for "poor", and "5" for "bad". To investigate the loss of taxonomic resolution during the Perlodes conversion, the number of unique taxa merged into a single entry was calculated. Both the initial taxa and the remaining taxa after the conversion were visualised in a bar chart. Additionally, the absolute and relative numbers of shared and exclusive species for all families (Fig. 3 and Appendix Table C.7) and orders (Appendix

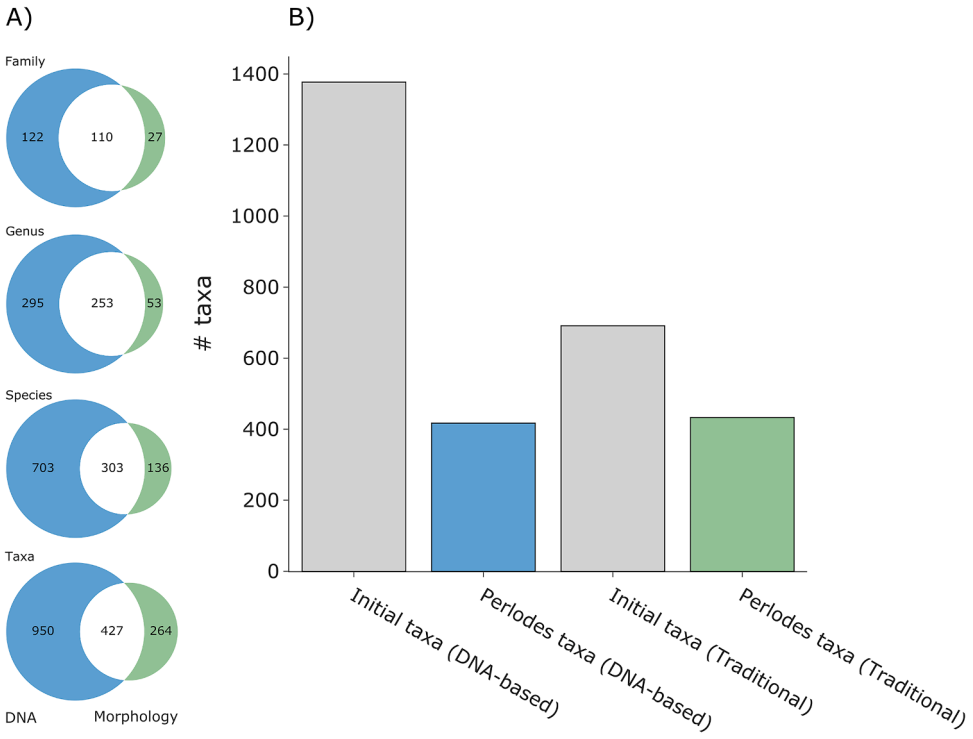


Fig. 2. Venn diagrams depicting shared and exclusive taxa between the DNA-based and traditional invertebrate dataset across different taxonomic levels. The bar charts show the number of initial taxa and the number of taxa remaining after the conversion to the Perlodes format for both datasets.

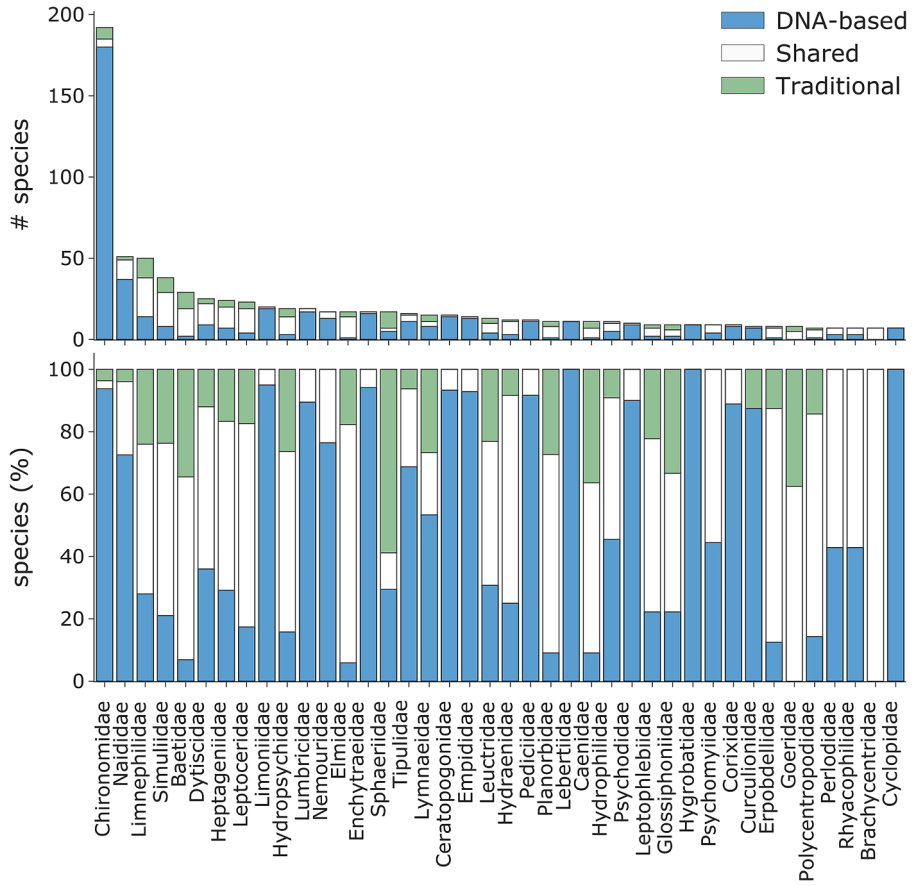


Fig. 3. Number (top) and proportion (bottom) of shared benthic invertebrate species per family between the morpho-taxonomic and DNA metabarcoding dataset.

Figure D.1 and Appendix Table C.8) between the two methods were calculated and visualised as bar charts.

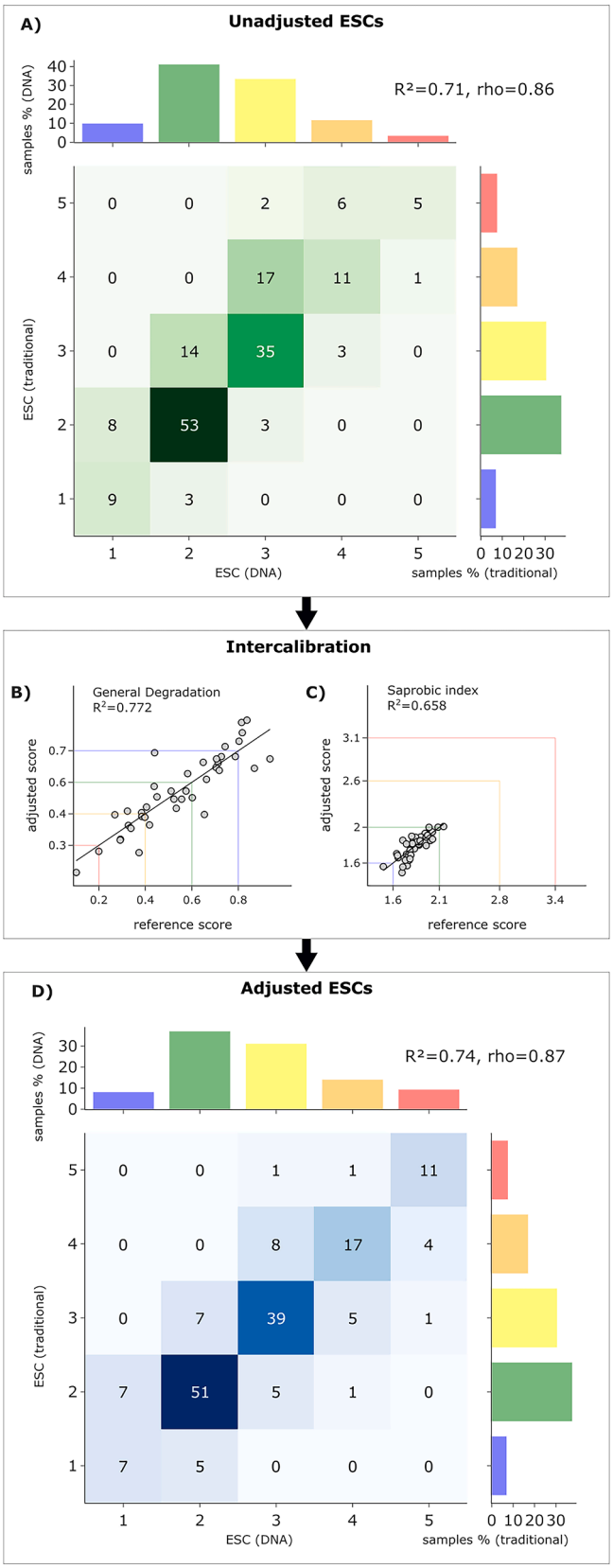
Several taxa were merged to a single entry during the conversion to the Perlodes format, which was depicted as bar charts for both datasets (Appendix Fig. D.2). The ESC results for the benthic invertebrate dataset were visualised in a matrix format to facilitate a comparative analysis of ESC values across corresponding samples. To explore the influence of abundance and presence/absence data, three separate matrix plots were generated (Fig. 4A and Appendix Fig. D.3). The first matrix plot compared the specimen abundance and read abundance data, while the second matrix plot focused on the taxon presence/absence datasets. Lastly, the third matrix plot involved a comparison between the abundance data derived from traditional datasets and the presence/absence data obtained from the DNA-based dataset. Spearman correlations were calculated for all three comparisons to evaluate the concordance between the DNA-based and traditional ESC results.

The individual metrics and modules used for the calculation of ESCs were investigated separately, according to the river type and taxonomic group. For the DNA-based and traditional results a total of 24 different river type specific indices were investigated. The ESC borders are specific to each river type (Appendix A.5). River type specific class changes were calculated and highlighted (Appendix Figs. D.4 and D.5). For each comparison, a Wilcoxon test was performed to investigate the concordance between the methods (Appendix Figure D.6 and D.7). To simplify comparisons, all indices were normalised to values between 0 and 100. Based on these normalised indices the deviation between the DNA metabarcoding and traditional monitoring results were calculated for each sample. Here, the mean deviation for each index was then plotted as a horizontal bar chart to identify the indices with the highest differences (Appendix Figure D.8A). Additionally, the proportion of samples with class changes for all investigated indices were calculated and plotted as horizontal bar chart (Appendix Figure D.8B).

To examine compositional differences among samples of different status classes inferred using DNA metabarcoding and morpho-taxonomic taxa lists, we conducted Principal Coordinate Analyses (PCoAs) using TaxonTableTools. These PCoAs were based on species-level Jaccard dissimilarity values. We calculated four PCoAs to assess compositional differences for river types 5 & 9 (Appendix Figure D.9A) and river types 14 & 15 (Appendix Figure D.9B) using both DNA metabarcoding and morpho-taxonomic species lists. Samples were color-coded according to their ESCs, and group differences were tested using an ANOSIM. To identify indicator species, linear regression was performed for each species' presence and principal coordinates 1 and 2 (PC1, PC2). Species that were significantly associated with PCs ($p \leq 0.05$) were visualized in the PCoAs as vectors, representing their directional influence on sample composition. This approach highlights species that contribute to compositional differences, calculated and visualized as implemented in TaxonTableTools.

2.6. Intercalibration procedure

To demonstrate that DNA metabarcoding is compliant with the WFD normative definitions and its class boundaries are in line with results from the traditional approach, an intercalibration exercise was performed by our team as described in the EC guidance document No. 30 (EC 2015) to fit a new or updated classification method. The intercalibration was conducted on the two main indices "Saprobic index" and "General Degradation" (MMI; multi-metric index) for each of the four river types. The "Acidification" index, which is relevant for river type 5, is based on five classes and could not be used in the intercalibration procedure. The individual indices used to calculate the MMI were also analysed, but due to insufficient coverage of the ecological range only



(caption on next column)

Fig. 4. Intercalibration procedure of DNA metabarcoding WFD data. Matrix comparing the ecological status classes (ESC) derived with the DNA metabarcoding and traditional, morphology-based methods, without intercalibration measures (A). The proportions of samples per status class are shown as bar charts. Spearman correlation was employed to assess the correlation between the two methods. Intercalibration was performed using ordinary least square (OLS) linear regression analyses to calculate the adjusted boundaries the General Degradation (B) and Saprobic index (C) DNA metabarcoding (here: example for river type 9) as requested by the official EU intercalibration manual (EC 2015). After the intercalibration ESC differences were more evenly distributed and the two methods showed higher congruence (D).

the calculated MMI was used for the intercalibration procedure.

For each metric the R^2 value was calculated, and ordinary least squares (OLS) regressions were used to calculate new class boundaries (Appendix E, Fig. 4B and 4C). In cases where the relation between the old and new method equates to $R^2 < 0.8$, the DNA metabarcoding boundaries were updated. Here, the intercept of the OLS regression line was used to define the new boundaries (see Fig. 4B and 4C). In cases of $R^2 \geq 0.8$ boundaries were not changed. Based on the updated boundaries new ESCs were calculated for the DNA metabarcoding dataset and plotted as ESC matrix (Fig. 4D). Additionally, mean weighted averages were calculated to display the distribution of samples across different ESCs were calculated for the unadjusted and adjusted boundaries (Appendix Figure D.10).

3. Results

3.1. DNA metabarcoding of benthic invertebrates

Sequencing for all 170 samples was successful, resulting in 739,258,391 raw reads. After quality filtering, 634,431,045 reads remained, which were clustered into 8,064 OTUs. Negative controls contained 787,175 reads (0.1 % of total reads) that were subtracted as cumulative read numbers per OTU and sample. The final dataset consisted of 3,846 OTUs and 590,095,956 reads after replicate merging and negative control subtraction. In total, 1,006 species were detected with metabarcoding and 439 with the traditional method. Of these, 303 (69 %) were shared between the two methods and 136 species were exclusive to the morphological identification (Fig. 2). The most species-rich taxa groups in the DNA-based dataset were (in decreasing order) Diptera (337 species), Trichoptera (123), Coleoptera (99), Ephemeroptera (59), Tubificidae (48), and Plecoptera (43). The most diverse taxa groups in the dataset derived with traditional methods were Trichoptera (113), Ephemeroptera (69), Coleoptera (54), Diptera (52), Pulmonata (20), and Plecoptera (20). On average 20.7 % of species per family were shared between both methods, while 16.5 % and 62.7 % were exclusive to morpho-taxonomic methods and DNA metabarcoding, respectively. Particularly for the families Chironomidae and Naididae, DNA metabarcoding detected substantially more exclusive species (180 and 37 species), compared to morpho-taxonomic methods (7 and 2), while only a small proportion was shared (5 and 12) (Fig. 3). Note here that identification according to Perlodes foresees identification of Chironomidae only to tribe level and identification of Oligochaeta at a maximum to family level, leading to the expected differences.

The 136 taxa exclusive to the morphological monitoring were split into 165 individual species that were queried against BOLDsystems' API. For 49 species no specific bias could be identified, so we suggest potential morphological misidentification, erroneous reference entries, primer bias, shorter marker fragment, incomplete lineage sorting, or hybridisation as potential bias source. Furthermore, for 50 species the discordance between the two approaches can be attributed to the usage of group names as result of several identification options in the morpho-taxonomic approach. Similar cases in the DNA metabarcoding dataset were trimmed to their most common recent ancestor. Another 20 species were not matched between the datasets due to the usage of synonyms. In

total 19 species were lacking both public and private hits and could thus not be detected with DNA metabarcoding, while another 5 species did not possess records from Central Europe. In total 4 cases of specific errors were identified. *Anomalopterygella chauviniana* was not detected, which can be explained by erroneous reference sequences (misidentification with *Drusus monticola*). *Chaetopteryx villosa* is known for its hybridisation and cannot be distinguished from *Chaetopteryx fusca*, based on the targeted marker. *Oligoneuriella rhenana* was suggested as a cryptic species complex (Laini et al. 2020) and can most likely not be identified using DNA metabarcoding. *Halesus tesselatus* was detected with both methods, however, spelling differed between the lists (*tesselatus* vs. *tessellatus*). Lastly, 18 species have records available in the BOLDsystems database, but were not identified using DNA metabarcoding and no further potential bias could be identified for these taxa.

3.2. Ecological status classification: Perlodes

The loss of taxonomic information during the conversion to the Perlodes format was higher for the DNA metabarcoding results, where originally 1,377 unique taxa were converted into 417 taxa (30 % retained) (Fig. 2). In the traditional dataset, 691 unique taxa were converted into 433 taxa (62 % retained). The taxonomic group with the highest loss rate (i.e., taxa that were trimmed to a higher taxonomic level) in the DNA-based dataset (Appendix Figure D.2) were the Chironomidae (Perlodes taxon: Chironomidae Gen. sp.) which consisted of 225 unique taxa, followed by Naididae and Tubificidae (Naididae/Tubificidae Gen. sp., 58 unique taxa), Heteroptera (Heteroptera Gen. sp., 37), and Oligochaeta (Oligochaeta Gen. sp., 35). Contrarily, the loss rate was lower for the traditional dataset, where 24 unique Chironomidae taxa were merged, followed by Naididae and Tubificidae (17 unique taxa), Limnephilidae (Limnephilidae Gen. sp., 15), Heteroptera (11), and *Simulium* (*Simulium* sp., 10).

The ESC estimates from Perlodes were overall slightly higher in the DNA-based dataset (Fig. 4), where 47 samples were rated higher, and 10 samples were rated lower compared to traditional methods. Interestingly, the pairwise ESCs were highly similar when comparing presence/absence data (rho = 0.87, $p \leq 0.05$; Appendix Figure D.3), abundance data (rho = 0.85, $p \leq 0.05$; Appendix Figure D.3), and morphological abundance and metabarcoding presence/absence data (rho = 0.86, $p \leq 0.05$; Fig. 4). Thus, the latter comparison was used for all downstream analyses since this represents the most-likely scenario how metabarcoding will be applied. A greater number of samples was assigned to status classes 1–3 in the DNA-based dataset (ESC high: 17, good: 70, moderate: 57) compared to the traditionally derived dataset (high: 12, good: 64, moderate: 52). Conversely, more samples were classified with status classes 4 and 5 in the traditional results (poor: 29, bad: 13), compared to the DNA metabarcoding results (poor: 20, bad: 6). Most samples shared the same ESC between methods (113 samples; 66.47 %), while 45 (26.47 %) and two samples (1.18 %) were classified better by one and two classes with DNA-based methods, respectively. Another 10 samples (5.88 %) were evaluated worse by one class.

Most differences in ESCs were observed for river types 14 and 15, for which 20 out of 37 and 6 out of 20 samples, respectively, were evaluated better with DNA metabarcoding, while only a single sample each was evaluated better with traditional methods (Appendix Figure D.4). This is also reflected in the highest differences in normalised values for these river types (Appendix Figure D.8A). For river types 5 (78 samples) and 9 (40 samples) a more homogeneous distribution of class deviations was observed, with 13 and 8 samples with better DNA metabarcoding status classes and 3 and 5 samples with better traditional status classes (Appendix Figures D.4, and D.8B). Overall, the highest differences for the individual metrics were generally observed for river types 14 and 15 and mostly indices that were based on absolute species numbers, such as the number of trichopteran taxa or number of EPTCBO taxa, relevant for river types 14 and 15 (Appendix Figures D.5). Consequently, also the river types that are more based on absolute species numbers showed the highest differences in ESCs, such as river types 14 (46.7 % class changes) and 15 (56.8 %), while river types 5 (20.5 %) and 9 (35 %) shower higher congruence. Despite these differences in exact ESC, all except two samples deviated by just one status class (Fig. 4).

Of the 18 relevant metrics, 12 showed significant differences when comparing the metabarcoding and morphology datasets (Wilcoxon test $p \leq 0.05$). For river type 5, significant differences were observed in the Saprobic index, which was lower on average for metabarcoding, and Fauna index, which was higher for metabarcoding (Appendix Figure D.6). River type 9 showed significantly lower Saprobic index values, proportions of EPT taxa, and proportions of metarhithral colonisers in the metabarcoding dataset (Appendix Figure D.7). Contrarily, the number of EPTCBO taxa was significantly higher in the metabarcoding dataset. River type 14 showed significantly higher numbers of Trichoptera taxa and proportions of EPT taxa, while the Saprobic index values were lower in the metabarcoding dataset (Appendix Figure D.6). Lastly, the metabarcoding samples of river type 15 showed significantly higher Fauna index values and significantly lower Saprobic index values (Appendix Figure D.7).

Principal Coordinate Analyses (PCoAs) showed significant compositional differences between the samples, with a general compositional shift from “high” to “bad” ecological status. This pattern was found in both subsets (i.e. for river types 5 & 9 as well as for 14 & 15), regardless of the identification methods (ANOSIM, $p \leq 0.001$). Group overlap existed across the ESC gradient as indicated by relatively low ANOSIM R values (ranging from 0.14 to 0.3; see Appendix Figure D.9), which, however were larger for DNA metabarcoding as compared to morpho-taxonomic lists.

3.3. Intercalibration procedure

The intercalibration of metric boundaries differed between the river types (see Appendix E). For river type 5 both the saprobic ($R^2 = 0.815$) and degradation index ($R^2 = 0.897$) did not require intercalibration. River type 9 required intercalibration for both the saprobic ($R^2 = 0.658$) and degradation index ($R^2 = 0.772$). Boundaries were also adjusted for river type 14, where particularly the saprobic index showed little concordance ($R^2 = 0.292$), compared to the degradation ($R^2 = 0.674$). The same pattern was observed for the saprobic ($R^2 = 0.16$) and degradation index ($R^2 = 0.758$) of river type 15. Across all the whole dataset more samples showed matching ESCs after the intercalibration procedure (113 to 125; Fig. 4C). Also, intercalibration of boundaries led to a small increase in both the R^2 value ($R^2 = 0.74$) and spearman correlation (rho = 0.87) compared to the non-adjusted analysis ($R^2 = 0.71$, rho = 0.86; Fig. 4A and 4D). The intercalibration success was more reflected in the distribution of ESC differences (Fig. 4D). While the unadjusted ESC differences were skewed towards a better evaluation with DNA metabarcoding (weighted average = −4.33), the ESC differences were more evenly distributed after the intercalibration (weighted average = −0.22).

4. Discussion

By adopting international standards (EN 17,136) along with supplementary guidance and training, our study proved the seamless adoption of DNA-conform sampling and shipping methods into the regular WFD monitoring, guaranteeing a 100 % processing rate in the laboratory. The analysis of 170 benthic invertebrate samples showed that despite method-specific differences in identified taxa and taxonomic resolution, the resulting ESCs for WFD-conform assessments were consistent between morphology-based and DNA metabarcoding analysis, even identical in two-thirds of the cases.

4.1. Reasons for differences in taxonomic lists from both methods

Generally, DNA metabarcoding detected more taxa compared to morphology-based methods. This pattern reflected in the detection of families (122 exclusive to metabarcoding, 11 shared, 27 exclusives to traditional methods), genera (295, 253, 53), and particularly on species level (703, 303, 136). Particularly when investigating the two datasets in detail, many assignments with discrepancies were observed. Overall, specifically two groups led to this substantial difference in detection: Chironomidae and Naididae, which where both detected in much higher resolution when using DNA metabarcoding. Chironomids ('non-biting midges') are renowned for their difficult morphological identification (Beermann et al. 2018) and have therefore been widely ignored in bioassessment metrics (Hering et al. 2018). Furthermore, 21 species were shared among the datasets but were not present in the same sample (Appendix Table C.9). This concerned several species of the stonefly genus *Leuctra* (i.e., *L. albida L. aurita*, and *L. digitata*) and *Simulium* (i.e., *S. argyreatum, S. aureum, S. intermedium, S. monticola*, and *S. tuberosum*). Another issue was the usage of special characters in the traditional monitoring dataset. Here, 44 taxa in total had special characters to include two or more potential species or groups, such as "*Potamophylax cingulatus / latipennis / luctuosus*" or "*Caenis pseudorivulorum* - Gruppe". These taxa aggregates are currently not comparable with the taxonomy used in the BOLDsystems database that needs distinct species names. Thus, such cases artificially increase discrepancies between the two methods. Another potential source of bias is the different identification of closely related species. For example, DNA metabarcoding detected the mollusc *Dreissena bugensis*, while traditional methods observed *Dreissena rostriformis* (single specimen). Both species occurred in the same sample ('501,580_Stever_NRW_2021') and are reported from Germany, rendering both assignments plausible. Also, both species have records available in BOLDsystems (*Dreissena bugensis*: 21 records, 1 BIN; *Dreissena rostriformis*: 48 records, 1 BIN) and can be genetically differentiated. Lastly, we observed several issues in the taxonomic assignment when using the BOLDsystems database as reference due to insufficient QA/QC measures. For example, as of the date of the taxonomic assignment for the DNA metabarcoding data, several published records of *Anomalopterygella chauviana* were incorrectly assigned as *Drusus monticola*, an alpine species not known to occur in North-Rhine Westphalia or Saxony. This incorrect database entry caused discrepancies between the two methods. Another issue is the usage of synonyms, which for example caused *Mystacides azurea* and *M. azureus* to be counted as separate species. These issues call for improved authoritative reference database curation.

Generally, it is important to note that neither of the two methods should be regarded as "truth" and differences in species-level identification are to be expected for known reasons. On one hand, DNA metabarcoding analyses are still affected by several challenges, such as potential primer bias, lack of reference barcodes, or incomplete and erroneous reference databases as exemplarily highlighted above. But also, additional taxa can be identified based on DNA based on molecules available as gut content or via juvenile specimen instars that cannot be identified using morphological methods. On the other hand, achieved morpho-taxonomic resolution may vary depending on the analyst's taxonomic expertise or lack of identification characteristics (e.g., in larvae of certain species), resulting in different taxonomic assignments of the same species. Furthermore, incomplete specimens can lack relevant identification characters, thereby hindering reliable identification.

Another simple reason for differences in the taxa lists is due to conversion of taxon names. Prior to ESC assessment, a conversion based on the official German operational taxon list had to be conducted, as only taxa listed on this list can be used for WFD assessment. This list currently includes 888 taxa with available autecological information specifically tailored for the calculation of ESCs within the Perlodes module. However, through the conversion process the taxonomic richness and resolution provided was significantly reduced because various detected taxa were not further considered at lower taxonomic levels like genus or species, such as Chironomidae (reduced to Chironomidae Gen sp.), Naididae, Tubificidae (Naididae/Tubificidae Gen. sp.), Heteroptera (Heteroptera Gen sp.), Oligochaeta (Gen sp.), and Enchytraeidae (Gen sp.). As an example, 225 individual taxa identified within Chironomidae were merged into a single entry to comply with the Perlodes taxa list. While it is understandable that the notoriously difficult to identify chironomids are not further considered in classical biomonitoring to reduce potential determination errors, it is well known that chironomids play a crucial role within aquatic ecosystems (Armitage et al. 2012) and are particularly useful bioindicators. Here, the integration of DNA metabarcoding presents a promising opportunity to overcome limitations that have historically restricted the study of pivotal taxonomic groups especially like chironomids within aquatic environments. Metrics to use such species could be further refined in future studies, as was also done e.g. for nematods (Sieriebriennikov et al. 2014), to improve our ability to effectively monitor biodiversity change and identify drivers of change to effectively safeguard freshwater ecosystems.

4.2. Presence/absence data provide reliable status class assessments

The WFD requests data on composition and abundance. As evidenced by our study and previous research (Elbrecht and Leese 2017; Elbrecht et al. 2017b; Zizka et al. 2020; Brantschen et al. 2021), DNA metabarcoding provides robust composition data for biodiversity assessments but quantitative inferences are imprecise at best due to many known issues (Thomas et al. 2016; Krehenwinkel et al. 2017; Ushio et al. 2018; Pont et al., 2023). However, while prevailing WFD assessment modules rely on specimen abundances to compute ESCs, such as the percentage of EPT taxa and saprobic index for benthic invertebrates, our study shows that presence absence data also provide reliable results with minor differences to quantitative (read amount) data. Notably, we revealed minimal discrepancies in ESCs when comparing presence/absence data (rho = 0.867, $p \leq 0.05$), abundance data (rho = 0.846, $p \leq 0.05$), and morphological abundance versus metabarcoding presence/absence data (rho = 0.859, $p \leq 0.05$). These findings align with earlier studies (Buchner et al. 2019), highlighting the applicability of presence/absence data with established indices, as utilised in the Perlodes modules. While the WFD currently requires abundance data, the question remains if the lack of these information when using DNA metabarcoding should be regarded as an exclusion criterion. Notably, between 2012 and 2018, 42 % of rivers remained unassessed for benthic invertebrates under the WFD (EEA 2018). Sole reliance on morpho-taxonomic methods is unlikely to bridge this gap. Thus, DNA-based methods should be regarded as complementary tools to assess surface waters that would otherwise remain unevaluated. While abundance data might not be collected, ESCs derived from DNA metabarcoding promise high concordance with status classes calculated based on abundance data. Generating data for these missing sites outweighs the absence of abundance data, given the importance of obtaining information about these previously unassessed locations.

4.3. Comparison of ecological status classes and indices

Our comparative analysis showed good concordance between DNA metabarcoding and traditional methods in determining ESCs using benthic invertebrates as bioindicators. While differences emerged in individual evaluation modules, the overall assignment of ESCs remained largely comparable. Notably, more than 65 % of samples exhibited identical ESCs across both methodologies, with only two samples deviating by two status classes. For the 170 samples analysed, DNA metabarcoding tended to result in a better ecological state (26.5 %) compared to traditional methods (5.9 %). It is important to consider that the here conducted methodological comparison did neither include further quality control steps nor intercalibration measures prior to the data analysis. Both methods were applied by independent experts without

prior inter-methodological calibrations to improve the concordance between the methods. In this context the observed concordance is to be evaluated even higher, especially when considering already large deviations within methodological workflows on their own. For example, Haase et al. 2010 reported a 35 % disparity between modules and a 16 % variance in final ESC assignment during a QA/QC audit of WFD assessment monitoring reliant on morpho-taxonomic identification of benthic invertebrates. Considering this acknowledged margin of error inherent in both morpho-taxonomic and DNA-based identifications and the strong agreement observed in our study, the process of intercalibration and devising correction factors for DNA metabarcoding assessments is a crucial next step.

4.4. Moving forward: Standardisation and intercalibration

The intercalibration of traditional methods used for the WFD assessment of the member states was an essential task to allow for partial regionally comparable status class assessments (Birk et al. 2012; Poikane et al. 2014). Therefore, a detailed intercalibration procedure of the national monitoring methods resulted in considerable fitting procedures (EC 2015). Based on the intercalibration criteria already over 400 national monitoring methods were evaluated and intercalibrated to meet the WFD assessment standards. Here we demonstrate a first successful pilot intercalibration exercise between DNA metabarcoding and traditional WFD data based on the official European Commission's Guidance Document. It was possible to adjust the classification method based on metabarcoding to the national classification method. These adjustments not only helped to increase the concordance between the two methods, but more importantly led to a more even distribution of class differences, which is in the range of the expected status class variation when assuming Poisson-distributed abundance data (Buchner et al., Fig. S3), or when transforming abundance data into presence-absence data (Buchner et al. 2019). After the boundary adjustment, more than 73.5 % of samples showed identical ESCs and 23.5 % (−1) and 11.2 % (+1) differed by only one ESC. However, the intercalibration could not be performed for all indices used in the ESC calculation (mainly MMI indices), due to insufficient sample size (river types 14 and 15) and insufficient coverage of ecological ranges for the specific indices. Nevertheless, we here demonstrate the path to successful intercalibration procedure, which can inform future national and international intercalibration tasks, covering further indices and river types and facilitate the official intercalibration exercises required to be done by the formal entities in charge of method acceptance and intercalibration in the EU, i.e. the working group ECOSTAT of the Common Implementation Strategy (CIS).

To ensure quality among DNA metabarcoding studies, standardisation of the method also becomes essential from sampling and storage, DNA extraction, amplification to sequencing and bioinformatic data processing. While there is no single best-practice workflow and many methods can lead to similar results there are obvious minimum standards along the process chain that must be outlined as part of formal standards, such as avoidance of cross contamination, adequate preservation, suitable primer choice and sequencing depth and in particular curated reference libraries. For Quality Control, Blackman et al. (2019) proposed the systematic use of reference materials for DNA metabarcoding. Similar quality control procedures are well established for diatom identification under the WFD in France and in environmental chemistry. For molecular BQE monitoring, laboratories would be required to demonstrate the performance of their custom lab and bioinformatic pipeline with calibration standards, e.g., mock communities of known composition. Only if a sufficient number of species in the mock community is detected, the data would be considered acceptable for WFD assessments. In this way, the focus would be set on the plausibility of the data obtained in the analyses rather than on standardising the diversity of approaches to selected methods (see Leese et al. 2023). Several national and international guidance documents and first official standards are already being established to facilitate the adoption of DNA-based methods, including the environmental DNA sampling and experiment manual (The eDNA Society), eDNA analysis for Great Crested Newt (Natural England), CEN/TR 17245, Water quality - Technical report for the routine sampling of benthic diatoms from rivers and lakes adapted for metabarcoding analyses, EN 17136 for DNA-conform sampling and storage or benthic invertebrates, as well as first ISO standards, e.g., ISO 21286:2019 (Soil quality - Identification of ecotoxicological test species by DNA barcoding) and ISODIS 17805 on sampling, capture and preservation of environmental DNA from water.

Furthermore, to ensure accurate reporting and comparability of species data in line with the WFD, the development of standardised formats and conversion tools are required to allow for a standardised translation of DNA metabarcoding data into the required input format. TaxonTableTools (Macher et al. 2021a) already provides a module that automatically conducts the required formatting and harmonisation tasks. Here, the conversion tool works reliably for the conversion of benthic invertebrate dataset for Perlodes. Until the implementation of DNA metabarcoding into the WFD, the agreement on a set of reliable conversion tools with updated taxonomic backbones will be required for all modules of all WFD member states. Next to curation and automated approaches there is a need to make taxon data available in a FAIR manner, i.e., to have the data findable, accessible, interoperable, and reusable.

5. Conclusions and outlook

In the context of WFD routine monitoring in Germany our study employed a co-designed approach to compare traditional bioassessment methods with DNA metabarcoding. This allowed us to validate several key criteria for successful implementation, as outlined by Hering et al. (2018). First, we showed that sampling can be seamlessly integrated into routine monitoring with only minor adjustments in preservation and sample handling. Second, we demonstrated that DNA-based and traditional status class assessment yield consistent results despite obvious and expected differences in taxa lists. While different methods will inevitably yield partly differing taxa lists, further methodological advancements - especially in reference database quality – will further improve the already good correlations. Third, our data demonstrated that species abundance has a minor impact on ESC assessment. This suggests that data solely on the presence or absence of species can yield comparable results as requested by WFD.

Until the end of the fourth River Basin Management Cycle in 2027 traditional morpho-taxonomic methods will continue to be used as the only accepted method for WFD BQE assessment. Lastly, we report a first successful intercalibration pilot for DNA metabarcoding WFD data, which increased the concordance with traditional methods and can inform future intercalibration tasks. Overall, the future potential and current advantages of DNA-based methods (either eDNA based or bulk sample based) speak for their implementation as complimentary or even alternative methods. Thus, also other EU legislation and policy, such as the "EU Regulation 1143/2014 on Invasive Alien Species", the EU Habitats Directive (Council Directive 92/43/EEC, 1992), and the Biodiversity strategy for 2030 (EC, 2020), can also be supported by DNA-based methods. But even for the WFD we must acknowledge that a substantial portion of surface waters, ranging from 24 % to 97 %, relevant for the WFD monitoring remained unassessed for certain BQEs (EEA 2018). This lack of information hinders our understanding of the ecological state of many freshwater systems in space and time. More efficient and scalable methods, such as DNA-based approaches, can help close these knowledge gaps, leading to better protection and management of surface waters. Implementing DNA-based approaches and embracing their potential for cost-effectiveness and scalability is crucial for advancing our understanding of aquatic ecosystems and supporting effective environmental conservation efforts. To achieve this, it is now essential to establish the necessary national and transnational pathways

of intercalibration and standardisation to allow the uptake of the methods at international scale.

Data accessibility

The raw data were deposited at the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) under the accession number PRJEB83622.

CRediT authorship contribution statement

Till-Hendrik Macher: Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Arne J. Beermann:** Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Jens Arle:** Writing – review & editing, Validation, Supervision, Project administration, Conceptualization. **Julia Foerster:** Writing – review & editing, Validation, Resources, Data curation. **Matthias Greyer:** Writing – review & editing, Validation, Resources, Data curation. **Demetrio Mora:** Writing – original draft, Validation, Methodology, Data curation, Conceptualization. **Jan Koschorreck:** Writing – review & editing, Validation, Supervision, Project administration, Conceptualization. **Peter Rolauffs:** Writing – review & editing, Validation, Data curation. **Anne Rother:** Writing – review & editing, Validation, Resources, Data curation. **Susanne Schüler:** Writing – review & editing, Validation, Resources, Data curation. **Jonas Zimmermann:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization. **Daniel Hering:** Writing – original draft, Validation, Funding acquisition, Data curation, Conceptualization. **Florian Leese:** Writing – original draft, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to the field teams and taxonomic experts for the WFD sampling and morphological identification. Special thanks to the advisory board of the GeDNA project, in particular Dr. Gabriele Eckartz-Vreden and Kristin Stolberg for their support in planning, Lina Frank and Sören Köhler for sample homogenisation, and Charlotte Frie for her valuable support in laboratory processing. The German Federal Environment Agency funded the REFOPLAN project GeDNA (FKZ 3719 24 2040) from 2019 to 2023.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2024.122987.

Data availability

Data accessibility statement can be found in the manuscript.

References

Armitage, P.D., Pinder, L.C., Cranston, P.S., 2012. The Chironomidae: Biology and Ecology of Non-Biting Midges. Springer Science & Business Media, p. 579.

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. Environ. Sci. Eur. 30, 26. https://doi.org/10.1186/s12302-018-0157-x.

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., Solimini, A., van de Bund, W., Zampoukas, N., Hering, D., 2012. Three hundred ways to assess Europe's surface waters: an almost complete overview of biological methods to implement the Water Framework Directive. Ecol. Indic. 18, 31–41. https://doi.org/10.1016/j.ecolind.2011.10.009.

Blackman, R.C., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., Egeter, B., Elbrecht, V., Filipe, A.F., Jones, J.I., Macher, J., Majaneva, M., Martins, F.M.S., Múrria, C., Meissner, K., Pawlowski, J., Yáñez, P.L.S., Zizka, V.M.A., Leese, F., Price, B., Deiner, K., 2019. Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. Metabarcod. Metagenom. 3, e34735. https://doi.org/10.3897/mbmg.3.34735.

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, e0257510. https://doi.org/10.1371/journal.pone.0257510.

Buchner, D., Leese, F., 2020. BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. Metabarcod. Metagenom. 4, e53535. https://doi.org/10.3897/mbmg.4.53535.

Buchner, D., Beermann, A.J., Laini, A., Rolauffs, P., Vitecek, S., Hering, D., Leese, F., 2019. Analysis of 13,312 benthic invertebrate samples from German streams reveals minor deviations in ecological status class between abundance and presence/absence data. PLoS. One 14, e0226547. https://doi.org/10.1371/journal.pone.0226547.

Buchner, D., Sinclair, J.S., Ayasse, M., Beermann, A.J., Buse, J., Dziock, F., Enss, J., Frenzel, M., Hörren, T., Li, Y., Monaghan, M.T., Morkel, C., Müller, J., Pauls, S.U., Richter, R., Scharnweber, T., Sorg, M., Stoll, S., Twietmeyer, S., Weisser, W.W., Wiggering, B., Wilmking, M., Zotz, G., Gessner, M.O., Haase, P., Leese, F., 2024. Upscaling biodiversity monitoring: metabarcoding estimates 31,846 insect species from Malaise traps across Germany. Mol. Ecol. Resour. 25, e14023. https://doi.org/10.1111/1755-0998.14023.

EC, 2015. Procedure to Fit New or Updated Classification Methods to the Results of a Completed Intercalibration Exercise. Guidance Document No 30. Publications Office, LU. Available from: https://data.europa.eu/doi/10.2779/158259 (November 21, 2024).

EC, 2020. EU Biodiversity Strategy for 2030. Bringing nature back into our lives. Brussels 20.5.2020 COM(2020) 380 final.

EEA (2018) European waters—Assessment of status and pressures. 2018, *EEA Report No* **7**/2018. Publications Office of the European Union Luxembourg.

Elbrecht, V., Leese, F., 2017. Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Front. Environ. Sci. 5. https://doi.org/10.3389/fenvs.2017.00011.

Elbrecht, V., Peinert, B., Leese, F., 2017a. Sorting things out: assessing effects of unequal specimen biomass on DNA metabarcoding. Ecol. Evol. 7, 6918–6926. https://doi.org/10.1002/ece3.3192.

Elbrecht, V., Vamos, E.E., Meissner, K., Aroviita, J., Leese, F., 2017b. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods Ecol. Evol. 8, 1265–1275. https://doi.org/10.1111/2041-210X.12789.

Haase, P., Pauls, S.U., Schindehütte, K., Sundermann, A., 2010. First audit of macroinvertebrate samples from an EU Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. J. N. Am. Benthological Soc. 29, 1279–1291. https://doi.org/10.1899/09-183.1.

Hering, D., Borja, A., Jones, J.I., Pont, D., Boets, P., Bouchez, A., Bruce, K., Drakare, S., Hänfling, B., Kahlert, M., Leese, F., Meissner, K., Mergen, P., Reyjol, Y., Segurado, P., Vogler, A., Kelly, M., 2018. Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water. Res. 138, 192–205. https://doi.org/10.1016/j.watres.2018.03.003.

Krehenwinkel, H., Wolf, M., Lim, J.Y., Rominger, A.J., Simison, W.B., Gillespie, R.G., 2017. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Sci. Rep. 7, 17668. https://doi.org/10.1038/s41598-017-17333-x.

Kuntke, F., de Jonge, N., Hesselsøe, M., Lund Nielsen, J., 2020. Stream water quality assessment by metabarcoding of invertebrates. Ecol. Indic. 111, 105982. https://doi.org/10.1016/j.ecolind.2019.105982.

Laini, A., Beermann, A.J., Bolpagni, R., Burgazzi, G., Elbrecht, V., Zizka, V.M., Leese, F., Viaroli, P., 2020. Exploring the potential of metabarcoding to disentangle macroinvertebrate community dynamics in intermittent streams. Metabarcod. Metagenom. 4, e51433.

Leese, F., Woppowa, L., Bálint, M., Höss, S., Krehenwinkel, H., Lötters, S., Meissner, K., Nowak, C., Rausch, P., Rduch, V., Rulik, B., Weigand, A.M., Zimmermann, J., Koschorrek, J., Zughart, W., 2023. DNA-basierte Biodiversitätsanalysen im Natur- und Umweltschutz: welche Optionen haben wir für eine Standardisierung? Eine Handlungsempfehlung aus Forschung und Praxis. https://doi.org/10.19217/skr666.

Macher, T.-H., Beermann, A.J., Leese, F., 2021a. TaxonTableTools: a comprehensive, platform-independent graphical user interface software to explore and visualise DNA metabarcoding data. Mol. Ecol. Resour. 21, 1705–1714. https://doi.org/10.1111/1755-0998.13358.

Macher, T.-H., Buchner, D., Leese, F., 2022. APSCALE: advanced pipeline for simple yet comprehensive analyses of DNA metabarcoding data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btac588 btac588.

Macher, T.-H., Buchner, D., Beermann, A.J., Werner, M.-T., Leese, F., 2021b. Standardized high-throughput biomonitoring using DNA metabarcoding: strategies for the adoption of automated liquid handlers. Environ. Sci. Ecotechnol. 8, 100122. https://doi.org/10.1016/j.ese.2021.100122.

Macher, T.-H., Schütz, R., Arle, J., Beermann, A.J., Koschorreck, J., Leese, F., 2021c. Beyond fish eDNA metabarcoding: field replicates disproportionately improve the

detection of stream associated vertebrate species. Metabarcod. Metagenom. 5, e66557. https://doi.org/10.3897/mbmg.5.66557.
Meyer, A., Boyer, F., Valentini, A., Bonin, A., Ficetola, G.F., Beisel, J.-N., Bouquerel, J., Wagner, P., Gaboriaud, C., Leese, F., Dejean, T., Taberlet, P., Usseglio-Polatera, P., 2020. Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: testing different combinations of markers and strategies of data filtering. Mol. Ecol. https://doi.org/10.1111/mec.15723 n/a.
Muri, C.D., Handley, L.L., Bean, C.W., Li, J., Peirson, G., Sellers, G.S., Walsh, K., Watson, H.V., Winfield, I.J., Hänfling, B., 2020. Read counts from environmental DNA (eDNA) metabarcoding reflect fish abundance and biomass in drained ponds. Metabarcod. Metagenom. 4, e56959. https://doi.org/10.3897/mbmg.4.56959.
Múrria, C., Wangensteen, O.S., Somma, S., Väisänen, L., Fortuño, P., Arnedo, M.A., Prat, N., 2024. Taxonomic accuracy and complementarity between bulk and eDNA metabarcoding provides an alternative to morphology for biological assessment of freshwater macroinvertebrates. Sci. Total Environ., 173243 https://doi.org/10.1016/j.scitotenv.2024.173243.
Nichols, R.V., Vollmers, C., Newsom, L.A., Wang, Y., Heintzman, P.D., Leighton, M., Green, R.E., Shapiro, B., 2018. Minimizing polymerase biases in metabarcoding. Mol. Ecol. Resour. 18, 927–939. https://doi.org/10.1111/1755-0998.12895.
Poikane, S., Zampoukas, N., Borja, A., Davies, S.P., van de Bund, W., Birk, S., 2014. Intercalibration of aquatic ecological assessment methods in the European Union: lessons learned and way forward. Environ. Sci. Policy. 44, 237–246. https://doi.org/10.1016/j.envsci.2014.08.006.
Pont, D., Meulenbroek, P., Bammer, V., Dejean, T., Erős, T., Jean, P., Lenhardt, M., Nagel, C., Pekarik, L., Schabuss, M., Stoeckle, B.C., Stoica, E., Zornig, H., Weigand, A., Valentini, A., 2023. Quantitative monitoring of diverse fish communities on a large scale combining eDNA metabarcoding and qPCR. Mol. Ecol. Resour. 23, 396–409. https://doi.org/10.1111/1755-0998.13715.
Pont, D., Valentini, A., Rocle, M., Maire, A., Delaigue, O., Jean, P., Dejean, T., 2021. The future of fish-based ecological assessment of European rivers: from traditional EU water framework directive compliant methods to eDNA metabarcoding-based approaches. J. Fish. Biol. 98, 354–366. https://doi.org/10.1111/jfb.14176.
Rivera, S.F., Vasselon, V., Jacquet, S., Bouchez, A., Ariztegui, D., Rimet, F., 2018. Metabarcoding of lake benthic diatoms: from structure assemblages to ecological assessment. Hydrobiologia 807, 37–51. https://doi.org/10.1007/s10750-017-3381-2.
Sieriebriennikov, B., Ferris, H., de Goede, R.G.M., 2014. NINJA: an automated calculation system for nematode-based biological monitoring. Eur. J. Soil. Biol. 61, 90–93. https://doi.org/10.1016/j.ejsobi.2014.02.004.
Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., Willerslev, E., 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol. Ecol. 21, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.
Thomas, A.C., Deagle, B.E., Eveson, J.P., Harsch, C.H., Trites, A.W., 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Mol. Ecol. Resour. 16, 714–726. https://doi.org/10.1111/1755-0998.12490.
Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., Yamanaka, H., Minamoto, T., Kondoh, M., 2018. Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing. Metabarcod. Metagenom. 2, e23297. https://doi.org/10.3897/mbmg.2.23297.
Vamos, E., Elbrecht, V., Leese, F., 2017. Short COI markers for freshwater macroinvertebrate metabarcoding. Metabarcoding. Metagenom. 1, e14625. https://doi.org/10.3897/mbmg.1.14625.
Vasselon, V., Rimet, F., Tapolczai, K., Bouchez, A., 2017. Assessing ecological status with diatoms DNA metabarcoding: scaling-up on a WFD monitoring network (Mayotte island, France). Ecol. Indic. 82, 1–12. https://doi.org/10.1016/j.ecolind.2017.06.024.
Vasselon, V., Rimet, F., Domaizon, I., Monnier, O., Reyjol, Y., Bouchez, A., 2019. Assessing pollution of aquatic environments with diatoms' DNA metabarcoding: experience and developments from France Water Framework Directive networks. Metabarcod. Metagenom. 3, e39646. https://doi.org/10.3897/mbmg.3.39646.
Zhao, B., van Bodegom, P.M., Trimbos, K.B., 2023. Environmental DNA methylation of *Lymnaea stagnalis* varies with age and is hypermethylated compared to tissue DNA. Mol. Ecol. Resour. https://doi.org/10.1111/1755-0998.13691, 23, 81-91.
Zimmermann, J., Glöckner, G., Jahn, R., Enke, N., Gemeinholzer, B., 2015. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol. Ecol. Resour. 15, 526–542. https://doi.org/10.1111/1755-0998.12336.
Zizka, V.M.A., Geiger, M.F., Leese, F., 2020. DNA metabarcoding of stream invertebrates reveals spatio-temporal variation but consistent status class assessments in a natural and urban river. Ecol. Indic. 115, 106383. https://doi.org/10.1016/j.ecolind.2020.106383.