PeerJ Preprints NOT PEER-REVIEWED

1 **Title: Assessing strengths and weaknesses of DNA metabarcoding based**
2 **macroinvertebrate identification for routine stream monitoring**

3

4 **Running Title (45 char max):** Stream bioassessment using DNA metabarcoding

5 **Word count:** 7766

6 **Authors:** Vasco Elbrecht[1], Ecaterina Edith Vamos[1], Kristian Meissner[2], Jukka Aroviita[3], Florian Leese[1,4*]

7

8 Affiliations:

9 1) Aquatic Ecosystem Research, Faculty of Biology, University of Duisburg-Essen, Universitätsstraße 5, 45141 Essen,
10 Germany

11 2) Finnish Environment Institute, Freshwater Centre, Survontie 9 A, 40500 Jyväskylä, Finland

12 3) Finnish Environment Institute, Freshwater Centre, PO Box 413, 90014 Oulu, Finland

13 4) Centre for Water and Environmental Research (ZWU) Essen, University of Duisburg-Essen, Universitätsstraße 2, 45141
14 Essen, Germany

15

16 *Corresponding author: florian.leese@uni-due.de, phone: +49.201-1834053

17

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

18 **Abstract**

19

20 1) DNA metabarcoding holds great promise for the assessment of macroinvertebrates in stream ecosystems. However, few
21 large-scale studies have compared the performance of DNA metabarcoding with that of routine morphological identification.
22 2) We performed metabarcoding using four primer sets on macroinvertebrate samples from 18 stream sites across Finland.
23 The samples were collected in 2013 and identified based on morphology as part of a Finnish stream monitoring program.
24 Specimens were morphologically classified, following standardised protocols, to the lowest taxonomic level for which
25 identification was feasible in the routine national monitoring.
26 3) DNA metabarcoding identified more than twice the number of taxa than the morphology-based protocol, and also yielded
27 a higher taxonomic resolution. For each sample, we detected more taxa by metabarcoding than by the morphological
28 method, and all four primer sets exhibited comparably good performance. Sequence read abundance and the number of
29 specimens per taxon (a proxy for biomass) were significantly correlated in each sample, although the adjusted R^2 were low.
30 With a few exceptions, the ecological status assessment metrics calculated from morphological and DNA metabarcoding
31 datasets were similar. Given the recent reduction in sequencing costs, metabarcoding is currently approximately as
32 expensive as morphology-based identification.
33 4) Using samples obtained in the field, we demonstrated that DNA metabarcoding can achieve comparable assessment
34 results to current protocols relying on morphological identification. Thus, metabarcoding represents a feasible and reliable
35 method to identify macroinvertebrates in stream bioassessment, and offers powerful advantage over morphological
36 identification in providing identification for taxonomic groups that are unfeasible to identify in routine protocols. To unlock
37 the full potential of DNA metabarcoding for ecosystem assessment, however, it will be necessary to address key problems
38 with current laboratory protocols and reference databases.

39

40

41 **Keywords:** biomass bias, high-throughput sequencing, macroinvertebrates, metabarcoding, ecological status

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

42 ## Introduction

43 The abundance and diversity of macroinvertebrates are used as key biological quality indicators in national and international
44 aquatic biomonitoring programs, which employ a variety of bioassessment protocols (Birk *et al.* 2012). In all current
45 protocols, however, biological quality components such as macroinvertebrates, diatoms, macroalgae, and fish, are identified
46 based only on morphological properties. Among benthic macroinvertebrates, the orders Ephemeroptera, Plecoptera,
47 Trichoptera, and Diptera are often regarded most sensitive to pollution and are thus ideal indicators of anthropogenic
48 stressor effects on stream ecosystems (Resh & Unzicker 1975; Buss *et al.* 2015). Unfortunately, the identification of benthic
49 taxa to species or even genus level is often difficult or impossible, and the accuracy of such identification is highly
50 dependent on the researcher's experience; consequently, misidentification is frequent (Sweeney *et al.* 2011). Accordingly,
51 classification is often performed only to a higher taxonomic level. However, the species within a higher taxonomic group
52 may exhibit diverse responses to stress (Macher *et al.* 2016), and these differences can go unnoticed in studies with low
53 taxonomic resolution. Misidentification, low comparability, and limited taxonomic resolution for difficult groups, such as
54 chironomids, can lead to inaccurate assessments and potentially to the mismanagement of stream ecosystems (Stein *et al.*
55 2013a). Moreover, the use of human experts for morphological identification is time-consuming and therefore expensive
56 (Yu *et al.* 2012; Aylagas *et al.* 2014).

57 In recent years, DNA-based taxon identification has emerged as a potential alternative to morphological methods.
58 The first DNA-based case studies highlighted the potential application of these methods to the assessment of freshwater
59 macroinvertebrates (Hajibabaei *et al.* 2011; Carew *et al.* 2013; Elbrecht & Leese 2015; 2016b). In particular, DNA
60 barcoding has often been advocated as a useful tool for ecosystem monitoring and assessment (Baird & Sweeney 2011;
61 Baird & Hajibabaei 2012; Taberlet *et al.* 2012). In metabarcoding, DNA is extracted from bulk samples, a standardised
62 marker gene amplified and sequenced using high throughput sequencing followed by comparison against reference
63 databases allowing for cost-efficient and reliable community assessments (Ratnasingham & Hebert 2007; Hajibabaei *et al.*
64 2011; Taberlet *et al.* 2012). Although several studies have established multiple benefits of DNA-based monitoring using
65 DNA metabarcoding, additional large-scale studies of complete freshwater macroinvertebrate samples are needed to
66 validate and improve metabarcoding protocols for routine monitoring. In marine, freshwater, and terrestrial ecosystems,
67 complete samples of arthropods and diatoms have been processed (Ji *et al.* 2013; Gibson *et al.* 2014; Zimmermann *et al.*
68 2014; Leray & Knowlton 2015) and used to obtain assessment metrics (Aylagas *et al.* 2016a). However, DNA
69 metabarcoding studies of complete macroinvertebrate samples from freshwater ecosystems have often been limited to one
70 or two sampling sites (Hajibabaei *et al.* 2011; 2012) or selected taxon groups (Carew *et al.* 2013). The only large-scale

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

71 study of 24 Canadian macrozoobenthos samples (Gibson *et al.* 2015) demonstrated that DNA metabarcoding outperforms
72 family- and order-level approaches to morphological identification. Although these results are promising, it should be noted
73 that in most European monitoring programs, taxa are identified generally to species level. For DNA metabarcoding to be
74 applied to routine stream monitoring, protocols for DNA-specific macrozoobenthos sampling and laboratory must be further
75 developed, optimised, and validated. We recently explored primer bias and tissue extraction protocols using a one-step PCR
76 metabarcoding protocol on the Illumina MiSeq sequencer and then employed this technique to examine mock invertebrate
77 samples of known composition (Elbrecht & Leese 2015). In addition, because we found that primer design is a critical
78 component for species detection, we developed primer sets specifically targeting freshwater macroinvertebrates (Elbrecht &
79 Leese 2016a,b). Although these BF/BR primers worked well in mock communities and initial tests based on two stream
80 benthos samples, they have not been tested in a larger-scale biomonitoring context (Elbrecht & Leese 2016b; Elbrecht *et al.*
81 2016). Further, the reliability and completeness of available reference data, e.g., the BOLD database for freshwater
82 macroinvertebrates (Ratnasingham & Hebert 2007), has not been fully explored in a metabarcoding context. Finally,
83 laboratory constraints specific to organisms and stream ecosystems may also exist. Therefore, it is important to further
84 explore and validate the potential of DNA metabarcoding for routine use in stream assessment.
85 In this study, we performed a one-to-one comparison of traditional morphological- and DNA metabarcoding–based
86 identification in the context of bioassessment of benthic macroinvertebrate communities at 18 sites across Finland. The
87 samples, which were collected through a national stream bioassessment program, were all morphologically identified by an
88 experienced taxonomist, and were thus ideally suited for comparing the performance of morphological- and DNA-based
89 identification protocols for bioassessments, as well as for critically evaluating the current limitations of both approaches.
90

91 Materials and Methods

92 Sample collection and processing

93 Benthic macroinvertebrates were collected in the fall of 2013 at 18 riffle sites across Finland as part of an official national
94 stream monitoring program (Figure S1, Table S1, Aroviita et al. 2014). Each monitoring sample was collected by taking
95 four 30-s kick-net subsamples covering most microhabitats at each site, following the national guidelines for Water
96 Framework Directive (WFD) monitoring (Meissner *et al.* 2016a). Samples were preserved in 70% ethanol in the field, and
97 all invertebrates in each sample were later sorted in the laboratory. Collected specimens were stored in 70% ethanol, which
98 was not replaced after collection, leading to an average ethanol concentration of 65.14% (SD = 2.83%) during long-term
99 storage. Samples were kept cool (8°C) for subsequent molecular analyses.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

100 All specimens were counted and identified based on morphology, mostly to species or genus level, with the
101 exception of Oligochaeta, Turbellaria, Nematoda, Hydracarina, and the dipteran families Chironomidae and Simuliidae,
102 which were counted but not identified to a lower taxonomic level. The level of identification followed the WFD monitoring
103 protocols targeting operational taxonomic units (OTUs) established by the Finnish Environment Institute SYKE [(Meissner
104 *et al.* 2016a), see page 29]. Identification was performed by a single experienced consultant, who scored higher than
105 average (i.e., >95%) in the most recent international macroinvertebrate taxonomic proficiency tests organised by Proftest of
106 SYKE in 2016 (Meissner *et al.* 2016b).
107

108 DNA extraction and tissue pooling

109 To remove ethanol, specimens from each sampling site were dried overnight in sterile Petri dishes. Specimens were placed
110 in sterile 20-mL tubes containing 10 steel beads (diameter, 5 mm) and homogenised by grinding at 4000 rpm for 30 min in
111 an IKA ULTRA-TURRAX Tube Drive Control System. From each sample, three aliquots each containing on average 14.32
112 mg (SD = 5.56 mg) of homogenised tissues were subjected to DNA extraction. Tissues were digested according to a
113 modified salt DNA extraction protocol (Sunnucks & Hales 1996). Next, 15 µL of DNA were pooled from each of the three
114 extraction replicates, digested with 1 µL of RNase A, and cleaned using a MinElute Reaction Cleanup Kit (Qiagen, Venlo,
115 Netherlands). DNA concentrations were quantified on a Fragment Analyzer™ Automated CE System (Advanced Analytical,
116 Heidelberg, Germany), and the concentrations of all samples were adjusted to 25 ng/µL DNA for PCR.
117

118 PCR amplification, high throughput sequencing and bioinformatics

119 All 18 samples were amplified in duplicate using four BF/BR freshwater macroinvertebrate fusion primer sets targeting
120 fragments internal of the Cytochrome c oxidase subunit I (COI) Folmer region, described previously (Elbrecht & Leese
121 2016b). Table S2 gives an overview of the combinations of fusion primers used for sample tagging with inline barcodes.
122 Each PCR reaction consisted of 1× PCR buffer (including 2.5 mM Mg^{2+}), 0.2 mM dNTPs, 0.5 µM each primer, 0.025 U/µL
123 HotMaster Taq (5Prime, Gaithersburg, MD, USA), 12.5 ng of DNA, and HPLC-grade H_2O to a final volume of 50 µL.
124 PCRs were run on a Biometra TAdvanced Thermocycler with the following program: 94°C for 3 min; 40 cycles of 94°C for
125 30 s, 50°C for 30 s, and 65°C for 2 min; and final extension at 65°C for 5 min. For a few of the samples, it was necessary to
126 use a larger PCR volume (250 µL) due to the presence of PCR inhibitors (see Table S2). PCR products were purified and
127 left-side size-selected using SPRIselect with a ratio of 0.76× (Beckman Coulter, Brea, CA, USA), and then quantified on a
128 Qubit Fluorometer (HS Kit, Thermo Fisher Scientific, Waltham, MA, USA) and Fragment Analyzer™ Automated CE

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

129 System (Advanced Analytical Technologies GmbH, Heidelberg, Germany). PCR products were pooled with equal molarity
130 and sequenced on two Illumina HiSeq 2500 lanes using the Rapid Run 250 bp PE v2 Sequencing Kit with 5% Phi-X spike-
131 in. Sequencing was carried out by GATC Biotech GmbH (Konstanz, Germany).
132 Bioinformatics processing was performed using the UPARSE pipeline in combination with custom R scripts
133 (Dryad DOI) for data processing (Edgar 2013). Scripts are available on http://github.com/VascoElbrecht/JAMP (JAMP
134 v0.10a). Reads were demultiplexed, and paired-end reads were merged using Usearch v8.1.1861 with the following settings:
135 -fastq_mergepairs with -fastq_maxdiffs 99, -fastq_maxdiffpct 99 and -fastq_trunctail 0 (Edgar & Flyvbjerg 2015). Primers
136 were removed using cutadapt version 1.9 with default settings (Martin 2011). Sequences were trimmed to the same 217-bp
137 region amplified by the BF1+BR1 primer set (and the reverse complement generated, if necessary) using fastx_truncate and
138 fastx_revcomp. Only sequences of 207–227 bp were used for further analysis (filtered with cutadapt). Low-quality
139 sequences were then filtered from all samples using fastq_filter with maxee = 0.5. Sequences from all samples were then
140 pooled, dereplicated (minuniquesize = 3), and clustered into molecular operational taxonomic units (MOTUs) using
141 cluster_otus with a 97% identity threshold (Edgar 2013) (includes chimera removal).
142 Pre-processed reads (Figure S2, step B) for all samples were de-replicated again using derep_fulllength, and
143 singletons were included to maximise the information extracted from the sequence data. Sequences from each sample were
144 matched against the MOTUs with a minimum match of 97% using usearch_global. Only OTUs with a read abundance
145 above 0.003% in at least one sequencing replicate were considered in downstream analyses, as this can remove some
146 ambiguous OUTs generated by PCR and sequencing errors (Elbrecht & Leese 2015). Taxonomic assignments for the
147 remaining MOTUs were determined using an R script to search against the BOLD and NCBI databases. Taxonomic
148 information was not further validated, and in the case of conflicting assignments between NCBI and BOLD databases, the
149 taxonomic level for which both databases returned identical results was used. For assignment to species level, a hit with 98%
150 similarity was required in at least one of the two databases; 95% similarity was required for assignment to genus level, 90%
151 for family level, and 85% for order level. Only MOTUs that matched macroinvertebrates were used in the statistical
152 analysis. In all further analyses, only MOTUs with a sequence abundance of at least 0.003% in both replicates of a sample
153 were included.
154

155 **Bioassessment metrics**

156 National bioassessment metrics were calculated from both morphology and DNA metabarcoding data using the protocol for
157 ecological status assessment for the 2nd cycle of WFD river basin management planning (Aroviita et al. 2012).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

158 For this comparison, the DNA-based species lists were reduced to the OTU list used in the Finnish monitoring protocols.
159 The assessment technique for stream macroinvertebrates includes three metrics: number of Type-specific Taxa (TT,
160 Aroviita et al. 2008), number of Type-specific EPT-families (T-EPTh, Aroviita et al. 2012) and PMA-index (Percent Model
161 Affinity, Novak & Bode 1992). Type-specific taxa are taxa typical for expected reference conditions in absence of human
162 disturbance in a given national stream type and region. TT and T-EPTh utilise presence/absence data whereas the PMA-
163 index is a percent similarity between observed and expected assemblages utilising information on taxon relative abundance.
164 The metrics are reported as normalised Ecological Quality Ratios (EQRs) that range from 0 (bad status) to 1 (high status)
165 and is a quotient between observed metric value and value expected in the reference conditions. Also a site-specific mean
166 EQR of the three metric EQRs was calculated.
167
168

169 ## Results

170 ### Sequencing run statistics

171 The HiSeq Rapid run yielded 260.75 million read pairs (raw data available at SRA, accession number SRR4112287). After
172 library demultiplexing, an average of 1.53 million (SD = 0.29 million) read pairs were retained (Figure S3). Unexpected
173 sample tagging combinations (potential tag switching) were uncommon, with only 12 of 136 unused combinations above
174 the 0.003% read abundance threshold and a maximum relative read abundance of 0.006% (Figure S4). After bioinformatic
175 processing, a total of 750 MOTUs remained, of which 49.3% were shared among all four primer sets (Figure S5). The
176 primer combination BF2+BR2 generated the highest number of MOTUs. Sequencing replicates for each sample yielded a
177 mean fold difference in sequence abundance of 2.05 (expected 1.0), indicating high variation in sequence abundance
178 between replicates (Figure S6). We detected a weak but significant negative correlation between relative read abundance per
179 MOTU and variation between replicates in 13 out of 72 total samples ($p \leq 0.05$, Figure S6), but the pattern was not
180 consistent across all samples, as some highly abundant MOTUs also exhibited large differences between replicates.
181

182 ### Taxonomic identification

183 Across all 18 samples, we identified a total of 126 taxa based on morphology, of which 61.1% were identified to species
184 level (Table S3). Eight species lacked public reference sequences in BOLD or NCBI (Table S3), and more taxa were
185 potentially missing at a lower taxonomic resolution (e.g., reference data for specimens only identified to family level). All

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

186 samples were dominated by a few common taxa, whereas rare taxa were only present in a subset of samples (mean Pielou's
187 evenness = 0.65, SD = 0.12, Figure S7).
188 A total of 750 MOTUs remained in the dataset after bioinformatic processing of the sequence data. Of these, we
189 further analysed 573 target invertebrate hits. The MOTU table for DNA metabarcoding with taxonomic assignments, along
190 with MOTU sequences, is available as supplementary Table S4. After taxonomic assignment using BOLD and NCBI, DNA
191 metabarcoding revealed the presence of 288 morphotaxa: 208 species, 47 genera, 23 families, and 10 order or coarser
192 resolution. Metabarcoding resolved more taxa at species level than morphology-based identification protocol (Figure S8).
193 Moreover, DNA metabarcoding consistently detected a substantially greater number of taxa than morphology-based
194 protocol across all samples with each primer combination (57.30% more taxa on average over all data, SD = 35.69%, Figure
195 1). For groups that were morphologically identified to species or genus level, DNA metabarcoding detected 25.3% more
196 taxa using OTUs. Despite enabling the identification of a substantial number of overlooked taxa, DNA metabarcoding did
197 not detect an average of 32.51% (SD = 9.71%) of the taxa identified based on morphology in each sample (Figure 2, see
198 Table S2 for undetected taxa). The proportion of detected taxa was similar for all primer pairs with 79.51% of 288 taxa
199 being detected with all 4 primer combinations and only 9 taxa (3.13%) being detected exclusively with the BF2+BR2
200 primer pair. Also in a principal component analysis (PCA) the primer pairs cluster closely together for all three stream types
201 (Figure S9). The number of reads assigned to each morphotaxon was significantly positively correlated with the number of
202 specimens per taxon for most samples and primer combinations (Figure 3). This correlation was significant for all 18
203 samples for the combination BF2+BR2, but for only 13 or 14 samples for the other primer combinations. However, despite
204 the positive correlations between read abundance and number of taxa, read abundance still varied by two orders of
205 magnitude, and this was also reflected in the low adjusted R^2 values for all primer sets (mean = 0.366 to 0.411).
206 Assessment metrics calculated from morphology and DNA metabarcoding data were generally similar (Figure 4,
207 Table S1), especially for TT and T-EPTh metrics which utilise presence/absence data only. For a few samples, however, the
208 status quality class changed with the DNA-based taxa lists. Most differing assessments were obtained with the PMA metric
209 utilising relative abundances, which assigned most samples to poorer status with DNA-identification than with the
210 morphological identification (Figure 4C). The overall status class (mean EQR) was generally similar between the two
211 approaches, and only 5 cases were one class lower with the DNA-identification (Figure 4D).
212

213 Discussion

214 Performance of DNA metabarcoding

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

215 Our results and other studies demonstrate that DNA-based identification methods can capture more diversity than routine
216 morphological identification protocols, even though several of the morphologically identified taxa were not recovered using
217 metabarcoding (Carew *et al.* 2013; Zimmermann *et al.* 2014; Lejzerowicz *et al.* 2015; Gibson *et al.* 2015; Clarke *et al.*
218 2017). Not all *Baetis* morphospecies which were very abundant across most samples (e.g. *B. niger*) were detected with
219 metabarcoding, potentially due to primer bias, morphological misidentification, recent speciation or likely because of lack
220 of barcode sequences or conflicting taxonomic information in reference databases, as the *Baetis* species complex is difficult
221 to identify based on morphology (Williams *et al.* 2006; Savolainen *et al.* 2007; Lucentini *et al.* 2011). DNA metabarcoding
222 was especially powerful for resolving taxon diversity in groups that are difficult or unfeasible in current morphology-based
223 biomonitoring protocols to distinguish morphologically in their larval stages, including dipteran families (chironomids and
224 simuliids), mites, Oligochaeta, and Limnephilidae. In addition, EPT taxa that were morphologically identified only to the
225 family (Limnephilidae) or genus level (e.g., *Eloeophila* and *Hydroptila*) could be identified to species level using DNA
226 metabarcoding, indicating a significant advantage of the metabarcoding to biomonitoring. Consistent with our observations,
227 higher taxonomic resolution of DNA-based methods in comparison with morphology-based identification has been
228 demonstrated in many previous studies (Baird & Sweeney 2011; Sweeney *et al.* 2011; Stein *et al.* 2013a; Gibson *et al.*
229 2015). All four of our macroinvertebrate-specific BF/BR primer combinations yielded similarly good performance,
230 consistent with our previous mock community tests (Elbrecht & Leese 2016b).
231 Morphology-based and DNA-based taxon lists yielded very similar results for the metrics used in WFD ecological
232 status assessment, indicating that metabarcoding can produce usable taxonomic data for current assessment techniques. This
233 finding is consistent with marine studies, which have demonstrated a good match between morphological assessments and
234 presence/absence data, as well as DNA-based taxon lists (Aylagas *et al.* 2014; Lejzerowicz *et al.* 2015; Aylagas *et al.*
235 2016a). Considering that the metrics used in this study were optimized for the routine morphological identification protocol,
236 in some cases considering coarser taxonomic levels than genus or species, future DNA-based assessment might indeed be
237 further improved by applying optimized metric calculation approaches and species-level trait databases (Mondy *et al.* 2012;
238 Schmidt-Kloiber & Hering 2015). DNA metabarcoding can provide much more accurate taxonomic identification than
239 morphology-based methods, and can even be used to detect cryptic species (Elbrecht & Leese 2015). The increase in
240 accuracy provides an opportunity to investigate potential differences in ecological preferences and detect stressors based on
241 indicator taxa when larval morphology alone is not sufficient (Macher *et al.* 2016). In future assessment techniques this
242 valuable additional information could be integrated by refining and expanding the taxa lists for expected reference
243 conditions. This might not only refine our conception of the condition of streams, but could also help to disentangle effects
244 of multiple stressors on ecosystems.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

245 While DNA metabarcoding has the advantages of increased reproducibility and taxonomic resolution, it also has
246 drawbacks, including the inability to quantify taxon abundance (Piñol *et al.* 2014; Elbrecht & Leese 2015). Although
247 approaches to adjust for primer bias have been developed (Thomas *et al.* 2015), these methods are unlikely to succeed in
248 complex communities; moreover, sequence abundance is affected by taxon biomass (Elbrecht *et al.* 2016). Nevertheless, in
249 most samples, we detected a significant linear relationship between the number of morphologically identified specimens and
250 the number of sequencing reads assigned to the respective OTUs. Although this could be interpreted as a potential means to
251 estimate taxon abundance, the poor fit and high scatter of up to two orders of magnitude (similar to comparisons in other
252 studies (Carew *et al.* 2013; Dowle *et al.* 2015; Leray & Knowlton 2015; Clarke *et al.* 2017)) prevent its practical
253 exploitation. While BF/BR primers exhibited less primer bias than the previously tested Folmer primers (Folmer *et al.* 1994;
254 Elbrecht & Leese 2015), the bias remained substantial (Elbrecht & Leese 2016b). In addition, sequence abundance is likely
255 further influenced by the different biomass of taxa and specimens of different sizes in a sample. Therefore, with exact
256 biomass data for each specimen, the relationship to sequence abundance might be stronger. Nonetheless, we argue that
257 particularly estimating biomass from PCR-based metabarcoding analyses could be useful when used in e.g. a semi-
258 quantitative way, even though primer bias hinders obtaining exact estimates.

259 Laboratory and sequencing costs are critical determinants of the viability of large-scale DNA-based monitoring (Ji
260 *et al.* 2013). In this study, the sequencing costs per sample using one primer pair and two replicates (~1.5 million sequences)
261 are 110 € (7900€ for the complete run at a commercial sequencing provider). Sequencing costs are likely to decline in the
262 future and could be further decreased by pooling more samples in each sequencing run and by pre-sorting samples
263 according to biomass (e.g. using sieves (Leray & Knowlton 2015; Aylagas *et al.* 2016b; Elbrecht *et al.* 2016)). All
264 laboratory steps from DNA extraction to library preparation currently accumulate to 70 € per sample, leading to a total cost
265 of 180 € per sample in this study, similar to previous estimates (Ji *et al.* 2013; Stein *et al.* 2014). Expenses related to
266 laboratory infrastructure and bioinformatics (which can be reduced by automation and parallelization) as well as kick
267 sample collection and sorting (Haase *et al.* 2004) may push the total costs per sample to 500–750 €, which is comparable to
268 current morphology-based monitoring costs (Buss *et al.* 2015). Kick sample collection and sorting makes a major
269 contribution to total expenses (up to 2/3rd in Finland), which might be substantially reduced by homogenising complete
270 kick samples without sorting or drying overnight. However, further optimisation of PCR inhibitor removal (e.g. using
271 commercial DNA purification kits) is necessary, as organic and anorganic substrates likely cause impact on amplification
272 efficiency. Environmental DNA is unlikely to be an alternative to sampling whole organisms for the detection of whole
273 macroinvertebrate communities, as DNA quantity and thus detection rates in eDNA metabarcoding are low (Aylagas *et al.*
274 2016a) and affected by additional biases (Barnes & Turner 2015).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

275

276 **Factors currently limiting DNA metabarcoding for ecosystem assessment**

277 The DNA metabarcoding protocol we used worked reliably across all 18 samples. However, we identified various
278 opportunities to further improve the performance of metabarcoding. Figure 5 provides an overview of the limitations of
279 DNA metabarcoding in relation to taxonomic assignment and the reference database, as well as the laboratory protocol
280 routines. Across all 18 samples used in this study, our metabarcoding approach was unable to detect 32% of the
281 morphologically identified taxa . Some of these omissions were linked to application of the precautionary principle,
282 specifically, the tendency of human experts to relegate the identification of small specimens to coarser taxa (e.g., genus
283 level) if higher taxonomic resolution cannot be established without doubt. In addition, the laboratory procedures used in
284 routine monitoring campaigns are not fully adequate for DNA extraction. For example, the low alcohol concentration
285 typically used for sample preservation during routine biomonitoring (typically 70% ethanol) may result in specimens still
286 viable for morphological detection, but containing highly degraded DNA, impairing their accurate molecular detection.
287 Collection and preservation of samples in 96% ethanol will likely prevent DNA degradation (Stein *et al.* 2013b). Further,
288 although unlikely given the proven proficiency of the expert who performed our morphological identification, it remains
289 possible that erroneous morphotaxonomical identification by the human expert may have introduced false taxa, contributing
290 to the discrepancy between the results of the identification methods. Several additional factors, listed in Figure 5, may have
291 influenced detection, either positively or negatively.

292 Laboratory methodology can strongly influence the absolute and relative amounts of invertebrates detected by
293 DNA metabarcoding. Because primer/template mismatches can prevent certain taxa from being amplified by PCR, primer
294 bias is one of the most serious concerns (Deagle *et al.* 2014; Piñol *et al.* 2014; Elbrecht & Leese 2015). The negative effects
295 of primer bias can be reduced by incorporating primer degeneracy and carefully choosing primer sets suited for the targeted
296 ecosystem and taxonomic groups (Elbrecht & Leese 2016a,b). However, even after primer optimisation, one-step PCR
297 methods will be affected by primer bias. Therefore, it is unlikely that all taxa present in a sample can be detected by DNA
298 metabarcoding, and primer bias makes it difficult to estimate abundance or biomass. PCR and sequencing errors, undetected
299 chimeras, and misidentified reference sequences can also lead to false positive detection. Moreover, specimens in a sample
300 can vary widely in biomass, depending on species and life stage. This not only prevents the estimation of taxon abundances,
301 but can also prevent detection of small and rare taxa (Elbrecht *et al.* 2016). Because 68.3% of the taxa detected in this study
302 were present in five or fewer samples, our data were likely affected by this bias. Primer bias and variation in taxon biomass
303 taken together, make it difficult to relate read abundance to taxon abundance. Although presence/absence data might already

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

304 be sufficient for ecosystem assessment (Aylagas *et al.* 2014), one must acknowledge that relative abundance -based
305 estimates are likely possible if identical protocols are used across all sample sites, thus leading to similar biases across
306 samples.

307 Some of our samples were also affected by PCR inhibition, a problem that could be solved by using larger PCR
308 volumes to dilute PCR inhibitors or with additional clean-up steps. However, because monitoring protocols must work in all
309 stream ecosystems, independent of environmental conditions, PCR inhibition remains a major challenge for the application
310 of DNA metabarcoding in this context. Ideally, methods should be developed and tested for purifying DNA from complete
311 kick-net samples without pre-sorting specimens from debris (e.g., sediment, small stones, leaves, and organic particles).
312 This would allow researchers to skip the time-consuming pre-sorting steps, during which up to 30% of specimens can be
313 missed (Haase *et al.* 2010). Thus, circumvention of pre-processing would allow inclusion of often overlooked small taxa,
314 potentially detecting more taxa.

315 Several other laboratory-specific factors might also affect metabarcoding. For example, tag switching is an issue
316 potentially generating additional MOTUs across several samples multiplexed in one library (Esling *et al.* 2015; Schnell *et al.*
317 2015). However, we did not observe such effects on our samples, or in our previous studies using the fusion primer system
318 with inline tagging. However, O´Donnell *et al.* (2016) showed that tags can lead to biases in read abundance, and our
319 samples are potentially affected by this phenomenon, as evidenced by the observed ~2-fold variation in read abundances
320 between the replicates for a given sample. It is of critical importance to minimise tag switching and determine the level of
321 tagging induced bias between replicates and its effect on the data. In our case, variation in read abundance might have
322 resulted in underestimation of diversity, because we conservatively discarded all reads not present in both replicates.
323 Although we obtained good taxonomic resolution, it is important to be aware of, and account for, these shortcomings, as
324 well as to solve these problems by modifying current protocols.

325 Clearly, DNA metabarcoding is not perfect, and of the many different protocols being developed, few have been
326 extensively validated. Method 'ground truthing' is essential to build trust in metabarcoding methods for monitoring, and the
327 various candidate protocols and modifications must be validated using the same standard invertebrate mock communities.
328 Sample sets specifically designed for such validation efforts would not only reveal biases, but could also be used to accredit
329 monitoring offices in order to ensure that their laboratory work meets quality standards, and that their results are comparable
330 with those of other accredited offices. Once a well-established standardised metabarcoding protocol is developed, the
331 analysis of high-throughput metabarcoding data could be carried out on cloud-based systems, facilitating comparisons and
332 easy updating of all bioinformatic analyses. Further, common metadata standards and central storage of all monitoring

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

PeerJ Preprints NOT PEER-REVIEWED

333 related metabarcoding data could serve as a valuable resource for research, e.g., by providing accurate maps of taxa
334 presence over a large geographic and temporal scales with unprecedented accuracy.
335 The second major factor influencing our results is database accuracy, along with the reliability of morphology-
336 based identification of specimens deposited in databases. Here, we specifically constrained our comparison to MOTUs with
337 assigned taxonomic information from the BOLD and NCBI reference databases and did not consider other MOTUs, despite
338 the potential to further increase assessment accuracy. Within the framework of the WFD, ecological assessment of aquatic
339 ecosystems in many countries currently evaluates taxa, associated traits and indicator values; therefore, metabarcoding must
340 compete on the same level. While we think it is feasible to infer traits by correlating MOTUs with abiotic data from
341 sampling locations, we currently lack metabarcoding datasets of sufficient size to verify this. Furthermore, it is desirable to
342 maintain and associate taxonomic information with MOTUs, to relate ecological information to obtained sequences, and to
343 associate correlative found traits and ecological preferences back to the taxa detected by metabarcoding. Currently,
344 available databases are still incomplete, and not all taxa have barcodes. Additionally, the accuracy of identification of larvae
345 and adult invertebrates varies depending on expert experience, and even databases like BOLD, specifically built for DNA
346 barcoding, contain misidentified taxa or conflicting taxonomic assignments for the same BIN (Barcode Index Number,
347 (Ratnasingham & Hebert 2013)). Databases require stricter standards and quality control, including incentives for data
348 providers and managers to better curate their data after the initial release. Sample degradation and misidentification could
349 have affected both our 18 samples and also the reference databases. Those errors could have further propagated into false
350 positives or negatives in both the morphologically-generated taxon list and our metabarcoding-based assessments. It is
351 imperative that taxonomical experts and molecular biologists come together to discuss and solve conflicting cases,
352 especially as traditional taxonomic expertise fades. DNA metabarcoding provides an excellent opportunity for traditional
353 taxonomists to contribute to reference databases and the increased taxonomic resolution make it possible to associate
354 ecological information with difficult groups such as Diptera.
355

356 **Conclusions**

357 We demonstrated that DNA metabarcoding is a viable alternative to morphology-based identification of macroinvertebrates,
358 as both the assessment results and costs are very similar for both methods. DNA metabarcoding detected more taxa than
359 morphology-based analysis in all samples examined. If combined with ecological species traits, DNA metabarcoding could
360 potentially improve assessment results over those obtained through morphological identification alone. Despite its merits,
361 the DNA-based approach has still minor technical issues, which, along with unreliability in reference databases, must be

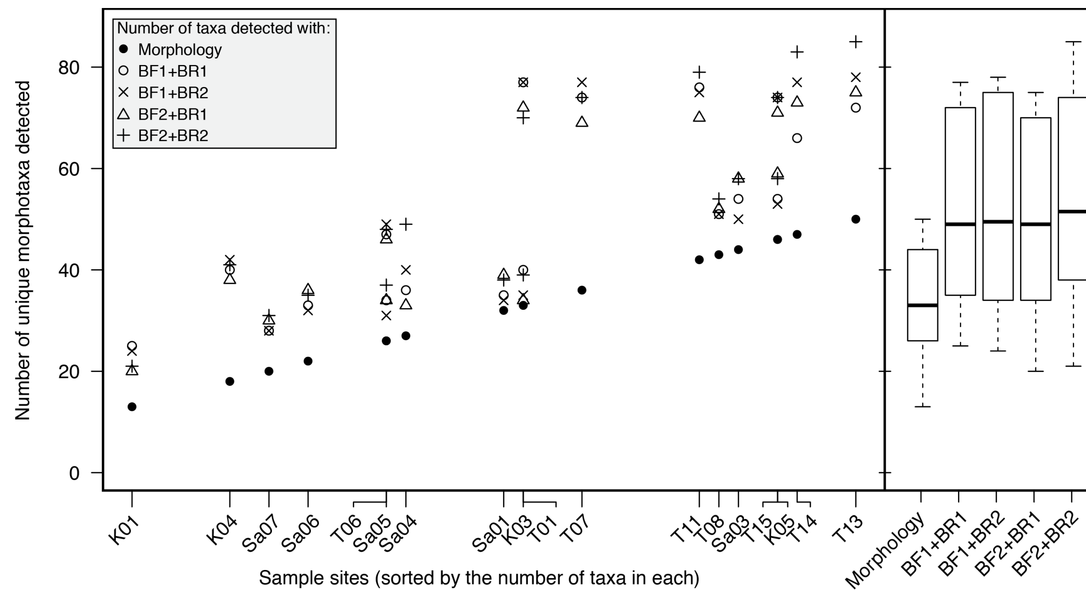
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

362 resolved before the full potential of DNA metabarcoding can be unlocked. This will require coordinated efforts such as the
363 DNAqua-Net project, which combines contributions from molecular biologists, ecologists and taxonomists (Leese *et al.*
364 2016).
365
366
367 **Acknowledgements:** FL and VE were supported by a grant of the Kurt Eberhard Bode Foundation to FL. VE, FL and KM
368 are supported by the EU COST Action DNAqua-Net (CA15219). KM was supported by the Academy of Finland project
369 DETECT (289104). The samples were collected with funding from the Finnish Ministry of Agriculture and Forestry. Two
370 anonymous reviewers provided helpful comments that greatly improved the quality of this manuscript.
371
372 **Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and
373 writing of the paper.
374
375 **Author contributions statement:** VE, KM, and FL conceived the ideas and designed the methodology. EV carried out the
376 laboratory work; VE performed bioinformatic analyses together with EV; JA calculated assessment metrics; and VE led the
377 writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.
378

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

PeerJ Preprints NOT PEER-REVIEWED

379 # Figures



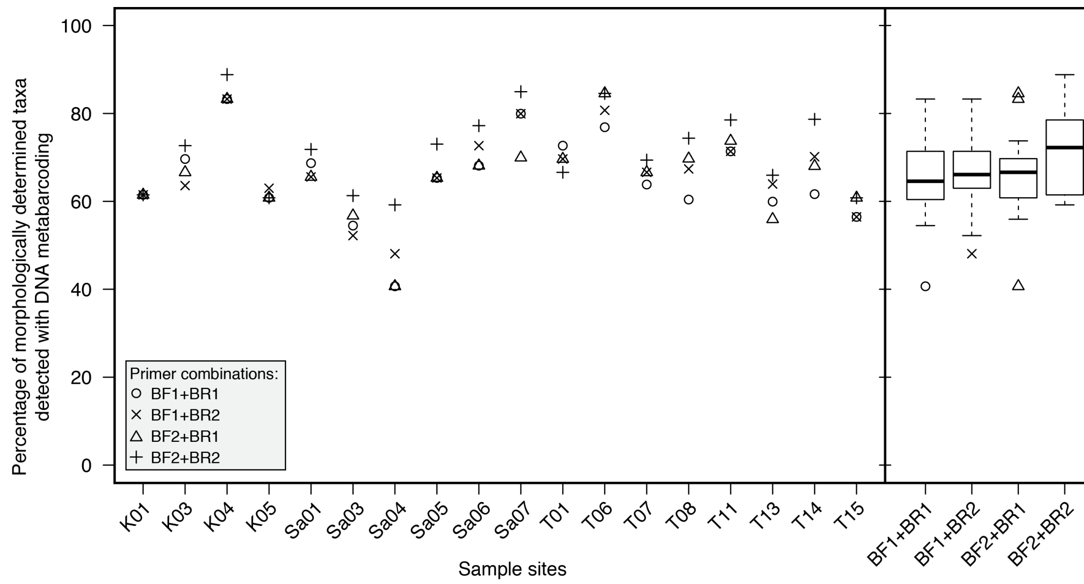
380

381 **Figure 1**: Number of morphotaxa detected by morphological and DNA-based identification methods across all 18 sample
382 sites. The number of taxa detected by DNA-based identification was compared among four primer pairs (different symbols).
383 The boxplot on the right compares the overall performances of DNA- and morphology-based identification across samples.

384

385

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017



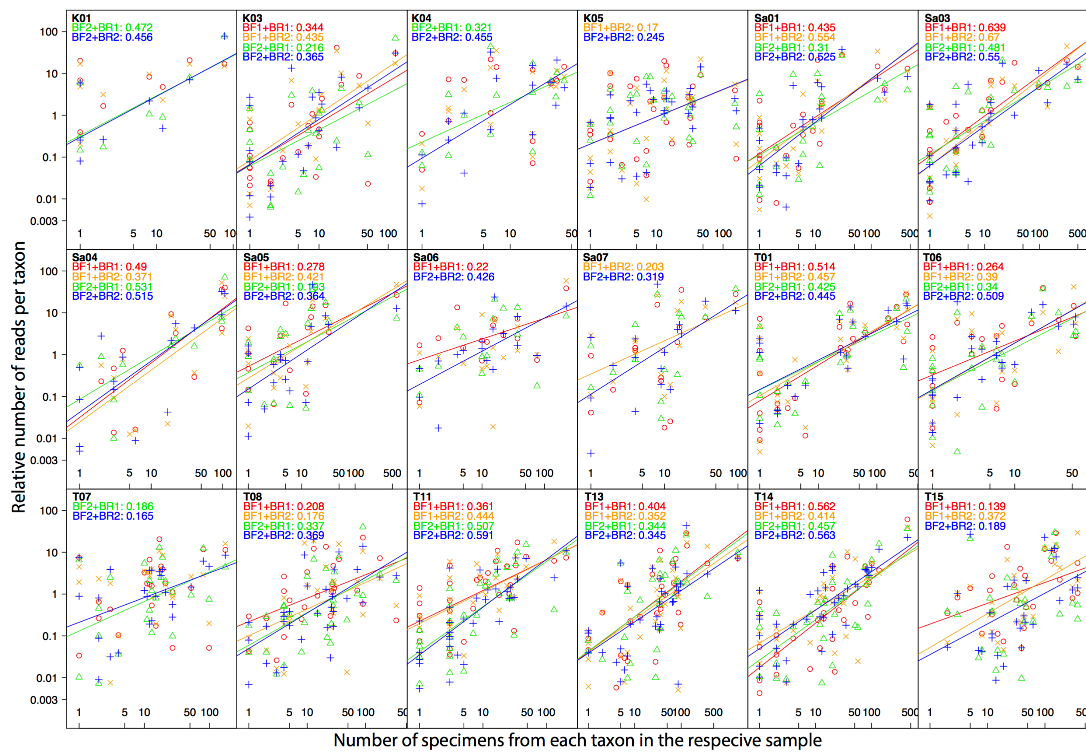
386

387 **Figure 2**: Percentage of morphologically-identified taxa detected with four different primer pairs across all 18 sample sites.

388 Primers pairs are indicated by different symbols, and overall detection rates for the primer pairs are shown on the right.

389

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017



390

391 **Figure 3:** Relative logarithmic sequence abundance plotted against logarithmic number of specimens from each

392 morphologically identified taxon for all 18 individual samples. The four primer combinations are indicated by colour, with a

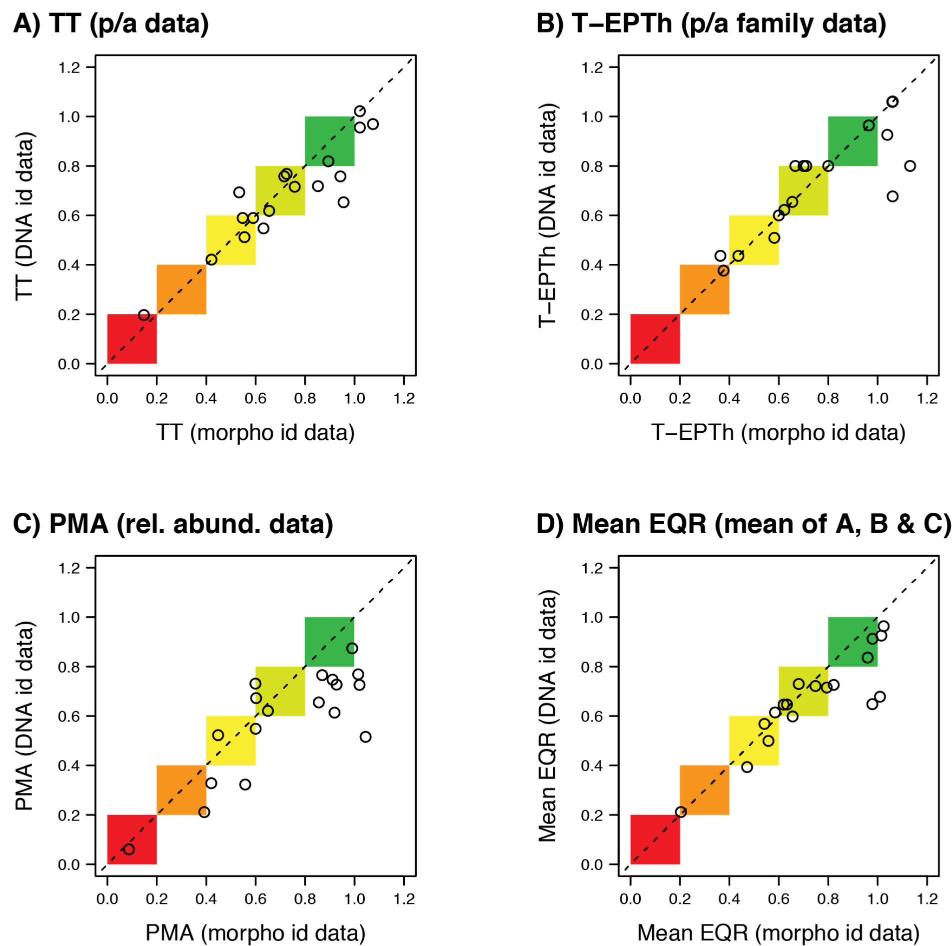
393 linear regression line plotted in case of a significant positive linear correlation (p=<0.05) and the adjusted R^2 value is given

394 for the respective primer pair.

395

396

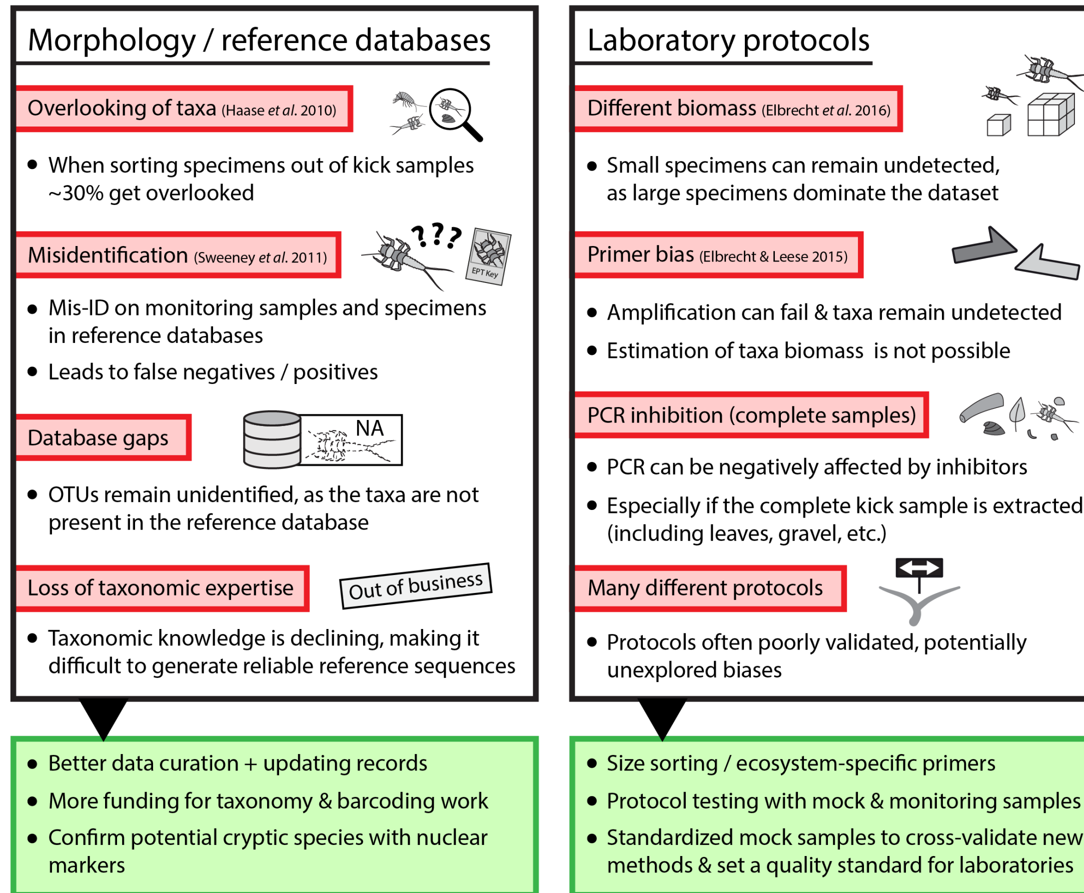
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017



Assessment: Bad Poor Moderate Good High

Figure 4: Comparison of Finnish macroinvertebrate WFD assessment metrics calculated with taxa lists based on morphological- and DNA-based (primers BF2+BR2) identification. The metrics are shown as normalised Ecological Quality Ratios (EQR) ranging from 0 (low status) to 1 (values can >1 if more taxa are observed than expected on average). For all four metrics, the results of morphological- and DNA-based assessments were significantly correlated (Pearson correlation, p > 0.0001). A) Occurrence of stream Type-specific Taxa (TT, based on presence/absence data). B) Occurrence of stream Type-specific EPT families (EPTh, based on presence/absence family data). C) Percent model affinity (PMA, based on relative abundance data). D) Mean EQR of the three metrics.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017



405

406 **Figure 5:** Overview of factors currently limiting the application of DNA metabarcoding to ecosystem assessment, with

407 potential solutions.

408

409

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

410

411 ## References

412 Aroviita J., Koskenniemi E., Kotanen J., Hämäläinen H., 2008. A priori typology-based prediction of benthic
413 macroinvertebrate fauna for ecological classification of rivers. Environmental Management 42, 894-906.

414 Aroviita J., Hellsten S., Jyväsjärvi J., Järvenpää L., Järvinen M., Karjalainen S.M. et al., 2012. Guidelines for the ecological
415 and chemical status classification of surface waters for 2012–2013 – updated assessment criteria and their application.
416 Environmental Administration Guidelines 7/2012, 1–144. (in Finnish) http://hdl.handle.net/10138/41788

417 Aroviita, J., Vuori, K.-M., Hellsten, S., Jyväsjärvi, J., Järvinen, M., Karjalainen, S.M., Kauppila, P., Korpinen, S., Kuoppala,
418 M., Mitikka, S., Mykrä, H., Olin, M., Rask, M., Riihimäki, J., Räike, A., Rääpysjärvi, J., Sutela, T., Vehanen, T.,
419 Vuorio, K. 2014. Ecological status and monitoring of surface waters polluted by agriculture and forestry. Reports of
420 the Finnish Environment Institute 12/2014: 1–96. (in Finnish) http://hdl.handle.net/10138/45017

421 Aylagas, E., Borja, A. & Rodríguez-Ezpeleta, N. (2014). Environmental status assessment using DNA metabarcoding:
422 towards a genetics based Marine Biotic Index (gAMBI). (S. Consuegra, Ed.). *PloS one*, **9**, e90529.

423 Aylagas, E., Borja, A., Irigoien, X. & Rodríguez-Ezpeleta, N. (2016a). Benchmarking DNA Metabarcoding for
424 Biodiversity-Based Monitoring and Assessment. *Frontiers in Marine Science*, **3**, 1809–12.

425 Aylagas, E., Mendibil, I., Borja, A. & Rodríguez-Ezpeleta, N. (2016b). Marine Sediment Sample Pre-processing for
426 Macroinvertebrates Metabarcoding: Mechanical Enrichment and Homogenization. *Frontiers in Marine Science*, **3**,
427 1452–8.

428 Baird, D.J. & Hajibabaei, M. (2012). Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-
429 generation DNA sequencing. **21**, 2039–2044.

430 Baird, D.J. & Sweeney, B.W. (2011). Applying DNA barcoding in benthology: the state of the science. *Journal of the North*
431 *American Benthological Society*, **30**, 122–124.

432 Barnes, M.A. & Turner, C.R. (2015). The ecology of environmental DNA and implications for conservation genetics.
433 *Conservation Genetics*, 1–17.

434 Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., Solimini, A., van de Bund, W., Zampoukas, N. &
435 Hering, D. (2012). Three hundred ways to assess Europe's surface waters: An almost complete overview of biological
436 methods to implement the Water Framework Directive. *Ecological Indicators*, **18**, 31–41.

437 Buss, D.F., Carlisle, D.M., Chon, T.-S., Culp, J., Harding, J.S., Keizer-Vlek, H.E., Robinson, W.A., Strachan, S., Thirion, C.
438 & Hughes, R.M. (2015). Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale
439 programs. *Environmental monitoring and assessment*, **187**, 4132.

440 Carew, M.E., Pettigrove, V.J., Metzeling, L. & Hoffmann, A.A. (2013). Environmental monitoring using next generation
441 sequencing: rapid identification of macroinvertebrate bioindicator species. *Frontiers in zoology*, **10**, 1–1.

442 Clarke, L.J., Beard, J.M., Swadling, K.M. & Deagle, B.E. (2017). Effect of marker choice and thermal cycling protocol on
443 zooplankton DNA metabarcoding studies. *Ecology and Evolution*, **7**, 873–883.

444 Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F. & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c
445 oxidase subunit I marker: not a perfect match. *Biology Letters*, **10**, 20140562–20140562.

446 Dowle, E.J., Pochon, X. & Banks, J.C. (2015). Targeted gene enrichment and high-throughput sequencing for
447 environmental biomonitoring: a case study using freshwater macroinvertebrates. *Molecular Ecology*.

448 Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, **10**, 996–
449 998.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

450 Edgar, R.C. & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads.
451 *Bioinformatics*, **31**, 3476–3482.

452 Elbrecht, V. & Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer
453 Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol (M. Hajibabaei, Ed.). *PloS*
454 *one*, **10**, e0130324–16.

455 Elbrecht, V. & Leese, F. (2016a). PrimerMiner: an R package for development and in silico validation of DNA
456 metabarcoding primers. *Methods in Ecology and Evolution*.

457 Elbrecht, V. & Leese, F. (2016b). Validation and development of freshwater invertebrate metabarcoding COI primers for
458 Environmental Impact Assessment. *PeerJ PrePrints*, 1–16.

459 Elbrecht, V., Peinert, B. & Leese, F. (2016). Sorting things out - assessing effects of unequal specimen biomass on DNA
460 metabarcoding. *PeerJ PrePrints*.

461 Esling, P., Lejzerowicz, F. & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-
462 sequencing. *Nucleic acids research*, **43**, 2513–2524.

463 Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial
464 cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology*, **3**,
465 294–299.

466 Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I. & Hajibabaei, M. (2015). Large-Scale
467 Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing (D. Fontaneto, Ed.). *PloS one*,
468 **10**, e0138432–15.

469 Gibson, J., Shokralla, S., Porter, T.M., King, I., van Konynenburg, S., Janzen, D.H., Hallwachs, W. & Hajibabaei, M.
470 (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through
471 DNA metasystematics. *Proceedings of the National Academy of Sciences*, **111**, 8007–8012.

472 Haase, P., Lohse, S., Pauls, S. & Schindehütte, K. (2004). Assessing streams in Germany with benthic invertebrates:
473 development of a practical standardised protocol for macroinvertebrate sampling and sorting. *Limnologica*, **34**, 349–
474 365.

475 Haase, P., Pauls, S.U., Schindehütte, K. & Sundermann, A. (2010). First audit of macroinvertebrate samples from an EU
476 Water Framework Directive monitoring program: human error greatly lowers precision of assessment results. *Journal*
477 *of the North American Benthological Society*, **29**, 1279–1291.

478 Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. & Baird, D.J. (2011). Environmental Barcoding: A Next-Generation
479 Sequencing Approach for Biomonitoring Applications Using River Benthos. *PloS one*.

480 Hajibabaei, M., Spall, J.L., Shokralla, S. & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic
481 macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol.
482 *BMC ecology*, **12**, 1–1.

483 Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P.,
484 Edwards, F.A., Larsen, T.H., Hsu, W.W., Benedick, S., Hamer, K.C., Wilcove, D.S., Bruce, C., Wang, X., Levi, T.,
485 Lott, M., Emerson, B.C. & Yu, D.W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via
486 metabarcoding. (M. Holyoak, Ed.). *Ecology letters*, **16**, 1245–1257.

487 Leese, F., Altermatt, F., Bouchez, A. & Ekrem, T. (2016). DNAqua-Net: Developing new genetic tools for bioassessment
488 and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*.

489 Lejzerowicz, F., Esling, P., Pillet, L., Wilding, T.A., Black, K.D. & Pawlowski, J. (2015). High-throughput sequencing and
490 morphology perform equally well for benthic monitoring of marine ecosystems. *Scientific Reports*, 1–10.

491 Leray, M. & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of marine

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

492 benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 201424997–6.

493 Lucentini, L., Rebora, M., Puletti, M.E., Gigliarelli, L., Fontaneto, D., Gaino, E. & Panara, F. (2011). Geographical and
494 seasonal evidence of cryptic diversity in the Baetis rhodani complex (Ephemeroptera, Baetidae) revealed by means of
495 DNA taxonomy. *Hydrobiologia*, **673**, 215–228.

496 Macher, J.N., Salis, R.K., Blakemore, K.S. & Tollrian, R. (2016). Multiple-stressor effects on stream invertebrates: DNA
497 barcoding reveals contrasting responses of cryptic mayfly species. *Ecological Indicators*, 159–169.

498 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, **17**, 10–
499 12.

500 Meissner, K., Aroviita, J., Hellsten, S., Järvinen, M., Karjalainen, S.M., Kuoppala, M. & Vuori, H.M.J.K.-M. (2016a).
501 *Jokien ja järvien biologinen seuranta–näytteenotosta tiedon tallentamiseen.(Biological monitoring of rivers and lakes–*
502 *from sampling to data recording) in Finish*. Finnish Environment ….

503 Meissner, K., Nygård, H., Björklöf, K., Jaale, M., Hasari, M., Laitila, L., Rissanen, J. & Leivuori, M. (2016b). Proficiency
504 Test 04/ 2016 Taxonomic identification of boreal freshwater lotic, lentic, profundal and North-Eastern Baltic benthic
505 macroinvertebrates. *Reports of the finnish environment institute*, **4**.

506 Mondy, C.P., Villeneuve, B., Archaimbault, V. & Usseglio-Polateraa, P. (2012). A new macroinvertebrate-based
507 multimetric index (I 2 M 2) to evaluate ecological quality of French wadeable streams fulfilling the WFD demands: a
508 taxonomical and trait apporach. *Ecological Indicators*, **18**, 452–467.

509 O'Donnell, J.L., Kelly, R.P., Lowell, N.C. & Port, J.A. (2016). Indexed PCR Primers Induce Template-Specific Bias in
510 Large-Scale DNA Sequencing Studies (A.R. Mahon, Ed.). *PloS one*, **11**, e0148698–11.

511 Piñol, J., Mir, G., Gomez-Polo, P. & Agustí, N. (2014). Universal and blocking primer mismatches limit the use of high-
512 throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular ecology resources*, **15**, 1–12.

513 Ratnasingham, S. & Hebert, P. (2007). BOLD: The Barcode of Life Data System (http://www. barcodinglife. org).
514 *Molecular Ecology Notes*, **7**, 355–364.

515 Ratnasingham, S. & Hebert, P.D.N. (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number
516 (BIN) System (D. Fontaneto, Ed.). *PloS one*, **8**, e66213–16.

517 Resh, V.H. & Unzicker, J.D. (1975). Water quality monitoring and aquatic organisms: the importance of species
518 identification. *Journal - Water Pollution Control Federation*, **47**, 9–19.

519 Savolainen, E., Drotz, M.K. & Hoffsten, P. (2007). The Baetis vernus group (Ephemeroptera: Baetidae) of northernmost
520 Europe: an evidently diverse but poorly understood group of mayflies. *Entomologica …*.

521 Schmidt-Kloiber, A. & Hering, D. (2015). www.freshwaterecology.info – An online tool that unifies, standardises and
522 codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecological Indicators*, **53**,
523 271–282.

524 Schnell, I.B., Bohmann, K. & Gilbert, M.T.P. (2015). Tag jumps illuminated - reducing sequence-to-sample
525 misidentifications in metabarcoding studies. *Molecular ecology resources*, **15**, 1289–1303.

526 Stein, E.D., Martinez, M.C., Stiles, S., Miller, P.E. & Zakharov, E.V. (2014). Is DNA Barcoding Actually Cheaper and
527 Faster than Traditional Morphological Methods: Results from a Survey of Freshwater Bioassessment Efforts in the
528 United States? (M. Casiraghi, Ed.). *PloS one*, **9**, e95525.

529 Stein, E.D., White, B.P., Mazor, R.D., Jackson, J.K. & Battle, J.M. (2013a). Does DNA barcoding improve performance of
530 traditional stream bioassessment metrics? *Freshwater Science*, **33**, 302–311.

531 Stein, E.D., White, B.P., Mazor, R.D., Miller, P.E. & Pilgrim, E.M. (2013b). Evaluating Ethanol-based Sample Preservation
532 to Facilitate Use of DNA Barcoding in Routine Freshwater Biomonitoring Programs Using Benthic Macroinvertebrates

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

533 (D. STEINKE, Ed.). *PloS one*, **8**, e51273–7.

534 Sunnucks, P. & Hales, D.F. (1996). Numerous transposed sequences of mitochondrial cytochrome oxidase I-II in aphids of
535 the genus Sitobion (Hemiptera: Aphididae). *Molecular biology and evolution*, **13**, 510–524.

536 Sweeney, B.W., Battle, J.M., Jackson, J.K. & Dapkey, T. (2011). Can DNA barcodes of stream macroinvertebrates improve
537 descriptions of community structure and water quality? *Journal of the North American Benthological Society*, **30**, 195–
538 216.

539 Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. (2012). Towards next-generation biodiversity
540 assessment using DNA metabarcoding. *Molecular Ecology*, **21**, 2045–2050.

541 Thomas, A.C., Deagle, B.E., Paige Eveson, J., Harsch, C.H. & Trites, A.W. (2015). Quantitative DNA metabarcoding:
542 improved estimates of species proportional biomass using correction factors derived from control material. *Molecular*
543 *ecology resources*, n/a–n/a.

544 Williams, H.C., Ormerod, S.J. & Bruford, M.W. (2006). Molecular systematics and phylogeography of the cryptic species
545 complex Baetis rhodani (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution*, **40**, 370–382.

546 Yu, D.W., Ji, Y., Emerson, B.C., Wang, X., Ye, C., Yang, C. & Ding, Z. (2012). Biodiversity soup: metabarcoding of
547 arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, **3**, 613–623.

548 Zimmermann, J., Glöckner, G., Jahn, R., Enke, N. & Gemeinholzer, B. (2014). Metabarcoding vs. morphological
549 identification to assess diatom diversity in environmental studies. *Molecular ecology resources*, 1–17.

550

551

552 Supporting information

553 **Figure S1.** Map of sample locations.

554

555 **Figure S2.** Flow chart detailing bioinformatics steps in our metabarcoding pipeline.

556

557 **Figure S3.** Scatterplot showing the number of reads obtained for the samples.

558

559 **Figure S4.** Matrix indicating potential tag switching.

560

561 **Figure S5.** Bar plot showing the number of OTUs shared among different primer sets.

562

563 **Figure S6.** Plot showing reproducibility between replicates.

564

565 **Figure S7.** Presence of morphotaxa across samples.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017

566

567 **Figure S8.** Comparison of taxonomic resolution between morphology and DNA metabarcoding.

568

569 **Figure S9.** Principal component analysis (PCA) comparing performance of the 4 used primer sets.

570

571 **Table S1.** Sample site coordinates and calculated assessment indices.

572

573 **Table S2.** Tagging combinations used in the metabarcoding library.

574

575 **Table S3.** Overview of morphotaxa identified based on morphology across samples.

576

577 **Table S4.** OTU table.

578

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v3 | CC BY 4.0 Open Access | rec: 22 Mar 2017, publ: 22 Mar 2017