PeerJ Preprints NOT PEER-REVIEWED

1 **Title: Strengths and weaknesses of DNA-based monitoring: Assessing**
2 **macroinvertebrates in 18 Finnish streams with metabarcoding and morphology**

3

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

5 **Word count:**

6 **Authors:** Vasco Elbrecht[1], 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, 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,
14 45141 Essen, Germany

15

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

17

18 **Abstract**

19 1) DNA metabarcoding holds great promise for assessment of stream ecosystems with macroinvertebrates. However,
20 few large-scale studies have compared the performance of DNA metabarcoding with that of routine morphological
21 identification.

22 2) We tested metabarcoding using 18 macroinvertebrate samples from Finland using four primer sets. The samples were
23 collected in 2013 and identified based on morphology as part of a Finnish stream monitoring program. Morphological
24 identification was performed to the taxonomic level at which identification was reliable following standardized
25 protocols.

26 3) We identified over twice the number of taxa, with greater species-level resolution, using DNA metabarcoding than
27 morphology-based identification. For each sample, we detected more taxa by metabarcoding than by previous
28 morphological methods, and all four primer sets showed similarly good performance. There was a significant linear
29 correlation between sequence abundance and the number of taxa in each sample, but the scatter was up to two orders of
30 magnitude. Ecological status assessment indices calculated from morphological and DNA metabarcoding datasets were



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

mostly similar, with a few exceptions. With the recent drop in sequencing costs per sample, both methods identification are currently equally expensive.

4) We used actual samples for monitoring to demonstrate that DNA metabarcoding can achieve similar results and better taxonomic resolution than current morphological identification methods. Metabarcoding has thus already become a viable and reliable invertebrate identification method for stream assessment. However, to unlock the full potential of DNA metabarcoding for ecosystem assessment key problems in current laboratory protocols and reference databases, specified in this work, will require further attention.

**Keywords:** Biomass bias, next generation sequencing, macroinvertebrates, metabarcoding, DNA barcoding, ecological status

42 ## Introduction

43 Macroinvertebrates are key biological quality indicators in national and international aquatic biomonitoring programs.
44 A variety of bioassessment protocols are used in these monitoring programs (Birk *et al.* 2012). However, in all current
45 protocols, biological quality components such as macroinvertebrates, diatoms, macroalgae, and fish, are identified
46 based only on morphological properties. For benthic macroinvertebrates, the orders Ephemeroptera, Plecoptera,
47 Trichoptera, and Diptera are often sensitive to pollution, and thus are ideal indicators of stressors affecting stream
48 ecosystems. Unfortunately, identifying benthic taxa to species or even genus level is often difficult or impossible, and
49 misidentification is frequent and highly dependent on experience (Sweeney *et al.* 2011). Thus, classification is often
50 only performed to a high taxonomic level. However, different responses to stressors are possible at the species level
51 (Macher *et al.* 2015), and these differences can go unnoticed with low taxonomic resolution. Furthermore, using human
52 experts for morphological identification is time time-consuming and therefore expensive (Yu *et al.* 2012; Aylagas *et al.*
53 2014). Misidentification, low comparability, and limited taxonomic resolution for difficult groups, such as chironomids,
54 can lead to inaccurate assessments and potentially to the mismanagement of stream ecosystems (Stein *et al.* 2013).
55 While also the use of computer vision based taxa identification has been explored (e.g. (Kiranyaz *et al.* 2011; Ärje *et al.*
56 2013; 2017)), in recent years DNA-based taxon identification has emerged as a potential alternative to morphological
57 methods. The first DNA based case studies highlight the potential application of these methods to the assessment of
58 freshwater macroinvertebrates (Hajibabaei *et al.* 2011; Carew *et al.* 2013; Elbrecht & Leese 2015; 2016b). DNA
59 barcoding in particular has often been promoted as a useful tool for ecosystem monitoring and assessment (Baird &
60 Sweeney 2011; Baird & Hajibabaei 2012; Taberlet *et al.* 2012). While several studies have established a number of
61 benefits of DNA-based monitoring using DNA metabarcoding, additional large-scale studies of complete freshwater
62 macroinvertebrate monitoring samples are needed. In marine, freshwater and terrestrial ecosystems, complete samples
63 of arthropods and diatoms have been processed (Ji *et al.* 2013; Gibson *et al.* 2014; Zimmermann *et al.* 2014; Leray &
64 Knowlton 2015) and used to obtain assessment indices (Aylagas *et al.* 2016). However, DNA metabarcoding studies of
65 complete macroinvertebrate samples from freshwater ecosystems are often limited to a few sampling sites (Hajibabaei
66 *et al.* 2011; 2012) or selected taxon groups (Carew *et al.* 2013). Only (Gibson *et al.* 2015) performed large-scale studies
67 of Canadian macrozoobenthos and demonstrated that DNA metabarcoding outperforms family- and order-level
68 morphological identification approaches. While these results are promising, it should be noted that, in most European
69 monitoring programs, taxa are identified to species level. For the application of DNA metabarcoding to routine stream
70 monitoring, further optimized DNA-specific macrozoobenthos sampling and laboratory protocols have to be developed
71 and validated. We recently explored primer bias and tissue extraction protocols using mock invertebrate samples of
72 known composition with a one-step PCR metabarcoding protocol using the Illumina MiSeq sequencer (Elbrecht &



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

73 Leese 2015). We identified primer design as a critical component for species detection and developed primer sets
74 specifically targeting freshwater macroinvertebrates (Elbrecht & Leese 2016b,a). While these BF/BR primers work well
75 on mock communities and initial tests based on two stream benthos samples, they have not been tested in a larger
76 biomonitoring context (Elbrecht & Leese 2016b; Elbrecht *et al.* 2016). Further, the sufficiency of available reference
77 data, e.g., the BOLD database for freshwater macroinvertebrates (Ratnasingham & Hebert 2007), has not been fully
78 explored in a metabarcoding context. Finally also laboratory constraints specific to organisms and stream ecosystems
79 are may exist. Thus, it is important to further explore and validate the potential of DNA metabarcoding for routine use
80 in stream assessment.
81 In this study, we performed a one-to-one comparison of traditional morphological- and DNA metabarcoding-based
82 identification in the context of bioassessment using benthic macroinvertebrate communities at 18 sites through a
83 national Finnish stream bioassessment program. All samples were morphologically identified by an experienced
84 taxonomist. Thus, the samples are ideally suited for performance comparisons of morphological and DNA-based
85 identification in bioassessments and enable the critical evaluation of the current limitations of both approaches.
86

87 ## Materials and Methods

88 ### Sample collection and processing

89 Benthic macroinvertebrates were collected in the fall of 2013 at 18 riffle sites in Finland as part of an official national
90 stream monitoring program (Figure S1, Table S1). At each site, four 30 s kick-net samples covering most microhabitats
91 at each site were obtained following national guidance for Water Framework Directive (WFD) monitoring (Meissner *et*
92 *al.* 2016a). The samples were preserved in 70% ethanol in the field, and the invertebrates in each sample were later
93 sorted in the laboratory. Collected specimens were stored in 70% ethanol, which was not replaced after collection,
94 leading to an average ethanol concentration of 65.14% (SD = 2.83%) during long-term storage.
95
96 All specimens were counted and identified based on morphology, mainly to the species or genus level, with the
97 exception of Oligochaeta, Turbellaria, Nematoda, Hydracarina, and the dipteran families Chironomidae and Simuliidae,
98 which were counted, but not identified to a lower taxonomic level. The level of identification followed the WFD
99 monitoring guidance targeting operational taxonomic units (OTUs) established by the Finnish Environment Institute
100 SYKE ((Meissner *et al.* 2016a), see page 29). Identification was performed by a single experienced consultant who
101 participated in the most recent international macroinvertebrate taxonomic proficiency tests organized by Proftest of
102 SYKE in 2016 and scored higher than average (i.e., >95%, (Meissner *et al.* 2016b)). For each site, four samples were



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

103 then pooled (hereafter referred to as a single sample). The samples were kept cool (8°C) for subsequent molecular
104 analyses.
105
106 **DNA extraction and tissue pooling**

107 To remove the ethanol, specimens from each sampling site were dried overnight in sterile Petri dishes. Specimens were
108 subsequently homogenized using an IKA ULTRA-TURRAX Tube Drive Control System with sterile 20 mL tubes and
109 10 steel beads (5 mm Ø) by grinding at 4000 rpm for 30 min. From each sample, three aliquots containing 15 mg (6.23
110 mg) of homogenized tissues were used for DNA extraction. The tissue was digested following a modified salt DNA
111 extraction protocol (Sunnucks & Hales 1996). Then, 15 µL of DNA was pooled from each of the three extraction
112 replicates, digested with 1 µL of RNAse A, and cleaned using a MinElute Reaction Cleanup Kit (Qiagen, Venlo,
113 Netherlands) following the manufacturer's instructions. DNA concentrations were quantified using the Fragment
114 Analyzer™ Automated CE System (Advanced Analytical, Heidelberg, Germany), and the concentrations of all samples
115 were adjusted to 25 ng/µL DNA for PCR.
116
117 **DNA metabarcoding and bioinformatics**

118 All 18 samples were amplified in duplicate with four BF/BR freshwater macroinvertebrate fusion primer sets, described
119 previously (Elbrecht & Leese 2016b). Table S2 gives an overview of the fusion primer combinations used for sample
120 tagging with inline barcodes. Each PCR reaction was composed of 1× PCR buffer (including 2.5 mM $Mg^{2+}$), 0.2 mM
121 dNTPs, 0.5 µM each primer, 0.025 U/µL HotMaster Taq (5Prime, Gaithersburg, MD, USA), 12.5 ng of DNA, and
122 HPLC $H_2O$ to obtain a total volume of 50 µL. PCRs were run using a Biometra TAdvanced Thermocycler with the
123 following program: 94°C for 3 min, 40 cycles of 94°C for 30 s, 50°C for 30 s, and 65°C for 2 min, and 65°C for 5 min.
124 For a few of the samples, it was necessary to use a larger PCR volume (250 µL) due to PCR inhibitors present in the
125 samples (see Table S2). PCR products were purified and left sided size-selected using SPRIselect with a ratio of 0.76×
126 (Beckman Coulter, Brea, CA, USA). They were quantified using a Qubit Fluorometer (HS Kit, Thermo Fisher
127 Scientific, Waltham, MA, USA) and Fragment Analyzer™ Automated CE System (Advanced Analytical Technologies
128 GmbH, Heidelberg, Germany). Samples were pooled to equal molarity, and sequened on two Illumina HiSeq 2500
129 lanes using a rapid Run 250 bp PE v2 Sequencing Kit and 5% PhiX spike-in. Sequencing was carried out by GATC
130 Biotech GmbH (Konstanz, Germany).
131 For bioinformatic processing, the UPARSE pipeline was used in combination with custom R scripts (Dryad DOI) for
132 data processing (Edgar 2013). Scripts are available on http://github.com/VascoElbrecht/JAMP (JAMP v0.10a). Reads
133 were demultiplexed and paired-end reads were merged using Usearch v8.1.1861 with the following settings: -



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

134 fastq_mergepairs with -fastq_maxdiffs 99, -fastq_maxdiffpct 99 and -fastq_trunctail 0 (Edgar & Flyvbjerg 2015).
135 Primers were removed using cutadapt version 1.9 with default settings (Martin 2011). Sequences were trimmed to the
136 same 217 bp region amplified by the BF1+BR1 primer set and the reverse complement build, if necessary, using
137 fastx_truncate and fastx_revcomp. Only sequences of 207–227 bp were used for further analysis (filtered with cutadapt).
138 Low-quality sequences were then filtered from all samples using fastq_filter with maxee = 0.5. Sequences from all
139 samples were then pooled, dereplicated (minuniquesize = 3), and clustered into molecular operational taxonomic units
140 (MOTUs) using cluster_otus with a 97% identity threshold (Edgar 2013) (includes chimera removal).
141 Pre-processed reads (Figure S2, step B) for all samples were dereplicated again using derep_fulllength, but singletons
142 were included. Sequences from each sample were matched against the MOTUs with a minimum match of 97% using
143 usearch_global. Only OTUs with a read abundance above 0.003% in at least one sequencing replicate were considered
144 in downstream analyses. Taxonomic assignments for the remaining MOTUs were determined using an R script to
145 search against the BOLD and NCBI databases. Taxonomic information was not further validated, and, in the case of
146 conflicting assignments between NCBI and BOLD databases, the taxonomic level where both databases returned
147 identical results was used. For assignment to the species level, a hit with 98% similarity was required in at least one of
148 the two databases, and 95% similarity was required for assignment to the genus level, 90% for the family level, and 85%
149 for the order level. Only MOTUs that matched macroinvertebrates were used in the statistical analysis. In all further
150 analyses, only MOTUs with a sequence abundance of at least 0.003% in both replicates of a sample were included.
151

152 # Results

153 ## Sequencing run statistics

154 The HiSeq Rapid run yielded 260.75 million read pairs (raw data available at SRA, accession number SRR4112287).
155 After library demultiplexing on average 1.53 million (SD = 0.29 million) read pairs were retained (Figure S3).
156 Unexpected sample tagging combinations (potential tag switching) were low with only 12 of 136 unused combinations
157 above the 0.003% read abundance threshold and a maximum relative read abundance of 0.006% (Figure S4). After
158 bioinformatic processing a total of 750 MOTUs remained, of which 49.3% were shared between all four primer sets
159 (Figure S5). The primer combination BF2+BR2 generated the highest number of MOTUs. Sequencing replicates for
160 each sample showed a mean difference in sequence abundance of a factor of 2.05 (expected 1.0), indicating high
161 variation in sequence abundance between the two replicates (Figure S6). There is a weak but significant negative
162 correlation between relative read abundance per MOTU and variation between replicates for 13 out of 72 total samples
163 ($p \leq 0.05$, Figure S6), but there was no consistent pattern across all samples, as also highly abundant MOTUs showed
164 large differences between replicates in some cases.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

165

## 166 Taxonomic identification

167 A total of 126 taxa were identified based on morphology across all 18 samples of which 61.1% were identified to
168 species level (Table S3). Five species lacked reference sequences in BOLD or NCBI (Table S2), and more taxa were
169 potentially missing at a lower taxonomic resolution (e.g., reference data for specimens only identified to the family
170 level). All samples were dominated by a few common taxa whereas rare taxa were only found in some samples (mean
171 Pielou's evenness = 0.65, SD = 0.12, Figure S7).

172 In total, 750 MOTUs remained in the dataset after bioinformatic processing of the sequence data. Of these, 573 target
173 invertebrate hits were further analyzed. The MOTU table for DNA metabarcoding with taxonomic assignments as well
174 as MOTU sequences is available as supplementary Table S4. After taxonomic assignment using BOLD and NCBI,
175 DNA metabarcoding revealed the presence of 288 morphotaxa (208 species, 47 genera, 23 family, 10 order or higher
176 level). More taxa were resolved at species level with metabarcoding than by morphological determination (Figure S8).
177 DNA metabarcoding consistently detected a substantially greater number of taxa than morphology-based identification
178 across all samples with each primer combination (57.30% more taxa on average over all data, SD = 35.69%, Figure 1).
179 The difference between methods was mainly explained by the coarse taxonomic level obtained using morphology-based
180 identification for dipteran families, mites, Oligochaeta, and Limnephilidae. For groups that were morphologically
181 identified to species or genus level, DNA metabarcoding detected 25.3% more taxa using OTUs. Despite enabling the
182 identification of a substantial number of overlooked taxa, DNA metabarcoding did not detect an average of 32.51% (SD
183 = 9.71%) of taxa identified based on morphology in each sample (Figure 2, see Table S2 for undetected taxa). The
184 proportion of detected taxa was similar for all primer pairs, with a slightly higher detection rate for the BF2+BR2 pair
185 than for other combinations. There was a significant positive correlation between the number of reads assigned to each
186 morphotaxon and the number of specimens per taxon for most samples and primer combinations (Figure S9). This
187 correlation was significant for all 18 samples for the combination BF2+BR2, while for other primer combinations it was
188 only significant for 13 or 14 samples. However, despite the positive correlations between read abundance and number
189 of taxa, there is still a variation of two orders of magnitude in read abundance. This was also reflected by the low
190 adjusted $R^2$ values (mean = 0.386, SD = 0.130).

191 Assessment indices calculated from morphology and DNA metabarcoding data were generally similar (Figure 3, Table
192 S1). For only a few samples the assessed ecological status changed with the DNA-based taxa lists. Most of the different
193 assessments were obtained for DNA-identification based calculation of the PMA (Percent model affinity index, (Novak
194 & Bode 1992)) which assigned many samples to only good status, whereas morphological identification did lead to



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

195 more high status assessment results (Figure 3C). Note that for this comparison, the DNA-based species lists had to be
196 reduced to the guidance target OTU list used for Finnish assessment.
197

## 198 Discussion

### 199 Performance of DNA metabarcoding

200 Our results clearly show that DNA-based identification methods can capture more diversity than routine morphological
201 ones despite the fact that several of the morphologically identified taxa were not recovered with the DNA-based
202 technique. DNA metabarcoding was especially powerful in resolving taxon diversity for those groups that are difficult
203 to determine morphologically in their larval stages such as many chironomids and simuliids. In addition, also EPT taxa
204 that were morphologically identified only to family (Limnephilidae) or genus level (e.g., *Eloeophila* and *Hydroptila*)
205 could be identified to species level using DNA metabarcoding. Our observed increased taxonomic resolution of DNA
206 over morphologically based identification has been demonstrated in many previous studies (Baird & Sweeney 2011;
207 Sweeney *et al.* 2011; Stein *et al.* 2013; Gibson *et al.* 2015). All four of our applied macroinvertebrate-specific BF/BR
208 primer combinations showed similarly good performance, consistent with our previous mock community tests (Elbrecht
209 & Leese 2016b).
210 When using current ecological status related metric calculation methods, results were very similar for morphology-
211 based and our DNA-based taxon lists, indicating that metabarcoding is on par with current assessment methods. A good
212 match of morphological assessments against presence/absence data as well as DNA-based taxon lists has also been
213 demonstrated in pioneering marine studies (Aylagas *et al.* 2014; 2016). Considering that the indices used in this study
214 were optimized for morphological identification, in some cases not considering species or genus level but even more
215 coarse taxonomic levels, future DNA-based assessment might be further improved by applying optimized metric
216 calculation approaches and species level trait databases (Schmidt-Kloiber & Hering 2015). DNA metabarcoding can
217 provide much more accurate taxonomic identification than can morphology and can even be used to detect cryptic
218 species (Elbrecht & Leese 2015). This gain in accuracy gives us a chance to investigate potential differences in
219 ecological preferences and allows stressor detection based on indicator taxa when larval morphology alone is not
220 sufficient (Macher *et al.* 2015). This valuable additional information could be integrated into future assessment indices,
221 and might not only clarify the condition of a stream, but can also potentially help to disentangle the precise stressors
222 affecting the ecosystem.
223 While DNA metabarcoding has the advantages of increased reproducibility and taxonomic resolution, it also has
224 drawbacks, including the inability to quantify abundance or biomass (Piñol *et al.* 2014; Elbrecht & Leese 2015).
225 Despite the development of approaches to adjust for primer bias (Thomas *et al.* 2015), these are unlikely to succeed for



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

226 complex diverse communities and are additionally affected by biomass (Elbrecht *et al.* 2016). Nevertheless, we found a
227 significant linear correlation for most samples between morphologically identified taxon abundance and number of
228 sequencing reads assigned to the respective OTUs. While this could be interpreted as an opportunity to estimate taxon
229 biomass, one has to acknowledge the poor fit and the scatter of up to two orders of magnitudes. While the BF/BR
230 primers show less primer bias than previously tested Folmer primers (Folmer *et al.* 1994; Elbrecht & Leese 2015), this
231 bias is still substantial (Elbrecht & Leese 2016b). Sequence abundance might be further influenced by specimen
232 biomass of different taxa. We could only determine the correlation between read number and taxon abundance, which is
233 a proxy for biomass. With exact biomass data for each specimen, the linear correlation and thus the estimate of biomass
234 with respect to abundance might be further improved. We nevertheless argue that estimating biomass from PCR-based
235 metabarcoding analyses remains very difficult owing to primer bias, but could provide utility.
236 Laboratory and sequencing costs are critical for the viability of large-scale DNA-based monitoring (Ji *et al.* 2013). In
237 this study, our ready to load library was sequenced using the HiSeq Rapid system, generating 3 million read pairs per
238 sample (1.5 per replicate), for a total cost of 7900€ by a commercial sequencing provider. We sequenced each sample
239 using four primer pairs to examine the performance of different primer pairs, resulting in a total of 72 samples being
240 sequenced with two replicates each. This results in a per-sample sequencing cost of 110€. For routine monitoring in the
241 future, only one best-working primer pair would be used. Sequencing costs can be further reduced by pooling more
242 samples in the same sequencing run and pre-sorting samples according to biomass, even though specimen sorting would
243 increase the cost of sample processing (Elbrecht *et al.* 2016). In the protocol used here, standard reagent costs for DNA
244 extraction, single-step PCR, library quantification, and clean-up were approximately 70€. This results in a combined
245 cost of reagents plus sequencing of 180€ per sample, which is in line with previously estimated costs (Stein *et al.* 2014).
246 Laboratory rent and maintenance, sample collection, personnel costs, and bioinformatics costs probably push the total
247 costs per sample to 500–750€, which is comparable to the current cost of Finnish morphology-based costs per site.
248 Sample collection is a large contributor to these expenses, as laboratory work and bioinformatics can be highly
249 parallelized and automated.
250

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

252 The DNA metabarcoding protocol we utilized worked reliably across all 18 samples. However, we identified various
253 opportunities to further improve the performance of metabarcoding. Figure 4 gives an overview of the limitations of
254 DNA metabarcoding related to taxonomic assignment and the reference database as well as to the laboratory protocols.
255 Across all 18 samples in this study, we were unable to detect 32% of taxa identified morphologically using our
256 metabarcoding approach. Some of these are linked to the practice of the cautionary principle of human experts to leave



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

257 the identification of small specimens to coarser taxa (e.g. genus level) if higher taxonomic resolution cannot be
258 established without doubt. Additionally, laboratory procedures used in routine monitoring campaigns may not be
259 adequate for DNA extraction. The low alcohol concentration typically used for sample preservation during routine
260 biomonitoring may result in specimens that are still viable for morphological detection, but have strongly degraded
261 DNA, impairing detection. Further, although unlikely, given the proven proficiency of the expert used in the
262 morphological identification, the introduction of false taxa through erroneous morphotaxonomical identification by the
263 human expert may have influenced the observed discrepancy between the results of the identification methods. Several
264 additional factors listed in Figure 4 may have influenced detection, positively or negatively.

265 Laboratory methodology can strongly bias the absolute and relative amounts of invertebrates detected by DNA
266 metabarcoding. Primer bias is one of the most common concerns, as primer and template mismatches can prevent
267 certain taxa from being amplified by PCR (Deagle *et al.* 2014; Piñol *et al.* 2014; Elbrecht & Leese 2015). By adding
268 primer degeneracy and carefully choosing primer sets suited for the targeted ecosystem and taxonomic groups, the
269 negative effects of primer bias can be reduced (Elbrecht & Leese 2016b,a). However, even with primer optimization,
270 one-step PCR methods will be affected by primer bias. Thus, it is unlikely that all taxa present in a sample can be
271 detected by DNA metabarcoding, and abundance or biomass estimation is difficult owing to primer bias. PCR and
272 sequencing errors, undetected chimeras, and misidentified reference sequences can further lead to false positive
273 detection. Additionally, specimens in a sample can vary widely in biomass, depending on species and life stage. This
274 not only prevents the estimation of taxon abundances, but can also lead to a lack of detection of small and rare taxa
275 (Elbrecht *et al.* 2016). Our data were likely affected by this bias, as 68.3% of taxa were present in just five or fewer
276 samples. The issue of primer bias and variation in taxon biomass makes it difficult to relate read abundance to taxon
277 abundance. While presence/absence data might already be sufficient for ecosystem assessment (Aylagas *et al.* 2014),
278 one has to acknowledge that relative abundance-based estimates might be possible if identical protocols are used across
279 all sample sites, leading to similar biases across samples. Some of our samples were additionally affected by PCR
280 inhibition, which could be solved by using larger PCR volumes to dilute PCR inhibitors or by additional clean-up steps.
281 However, PCR inhibition is a major issue for the application of DNA metabarcoding to monitoring, as the protocol
282 should work in all stream ecosystems, independent of environmental conditions. Ideally, methods to purify DNA from
283 complete kick samples are developed and tested without pre-sorting specimens from debris (e.g., sediment, small stones,
284 leaves, and organic particles). This would allow to skip the time consuming pre-sorting steps, during which, up to 30%
285 of the specimens can be missed (Haase *et al.* 2010). Thus, circumventing preprocessing would include a large
286 proportion of the currently potentially unstudied diversity.

287 Metabarcoding might be affected by several other laboratory-specific factors. For example, tag switching is an issue
288 when several samples are multiplexed in one library (Esling *et al.* 2015; Schnell *et al.* 2015). However, we did not



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

289 observe these effects on our samples or in our previous studies using the fusion primer system with in line tagging.
290 However, (O'Donnell *et al.* 2016) showed that tags can lead to biases in read abundance, and our samples are
291 potentially affected by this, as evidenced by the observed variation in read abundances of a average factor of ~2
292 between the two replicates for the same sample. It is of critical importance to validate that tag switching is at a
293 minimum and the level of bias between replicates and its effect on the data are known. In our case, variation in read
294 abundance will potentially result in minimal underestimation of diversity, as we conservatively discarded all reads not
295 present in both replicates. Although we obtained good taxonomic resolution, it is important to be aware of and account
296 for these shortcomings and to develop modifications to the protocols to solve these problems. DNA metabarcoding is
297 not perfect; many different protocols are being developed, of which few have been extensively validated. Method
298 groundtruthing is essential to build trust in metabarcoding methods for monitoring, and different protocols and
299 modifications should be validated using the same standard invertebrate mock communities. Such validation samples
300 would not only reveal biases, but could also be used to accredit monitoring offices to ensure that their laboratory work
301 meets quality standards and that results are comparable with those of other accredited offices. Once a well-established
302 standardized metabarcoding protocol is developed, the analysis of high-throughput metabarcoding data could be carried
303 out on cloud-based systems. This would facilitate comparisons and easy updating of all bioinformatic analyses. Further,
304 common metadata standards and central storage of all monitoring related metabarcoding data could be a valuable
305 resource for research, e.g. by providing accurate maps of taxa presence over a large geographic and temporal scales
306 with unprecedented accuracy.
307 The second major factor influencing our results is database accuracy and the reliability of morphology-based
308 identification of specimens deposited in databases. Here, we specifically constrained our comparison to MOTUs with
309 assigned taxonomic information from the BOLD and NCBI reference databases and did not consider other MOTUs,
310 despite the potential to further increase assessment accuracy. Within the framework of the Water framework Directive
311 ecological assessment of aquatic ecosystems in many countries is currently taxa and associated traits or indicator values;
312 thus metabarcoding has to compete on the same level. While we think it is feasible to infer traits by correlating MOTUs
313 with abiotic data from sampling locations, we currently lack appropriately large metabarcoding datasets to verify this.
314 Further, it is desirable to keep and associate taxonomic information with MOTUs, to relate ecological information to
315 obtained sequences, but also associate correlative found traits and ecological preferences back to the taxa detected with
316 metabarcoding. Currently, available databases are still incomplete and not all taxa have barcodes. Additionally, the
317 identification accuracy for larvae and adult invertebrates varies depending on expert experience, and even databases like
318 BOLD, specifically built for DNA barcoding, contain misidentified taxa or conflicting taxonomic assignments for the
319 same BIN. Databases require stricter standards and quality control, including incentives for data providers and
320 managers to better curate their data after the initial release. Sample degradation and misidentification in the 18 samples,



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

321 but also in the reference databases, could have led to false positives or negatives in both the morphologically generated
322 taxon list and our metabarcoding-based assessments. It is imperative that taxonomical experts and molecular biologists
323 come together to discuss and solve conflicting cases, especially as traditional taxonomic expertise is fading. DNA
324 metabarcoding provides an excellent opportunity for traditional taxonomists to contribute to reference databases and,
325 thanks to the increased taxonomic resolution using DNA barcoding, to associate ecological information with difficult
326 groups, like Diptera.
327

## 328 Conclusions

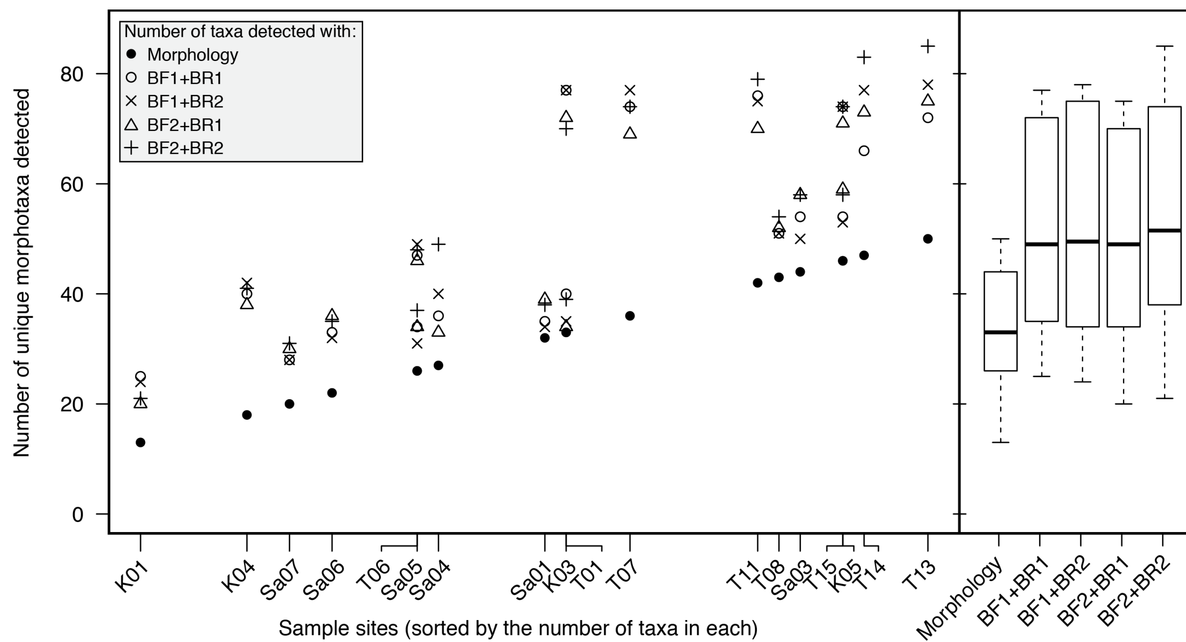
329 We demonstrated that DNA metabarcoding is a viable alternative to morphology-based identification of
330 macroinvertebrates as both costs and assessment results are very similar. DNA metabarcoding detected more taxa in all
331 samples, which ,if linked with ecological species traits, could potentially improve assessment results over those
332 obtained through morphological identifications. Despite its merits, several shortcomings of DNA-based method as well
333 as challenges with reference databases have to be tackled before unlocking the full potential of DNA metabarcoding.
334 This will require coordinated efforts such as the DNAquanet project combining work from molecular biologists,
335 ecologists and taxonomists (Leese *et al.* 2016).
336
337
338 **Acknowledgements:** FL and VE were supported by a grant of the Kurt Eberhard Bode Foundation to FL. KM was
339 supported by the Academy of Finland project DETECT (289104). We would also like to thank XXXXX for
340 proofreading this manuscript.
341
342 **Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content
343 and writing of the paper.
344
345 **Author contributions statement:** VE, KM, and FL conceived the ideas and designed the methodology. EV carried out
346 the laboratory work; VE performed bioinformatic analyses together with EV; JA calculated assessment indices; VE led
347 the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.
348
349



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

PeerJ Preprints NOT PEER-REVIEWED

350 ## Figures




351

352 **Figure 1**: Number of morphotaxa detected by morphological and DNA-based identification methods across all 18

353 sample sites. The number of taxa detected by DNA-based identification was compared among four primer pairs

354 (different symbols). The boxplot on the right compares the overall performances of DNA- and morphology-based

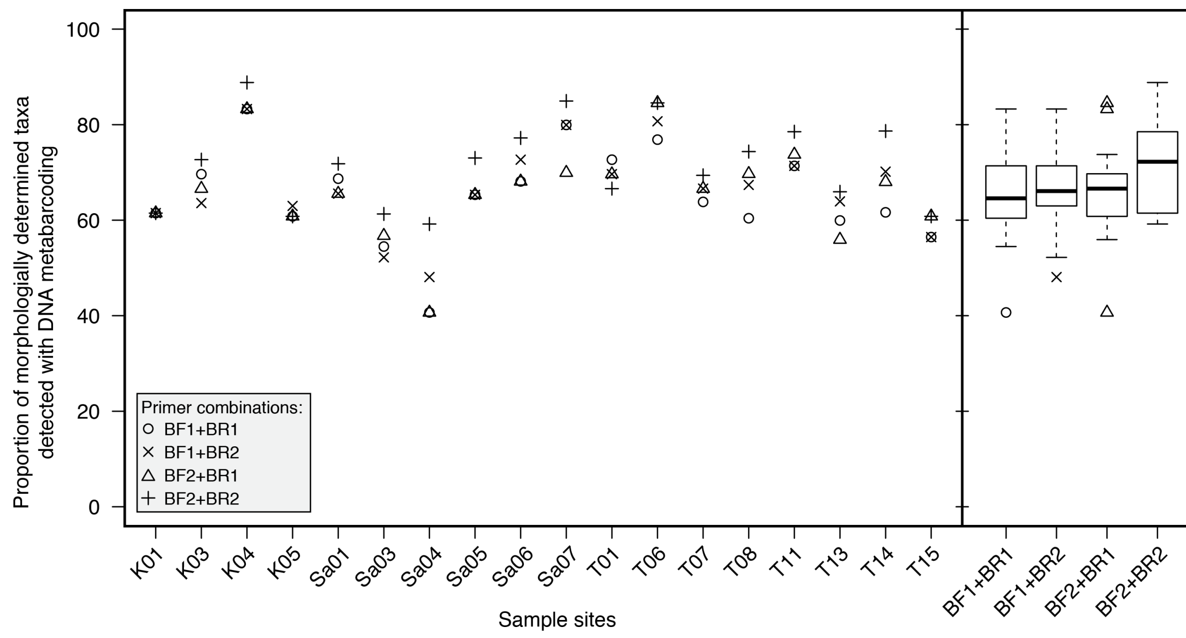
355 identification across samples.

356

357



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017




358

359 **Figure 2**: Proportion of morphologically identified taxa detected with the four different primer pairs across all 18

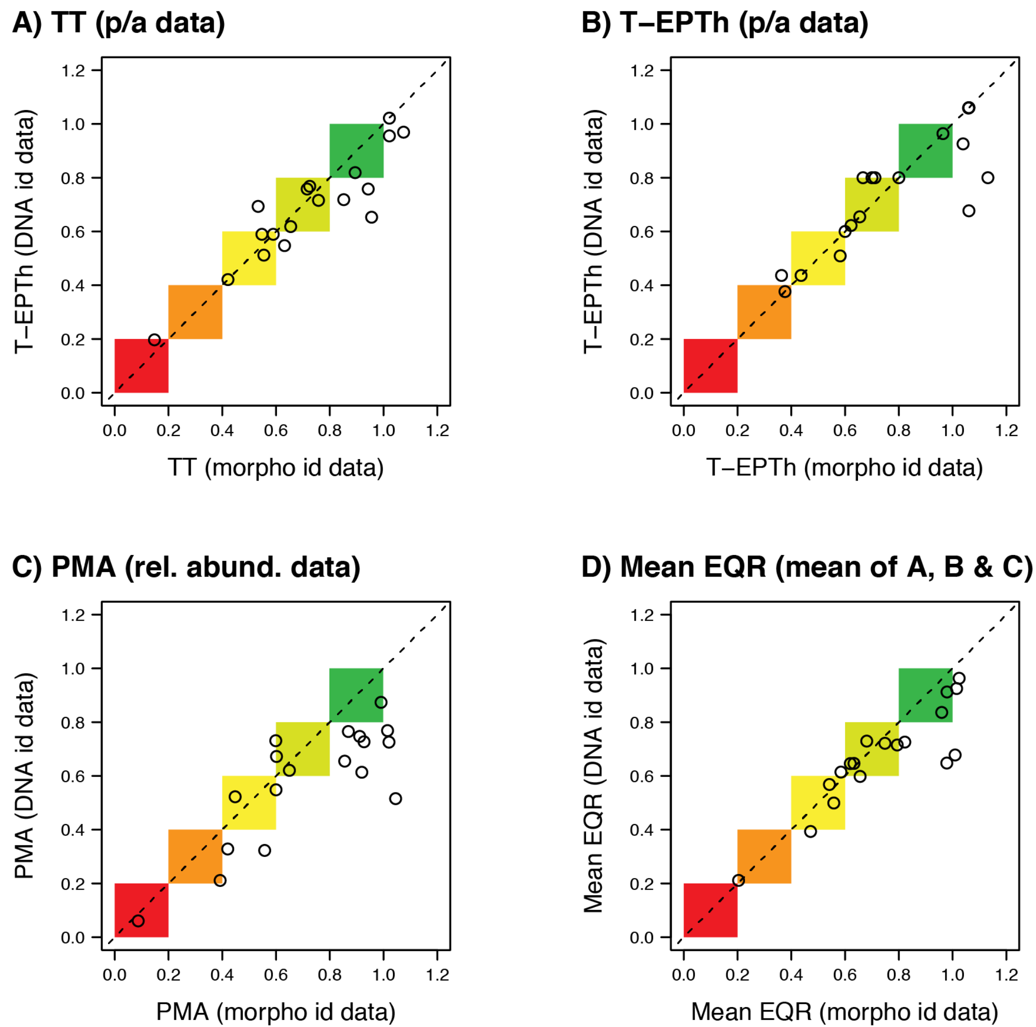
360 sample sites. Primers pairs are indicated by different symbols, and overall detection rates for the primer pairs are shown

361 on the right.

362



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

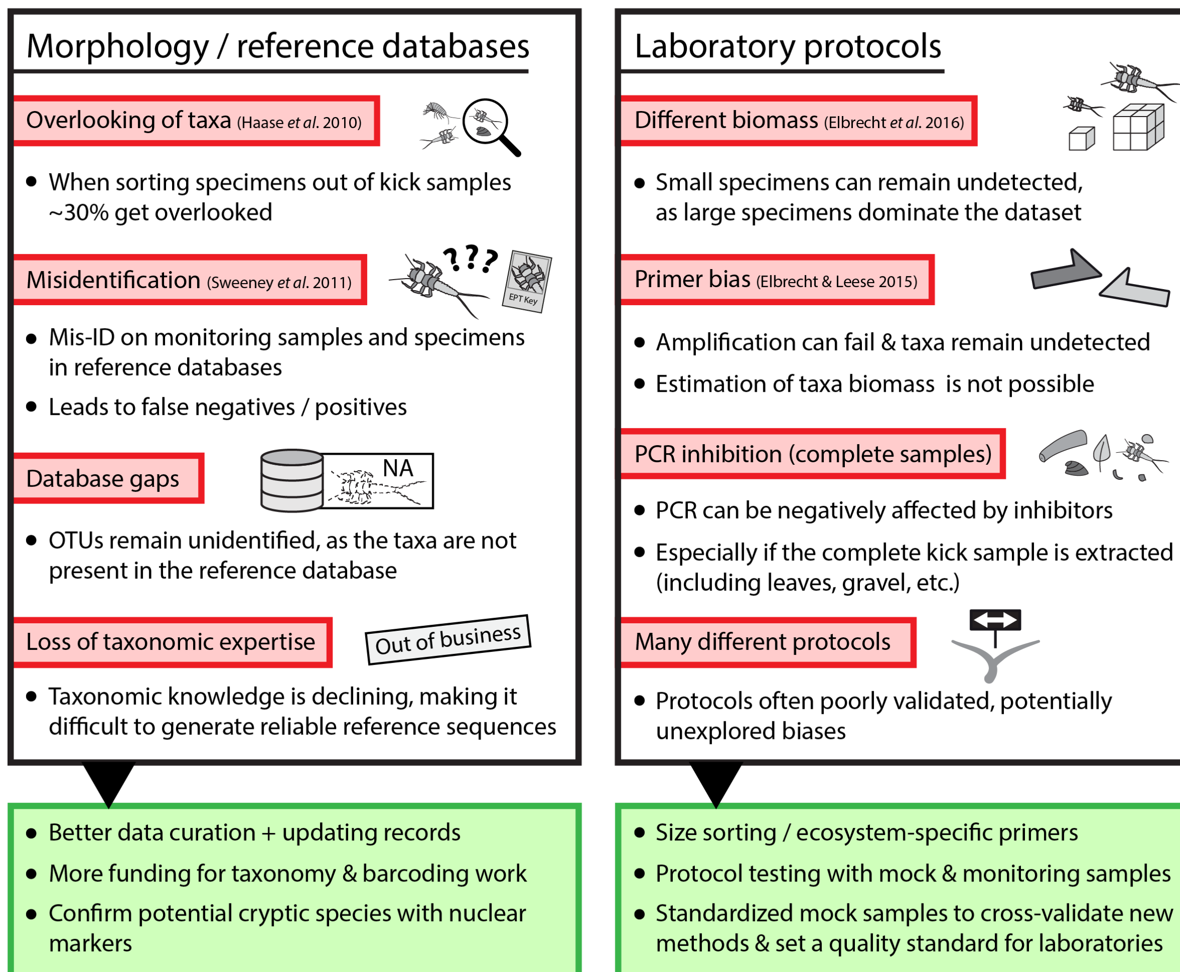



363

364 **Figure 3:** Comparison of Finnish macroinvertebrate WFD assessment indices calculated with taxa lists based on
365 morphological- and DNA-based (BF2+BR2 primer) identification. The three indices are shown as normalized
366 Ecological Quality Ratios (EQR) ranging from 0 (Bad status) to 1 (High status with no anthropogenic alteration). For all
367 four indices, there was a significant correlation between morphological- and DNA-based assessments (Pearson
368 correlation, $p > 0.0001$). A) Occurrence of river Type-Specific Taxa (TT, based on p/a data). B) Occurrence of river
369 Type-Specific EPT-families (EPTh, based on p/a data). C) Percent model affinity (PMA, based on relative abundance
370 data). D) Mean EQR of the three indices.



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017




371

372 **Figure 4:** Overview of factors currently limiting the application of DNA metabarcoding for ecosystem assessment, with

373 potential solutions.

374

375



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

376

377 **References**

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

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

382 Ärje, J., Kärkkäinen, S., Meissner, K. & Iosifidis, A. (2017). The effect of automated taxa identification errors on
383 biological indices. *Expert Systems with …*.

384 Ärje, J., Kärkkäinen, S., Turpeinen, T. & Meissner, K. (2013). Breaking the curse of dimensionality in quadratic
385 discriminant analysis models with a novel variant of a Bayes classifier enhances automated taxa identification of
386 freshwater macroinvertebrates. *Environmetrics*.

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

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

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

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

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

398 Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, **10**,
399 996–998.

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

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

405 Elbrecht, V. & Leese, F. (2016a). PrimerMiner: an rpackage for development and in silicovalidation of DNA
406 metabarcoding primers (M. Bunce, Ed.). *Methods in Ecology and Evolution*, 1–5.

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

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

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

413 Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial
414 cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology*,
415 **3**, 294–299.

416 Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I. & Hajibabaei, M. (2015). Large-Scale
417 Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing (D. Fontaneto, Ed.). *PloS*



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

418 *one*, **10**, e0138432–15.

419 Gibson, J., Shokralla, S., Porter, T.M., King, I., van Konynenburg, S., Janzen, D.H., Hallwachs, W. & Hajibabaei, M.
420 (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods
421 through DNA metasystematics. *Proceedings of the National Academy of Sciences*, **111**, 8007–8012.

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

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

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

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

434 Kiranyaz, S., Ince, T., Pulkkinen, J., Gabbouj, M., Ärje, J., Kärkkäinen, S., Tirronen, V., Juhola, M., Turpeinen, T. &
435 Meissner, K. (2011). Classification and retrieval on macroinvertebrate image databases. *Computers in biology and*
436 *medicine*, **41**, 463–472.

437 Leese, F., Altermatt, F., Bouchez, A. & Ekrem, T. (2016). DNAqua-Net: Developing new genetic tools for
438 bioassessment and monitoring of aquatic ecosystems in Europe. *RESEARCH IDEAS …*.

439 Leray, M. & Knowlton, N. (2015). DNA barcoding and metabarcoding of standardized samples reveal patterns of
440 marine benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America*,
441 201424997–6.

442 Macher, J.N., Salis, R.K., Blakemore, K.S., Tollrian, R., Matthaei, C.D. & Leese, F. (2015). Multiple-stressor effects on
443 stream invertebrates: DNA barcoding reveals contrasting responses of cryptic mayfly species. *Ecological*
444 *Indicators*, 1–11.

445 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, **17**,
446 10–12.

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

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

453 Novak, M.A. & Bode, R.W. (1992). Percent model affinity: a new measure of macroinvertebrate community
454 composition. *Journal of the North American …*.

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

457 Piñol, J., Mir, G., Gomez-Polo, P. & Agustí, N. (2014). Universal and blocking primer mismatches limit the use of
458 high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular ecology resources*,
459 **15**, 1–12.

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

462 Schmidt-Kloiber, A. & Hering, D. (2015). www.freshwaterecology.info – An online tool that unifies, standardises and



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

463 codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecological Indicators*,
464 **53**, 271–282.

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

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

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

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

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

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

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

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

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

486

487

488 **Supporting information**

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

490

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

492

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

494

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

496

497 **Figure S5.** Plot showing the numbers of shared OTUs between primer sets.

498

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



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017

500

501 **Figure S7.** Morphotaxa presence across samples.

502

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

504

505 **Figure S9.** Correlation between sequence abundance and morphotaxon abundance.

506

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

508

509 **Table S2.** Tagging combination used in the metabarcoding library.

510

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

512

513 **Table S4.** OTU table.

514



PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2759v1 | CC BY 4.0 Open Access | rec: 28 Jan 2017, publ: 28 Jan 2017