OPINION

WILEY Freshwater Biology

Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring

Vasco Elbrecht[1] | Dirk Steinke[1,2]

[1]Centre for Biodiversity Genomics, University of Guelph, Guelph, Ontario, Canada

[2]Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

Correspondence
Vasco Elbrecht, Centre for Biodiversity Genomics, University of Guelph, Guelph, ON, Canada.
Email: elbrecht@uoguelph.ca

Abstract

1. The viability of DNA metabarcoding for assessment of freshwater macrozoobenthos has been demonstrated over the recent years. The method has matured to a stage where it can be applied to monitoring at a large scale, keeping pace with increased high-throughput sequencing capacity. However, workflows and sample tagging need to be optimised to accommodate for hundreds of samples within a single sequencing run.

2. Here, we conceptualise a streamlined metabarcoding workflow, in which samples are processed in 96-well plates. Each sample is replicated starting with tissue extraction. Negative and positive controls are included to ensure data reliability. With our newly developed fusion primer sets for the BF2 + BR2 primer pair up to three 96-well plates (288 wells) can be uniquely tagged for a single Illumina sequencing run. By including Illumina indices, tagging can be extended to thousands of samples.

3. We hope that our metabarcoding workflow will be used as a practical guide for future large-scale biodiversity assessments involving freshwater invertebrates. However, as this is just one possible metabarcoding approach, we hope this article will stimulate discussion and publication of alternatives and extensions to this method.

KEYWORDS
biomonitoring, fusion primer, high throughput sequencing, metabarcoding workflow, replication

1 | INTRODUCTION

Reliable monitoring of freshwater macroinvertebrate diversity is a key component in the assessment and management of stream ecosystems (Dudgeon et al., 2006; Vörösmarty et al., 2010). DNA-based identification methods such as metabarcoding provide promising alternatives (Baird & Hajibabaei, 2012) to morphological identification, which is often limited in resolution and dependent on taxonomic expertise (Sweeney, Battle, Jackson, & Dapkey, 2011). In addition to reducing human bias, DNA-based identifications can also lead to improved stream assessment (Stein, White, Mazor, Jackson, & Battle, 2013). Over the past few years, several studies have demonstrated the feasibility of metabarcoding-based monitoring of freshwater macroinvertebrates (Andújar et al., 2018; Carew, Pettigrove, Metzeling, & Hoffmann, 2013; Gibson et al., 2015; Hajibabaei, Shokralla, Zhou, Singer, & Baird, 2011). Despite some methodological limitations (presence/absence data, primer bias) (Piñol, Mir, Gomez-Polo, & Agustí, 2014; Elbrecht & Leese, 2015), assessment results are at least comparable if not superior to conventional morphology-based stream monitoring approaches (Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017; Emilson et al., 2018; Gibson et al., 2015). Some macroinvertebrate reference databases are already fairly comprehensive especially for common taxa (Carew et al., 2017; Curry, Gibson, Shokralla, Hajibabaei, & Baird, 2018).
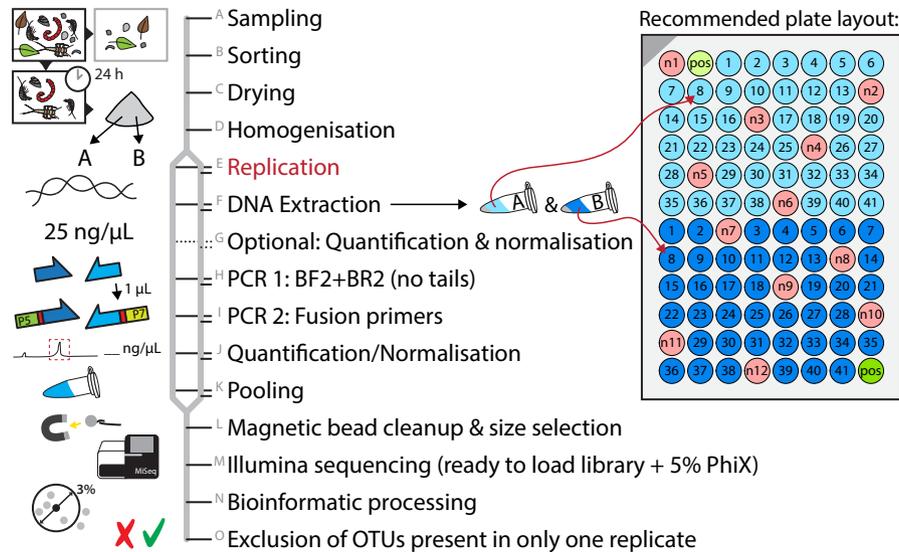
FIGURE 1 Overview of the proposed metabarcoding work flow for macroinvertebrates using a 96-well plate format and replication for each sample (in light and dark blue). Twelve negative controls are included (n1–n12) at the DNA extraction stage (F) to detect potential cross-contamination as well as tag switching. One positive control in replication (pos) can be used to estimate the overall performance of the metabarcoding run. Failed extractions or polymerase chain reactions (PCRs) can be excluded from the sequencing run, and repeated on a new plate. With the newly designed BF2 + BR2 fusion primers developed in this publication up to three 96-well plates can be multiplexed for a single run

Furthermore, many aspects of the metabarcoding approach have been thoroughly validated (Andújar et al., 2018; Carew et al., 2013; Elbrecht & Leese, 2015, 2017; Elbrecht, Peinert, & Leese, 2017; Emilson et al., 2018; Gibson et al., 2015; Hajibabaei, Spall, Shokralla, & van Konynenburg, 2012). Consequently, many countries are now actively working towards the use of DNA metabarcoding for routine monitoring of macroinvertebrates (Leese et al., 2018).

Routine stream monitoring requires the collection and identification of thousands of kick samples (Buss et al., 2015); however, current metabarcoding studies are typically limited to a few dozen samples (Andújar et al., 2018; Elbrecht, Vamos, et al., 2017; Emilson et al., 2018). If DNA metabarcoding is to be used in routine large-scale monitoring projects, a substantial scale up of laboratory protocols is needed in a way that ensures a high level of reliability and quality of data.

We propose a streamlined metabarcoding approach that runs up to 288 individual samples on a single Illumina sequencing run (Figure 1), using the BF2 + BR2 fusion primer system (Elbrecht & Leese, 2017), which has been shown to work well with macroinvertebrate monitoring samples (Elbrecht, Vamos, et al., 2017) but is currently limited to multiplexing of up to 72 samples per sequencing run (Elbrecht & Leese, 2017). Our new extended primer set allows for flexible multiplexing of samples in up to three 96-well plates thereby simplifying sample handling and reducing the risk of cross-contamination, while allowing for replication, as well as negative and positive controls. By incorporating replicates and positive and negative controls at the tissue homogenisation stage, samples affected by laboratory issues can be reliably detected and if necessary excluded from subsequent analysis. Furthermore, by minimising the number of validation steps throughout the protocol and the use of

streamlined fusion primer tags in a 96-well format, we ensure practicality of the protocol. Rather than continuously validating or replicating every step of the workflow, we recommend using controls and replicates in a manner that highlights samples affected by errors.

2 | SAMPLE COLLECTION, HOMOGENISATION AND DNA EXTRACTION

After samples are collected using a standardised protocol (Figure 1A, Buss et al., 2015), invertebrate specimens are usually separated from any debris such as substrate and non-target organic matter as part of the morphological identification process (Figure 1B). While this increases the chance that some taxa and specimens will be overlooked (Haase, Pauls, Schindehütte, & Sundermann, 2010), in most metabarcoding studies of freshwater macroinvertebrates, specimens were separated from debris (Carew et al., 2013; Elbrecht, Vamos, et al., 2017; Emilson et al., 2018; Gibson et al., 2015). This is often done as part of preceding morphological identifications, or out of the concern that homogenising an entire sample might introduce polymerase chain reaction (PCR) inhibitors and complicate standardisation. Although less work-intensive methods such as specimen flotation are currently being explored (Andújar et al., 2018), there is not enough evidence yet to decide if homogenisation of full kick samples without separating invertebrate specimens from collected substrate is also feasible. Once specimens are separated from debris they can be dried (Figure 1C) and homogenised (Figure 1D). Alternatively, DNA of bulk samples has also been extracted directly from the preservation ethanol (Hajibabaei et al., 2012), through homogenisation of the wet sample (Hajibabaei et al., 2011), or by lysing

the complete sample (Braukmann et al. 2018). To ensure complete homogenisation, it is recommended that dried bulk samples are ground using e.g. bead mills, as this allows DNA extraction of the entire community using just a small quantity of tissue powder (10–15 mg) (Elbrecht & Leese, 2015; Elbrecht, Vamos, et al., 2017). Two replicates per sample should be used for DNA extraction (Figure 1E), both of which are metabarcoded to facilitate the detection of insufficient tissue homogenisation. If homogenisation was incomplete taxon composition between replicates will vary substantially. Any DNA extraction method yielding high quality DNA can be used (e.g. silica-based spin columns, Figure 1F). However, as tissue powder is easily electrically charged, direct transfer of powder into the 96-well plate should be avoided. Rather, the powder should be incubated in individual 1.5-ml reaction tubes that may already contain lysis buffer to reduce electric charging. The tissue can then be incubated according to extraction protocol and the lysate safely transferred into the 96-well plate, to reduce the risk of cross-contamination. A strong adhesive plate sealing tape (if necessary, detergent resistant) should be used throughout the entire workflow to prevent spilling of samples. Additionally, plates should always be centrifuged before opening and sealed with fresh sealing tape (ideally tightened with a plastic squeegee). To facilitate detection of cross-contamination each row and each column needs to contain an extraction blank that will be included in PCR and sequencing (Figure 1). A positive control can be included to verify the consistency between runs, and if sufficiently different from the target community of the study also used to detect tag switching. For studies with a high contamination risk, e.g. environmental DNA (eDNA), mock samples and even synthesised DNA (Wilson, Wozney, & Smith, 2015) can be could be used as positive control. However, generating and thoroughly characterising these can be costly and time consuming. As an alternative, tissue powder from a previous metabarcoding run can be used as positive control throughout the metabarcoding workflow. In order to minimise variability introduced by insufficient homogenisation it is recommended to thoroughly homogenise the positive control sample (e.g. with liquid nitrogen) to ensure homogeneity if used across several experiments (Elbrecht & Leese, 2015). Also, individual specimens that are
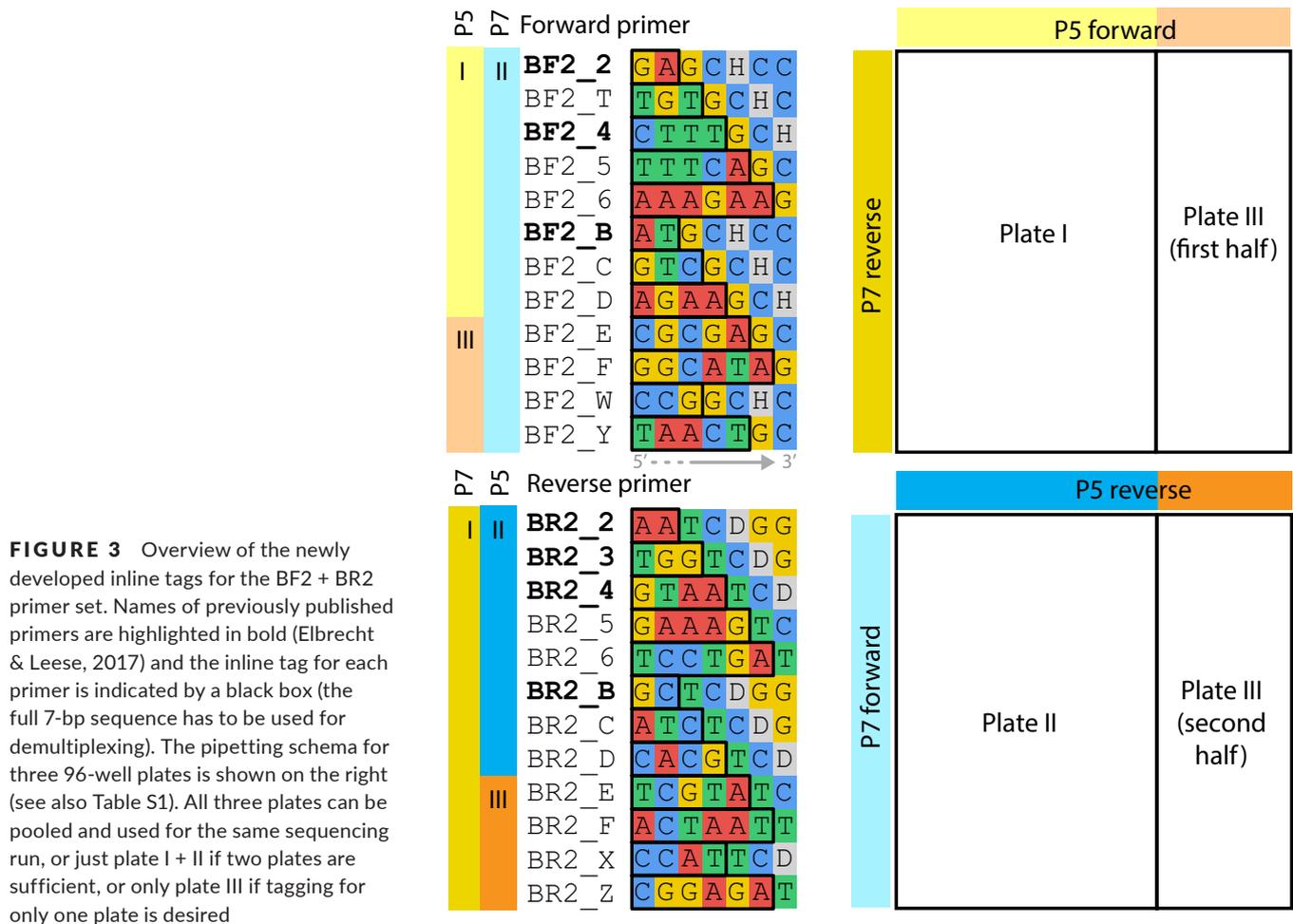
not expected to occur in the samples can be used to fill a few empty slots in the extraction plate, but these slots can also be left empty as additional negative controls. To increase PCR success and for easier troubleshooting DNA extracts could be normalised to identical concentrations (e.g. if samples are differently conserved). DNA can be quantified by fluorometric methods (Qubit, Life Technologies, USA) or by using a chromatogram-based approach (e.g. Fragment Analyzer, Advanced Analytical, USA), which measures DNA quality at the same time. However, in most cases, DNA normalisation might not be needed and DNA concentration can always be adjusted in response to low PCR success.

3 | AMPLIFICATION AND TAGGING: TWO-STEP PCR PROTOCOL

After the DNA is extracted, the barcode marker can be amplified. For freshwater macrozoobenthos, the cytochrome oxidase subunit I (COI) gene is usually used, but some authors also recommend ribosomal markers (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). We think ribosomal markers do not offer any advantages over well-designed degenerated COI primer sets (Elbrecht & Leese, 2017). Additionally, ribosomal markers often lack adequate reference data (Elbrecht et al., 2016). The use of highly degenerated primer sets is recommended, e.g. the BF2 + BR2 primer set, as it was specifically designed for freshwater macrozoobenthos and has already been evaluated using both mock and kick samples (Elbrecht & Leese, 2017; Elbrecht, Vamos, et al., 2017). Further PCR and primer modifications are dependent on the strategy used to multiplex several uniquely tagged samples for a sequencing run. We recommend the use of a two-step PCR protocol, in which the first PCR amplifies the target fragment utilising universal primers, while the second PCR uses fusion primer versions of the same primer sets, which include an inline tag and Illumina sequencing tails (Figure 2). Fusion primers can be used directly in a single PCR approach (Caporaso et al., 2012; Fadrosh et al., 2014; Kozich, Westcott, Baxter, Highlander, & Schloss, 2013), but a two-step PCR setup is less susceptible to PCR
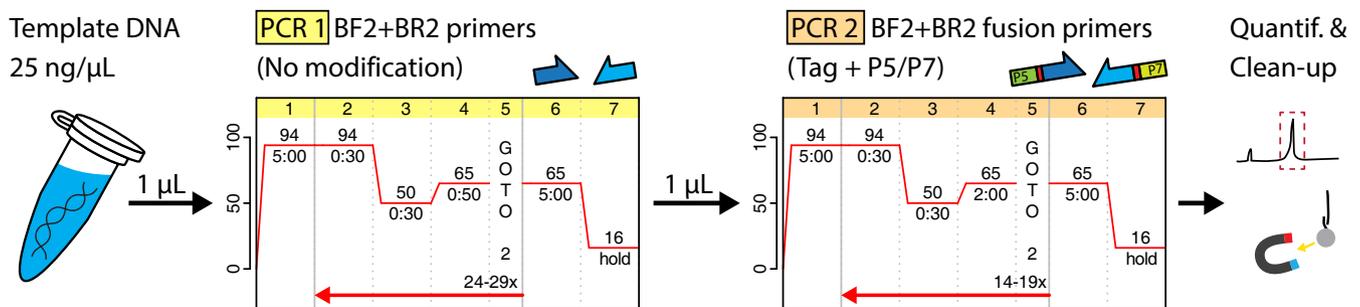
FIGURE 2 Overview of the two-step metabarcoding polymerase chain reaction (PCR) protocol (using HotMaster *Taq*, QuantaBio, U.S.A. in this example). The first PCR uses the standard BF2 + BR2 primers without modifications, thereby increasing amplification efficiency. Subsequently, 1 µl of amplicon product from the first PCR is used (without clean-up) as template for the second PCR step utilising fusion primers, which adds inline tags as well as Illumina sequencing adaptors. Note that the extension time is increased for the second PCR to ensure the entire fusion primer gets amplified. After the second PCR, the product can be prepared for sequencing (quantification, pooling with other amplicons and clean-up)

FIGURE 3 Overview of the newly developed inline tags for the BF2 + BR2 primer set. Names of previously published primers are highlighted in bold (Elbrecht & Leese, 2017) and the inline tag for each primer is indicated by a black box (the full 7-bp sequence has to be used for demultiplexing). The pipetting schema for three 96-well plates is shown on the right (see also Table S1). All three plates can be pooled and used for the same sequencing run, or just plate I + II if two plates are sufficient, or only plate III if tagging for only one plate is desired

inhibition (Schnell, Bohmann, & Gilbert, 2015). Additionally, fusion primers greatly reduce chances of tag switching (Elbrecht, Vamos, et al., 2017), especially when used in a one-step PCR. With a two-step PCR approach, cross-contaminations are still possible after the first PCR step, especially if good laboratory practice is not followed. Alternative approaches that add tagging to the first PCR step could be considered as well (Kitson et al., 2018). However, any tagging approach should be carefully chosen and experimentally evaluated. Some studies suggest that tag switching can become an issue with other more complex modular approaches (Esling, Lejzerowicz, & Pawlowski, 2015; Schnell et al., 2015).

Inline tags of different length and parallel sequencing in forward and reverse direction can substantially increase sequence diversity, which in turn leads to better results on Illumina machines and allows for a reduced spike-in of ~5% PhiX (Elbrecht & Leese, 2015; Wu et al., 2015). That said, fusion primers can be quite costly, as many versions with different in-line tags are needed. They also need to be developed and ordered for each new primer set (thus using commercial indexing kits for small projects might be more cost-effective). However, if the same fusion primer set is used more frequently, it can become highly cost-effective. A single primer costs around US$ 50 and yields over 100 μl with a 100 pmol/μl concentration, of which 25 pmol are used per 50 μl PCR reaction. Such a set

(forward + reverse primer) can be used to tag 400 samples at a cost of US$ 100 (US$ 0.25 per reaction).

Previously developed BF2 + BR2 fusion primer sets were limited to tagging a maximum of 72 samples (Elbrecht & Leese, 2017), which will quickly become insufficient for large-scale metabarcoding projects. Therefore, we developed new fusion primer sets that allow unique tagging and multiplexing of up to 288 samples on three 96-well microplates within the same run (Figure 3, see Supporting Information Figure S1 for full primer sequences and Supporting Information Table S1 for plate layouts). These new tags use a 7-bp sequence for both forward and reverse primers, while avoiding inline tags of 0–1 bp length, which are easily affected by insertions or deletions caused by sequencing errors (Faircloth & Glenn, 2012). Because the manual development of large numbers of different tags is difficult, we employed an R script that we used to randomly generate 100,000 tagging sets (Supporting Information Data S1). Seven previously developed primer pairs were incorporated into the design process, but the overall base composition was kept similar where possible (Supporting Information Figure S2). The similarity between tags of each generated set was subsequently visualised (Supporting Information Figure S3), and the primer set with most divergent tags was chosen in order to reduce potential tag switching through sequencing errors. All newly developed fusion primers had an inline

Sequencer	MiSeq			HiSeq (1 of 2 lanes)
Sequencing kit	250 PE v2 Nano[a]	250 PE v2	300 PE v3	250 PE v2 rapid run
Throughput (max)	1,000 k	15,000 k	25,000 k	150,000 k
Number of plates sequenced:				
One (96 wells)	10.4 k	156.2 k	260.4 k	1562.5 k
Two (192 wells)	5.2 k	78.1 k	130.2 k	781.3 k
Three (288 wells)	3.5 k	52.1 k	86.8 k	520.8 k

[a]Approximately two-thirds of the cost of the 250 PE v2 kit, too expensive.

tag of at least 3 bp, with the exception of four fusion primers that had only a 2-bp insert. We also calculated the Levenshtein distance utilising the R package stringdist v0.9.4.6 (Van der Loo 2014) to ensure single insertions or deletions (indels) will not lead to tag switching (Supporting Information Figure S4, Faircloth & Glenn, 2012). The Levenshtein distance was always 2 or higher, which should be sufficient given that Illumina sequencers are relatively unaffected by indels (Salipante et al. 2014). For PCR, we recommend using a reaction volume of 25–50 μl with a high-quality standard *Taq*. It is our experience that proof-reading *Taq* enzymes often struggle with primer degeneracy and long primer tails. For the first PCR (Figure 2), a master mix using the standard BF2 + BR2 primers is added to each 96-well plate. As the extracted DNA (including negative/positive controls) is already present in a 96-well format, ~25 ng DNA can be easily transferred to the PCR plate (Figure 1H). After the initial PCR, 1 μl of amplicon is used as template for the second PCR that individually tags each sample (Figure 1I). The number of cycles needed in each PCR might have to be optimised depending on how strongly samples are inhibited. While the cycle number should be kept as low as possible, studies on barcoding data show that a high number of cycles is not necessarily compromising data quality (Krehenwinkel et al., 2017; Vierna, Doña, Vizcaíno, Serrano, & Jovani, 2017). PCR success of the first and second PCR can be verified by electrophoresis; however, bands might only be visible after the second PCR depending on cycle number. PCR reactions that failed or showed only weak amplification should be excluded from sequencing or re-run with 5–15 additional PCR cycles (e.g. if the band on the gel was only barely visible).

4 | LIBRARY PREPARATION AND POOLING

Amplicons of the second PCR can be directly used for sequencing after chromatographic quantification (Figure 1J) and clean-up (to remove residual primers and other PCR components). As long as it is possible to measure the concentration of amplicons independently from primer dimers, samples can be pooled first and then subjected to clean-up. Otherwise, each individual sample will need to be cleaned separately before quantification. Usually, all samples are pooled with identical amplicon concentration to ensure similar sequencing depth across all of them. An alternative to

manual quantification and pooling could be quantitative DNA binding (SequalPrep kit; Invitrogen) to normalise DNA concentrations (Harris et al., 2010). In some cases, sample concentrations should be adjusted, e.g. if amplicons of different length are sequenced on the same run (Elbrecht & Leese, 2017) or if the number of specimens across samples is highly variable (Beermann, Zizka, Elbrecht, Baranov, & Leese, 2018; Theissinger et al., 2018). It should also be stressed that both the quantification and pooling steps are absolutely essential for the desired sequencing depth across samples, and the accuracy of any used quantification method should be verified prior to any experiments (Elbrecht, Vamos, et al., 2017). As negative controls are difficult to quantify due to low concentration, any adjustment to the concentration of other samples would lead to a strong over-representation. We therefore recommend adding each negative control to the library in volumes equal to the average volume of the samples used for pooling.

An effective solution for clean-up is magnetic bead purification as it also allows for removal of amplicons that do not match the targeted marker length (Figure 1L). Usually, a left-sided size selection is sufficient as long as no strong double bands are present. Alternative clean-up methods (e.g. spin column based) will be needed if bovine serum albumin was included as a PCR enhancer, as it can prevent re-suspension of magnetic beads (Elbrecht, Peinert, et al., 2017). The clean library can then be directly loaded onto an Illumina sequencer. As only inline barcodes are used for sample tagging, both Illumina indexing read steps can be skipped (Figure 1M). Following sequencing, reads are demultiplexed using the first 7 bp of read one and two (e.g. implemented in the R package "JAMP", http://github.com/VascoElbrecht/JAMP).

5 | SEQUENCING DEPTH

The number of samples (or plates) that can be sequenced on the same run depends on the number of sequences a platform produces as well as on the desired sequencing depth for each sample. The lower the sequencing depth, the more taxa will remain undetected, especially those with low abundance, low biomass, and those strongly affected by primer bias (Alberdi, Aizpurua, Gilbert, & Bohmann, 2017; Elbrecht, Peinert, et al., 2017). For macrozoobenthos bulk samples, we recommend a sequencing depth of at least

100,000 sequences per replicate but this can be more or less dependent on sample biomass. As the BF2 + BR2 primer set amplifies a 421 bp region, paired-end sequencing with at least 250 bp sequence length is necessary. Table 1 shows an overview of currently available Illumina sequencers that meet these criteria (end of 2018) and the expected sequencing depth they can produce per well. A library can be easily re-sequenced when sequencing depth turns out to be insufficient. Additionally, sequencing depth between samples might vary depending on quantification accuracy for individual samples. Samples with insufficient sequencing depth can be recovered, e.g. by adding additional PCR product to the affected samples in a library for a re-run (alternatively respective samples can be excluded from the dataset if only a few are affected).

6 | BIOINFORMATICS PROCESSING AND TROUBLESHOOTING

The choice of bioinformatics pipelines and clustering settings can drastically affect the resulting taxon list, especially when it comes to rare taxa (Figure 1N) (Kopylova et al., 2016). However, as long as data is strictly filtered (removal of singletons, abundance-based filtering of operational taxonomic units [OTUs]) and an appropriate OTU-clustering algorithm is used for the pool of all samples, results should be reliable (see e.g. Elbrecht, Vamos, et al., 2017). While it is out of the scope of this manuscript to describe the bioinformatic process in detail, we have highlighted a few key points related to our laboratory setup. These should be considered for analysis of metabarcoding data generated with the laboratory methods proposed. Only samples with sufficient sequencing depth should be used in such analysis, and if samples vary strongly in sequencing depth, rarefaction should be applied across all samples to ensure equal sequencing depth. If a single replicate is of insufficient sequencing depth, the entire sample should be removed from the dataset. Both replicates for each sample should be very similar in OTU composition, which can be verified by calculating Bray–Curtis (accounts for abundance) and Jaccard dissimilarities (presence–absence). Any discrepancies could indicate problems, e.g. caused by insufficient tissue homogenisation, cross-contamination or PCR and sequencing errors (Lange et al., 2015; Zepeda- Mendoza, Bohmann, Baez, & Gilbert, 2016). Any reads that are only present in one replicate should be removed to reduce influence of such errors, or the complete sample should be discarded and re-run if both replicates are too different (Figure 1O). However, these samples and OTUs should still be included and highlighted when reporting the raw data, ideally in form of an OTU table. Strong cross-contamination and tag switching can also be detected by discrepancies between the replicates and reads present in the 12 negative or positive controls, especially if the contamination is patchy and not systematic (Kelly, Pereira-Maxwell, Carnaby, & White, 2005). Some tag switching might be observed, but usually only at a very low abundance and therefore it is not of concern (Elbrecht, Vamos, et al., 2017). We suggest subtracting the maximum read count of each OTU in the 12 negative controls from all other samples in order to reduce the effects of low abundance tag switching on the data set. Depending on the accuracy needed, read counts in the negative controls can be multiplied by 2 or more before subtraction. However, if severe tag switching or cross-contamination is detected, the entire metabarcoding run might have to be repeated (starting from the DNA extraction stage if the source of contamination is unknown). Positive controls can also be used to confirm protocol consistency between plates and sequencing runs. While there are different ways of using positive and negative controls to reduce noise or to validate a metabarcoding run, the exact strategy depends on the accuracy needed. Contaminations or deviating positive controls might indicate issues and can provide pointers to the source of the problem. Potential issues should be reported in the respective study to ensure transparency. The decision to exclude samples or taxa or to repeat runs depends on the scope of a study and the intention of the researchers. For instance, the detection of invasive or rare species utilising eDNA requires very strict and conservative processing to reduce detections of false positives (Bista et al., 2017; Yamamoto et al., 2016), whereas studies investigating general patterns of biodiversity might be less affected by noise in the dataset. It is often useful to discard OTUs and reads with low relative abundance (Bista et al., 2017; Elbrecht, Vamos, et al., 2017; Yamamoto et al., 2016), as these can be heavily affected by stochastic effects (Leray & Knowlton, 2017). The minimum abundance used can, however, vary based on sample type and study goal (0.003% for Elbrecht, Peinert, et al., 2017; Elbrecht, Vamos, et al., 2017 and 1.5% for Yamamoto et al., 2016).

7 | DISCUSSION AND CONCLUSIONS

Our proposed metabarcoding strategy is based on sufficiently validated the laboratory processes using controls, while still keeping the workflow simple and scalable. By working with 96-well microplates high sample throughput can be easily achieved while at the same time reducing the risk of cross-contamination. By running two replicates starting at the DNA extraction stage, together with negative and positive controls, we ensure that errors are still detectable despite the reduced need to validate each individual laboratory step. The BF2 + BR2 fusion primer sets which are extended here (Supporting Information Figure S1) are well tested for macroinvertebrate communities (Elbrecht & Leese, 2017; Elbrecht, Vamos, et al., 2017), enabling the tagging and sequencing of up to 288 wells in a single sequencing run. We are confident that this metabarcoding workflow will produce reliable results for up to 123 replicated samples per sequencing run (Figure 1) utilising a simplified fusion primer-based sample tagging process.

The number of samples that can be multiplexed with our tagging system is optimised for the currently available Illumina platforms. However, the throughput of sequencers continues to increase with new sequencers and kits being introduced frequently. Already a shorter COI fragment could be used to amplify DNA from macrozoobenthos bulk samples (Meusnier et al., 2008),

which would allow for sequencing at ~50× increased throughput (e.g. HiSeq versus NovaSeq). Such an approach would require thousands of samples being uniquely tagged and multiplexed for a single sequencing run. Although our inline tags are only able to tag 288 wells, they could be extended to several thousand tagging combinations by incorporating Illumina indexing into the fusion primers.

While we are convinced that our metabarcoding approach is efficient and reliable it needs to be validated in practice and thoroughly compared to other protocols. We hope that this manuscript will encourage discussion and help to find better approaches for the scale-up of metabarcoding for biodiversity assessment. Variations of our proposed workflow as well as comparisons to alternative metabarcoding protocols are explicitly encouraged. This could be done by sequencing macroinvertebrate samples of known composition (mock samples, specimens identified based on morphology or DNA barcoding (Hebert et al. 2003)). Also, a sequencing run with tissue from individually tagged specimens could be used to investigate the robustness of our metabarcoding approach to cross-contamination. We also think the presented approach in this study could be extended to other groups and ecosystems, such as terrestrial arthropods, algae or eDNA studies.

ACKNOWLEDGMENTS

AUTHORS' CONTRIBUTIONS

V.E. developed the laboratory workflow, fusion primer set and wrote the manuscript. D.S. revised the manuscript.

REFERENCES

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2017). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, *9*, 134–147.

Andújar, C., Arribas, P., Gray, C., Bruce, K., Woodward, G., Yu, D. W., & Vogler, A. P. (2018). Metabarcoding of freshwater invertebrates to detect the effects of a pesticide spill. *Molecular Ecology*, *27*(1), 146–166.

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

Beermann, A. J., Zizka, V. M. A., Elbrecht, V., Baranov, V., & Leese, F. (2018). DNA metabarcoding reveals the complex and hidden responses of chironomids to multiple stressors. *Environmental Sciences Europe*, *30*, 1–15.

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., … Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, *8*, 1–11.

Braukmann, T. W., Ivanova, N. V., Prosser, S. W., Elbrecht, V., Steinke, D., Ratnasingham, S., et al. (2018). Revealing the Complexities of Metabarcoding with a Diverse Arthropod Mock Community. *bioRxiv*, 1–40. https://doi.org/10.1101/433607

Buss, D. F., Carlisle, D. M., Chon, T. S., Culp, J., Harding, J. S., Keizer-Vlek, H. E., … Hughes, R. M. (2015). Stream biomonitoring using macroinvertebrates around the globe: A comparison of large-scale programs. *Environmental Monitoring and Assessment*, *187*, 4132.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., … Knight, R. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal*, *6*, 1621–1624.

Carew, M. E., Nichols, S. J., Batovska, J., St Clair, R., Murphy, N. P., Blacket, M. J., & Shackleton, M. E. (2017). A DNA barcode database of Australia's freshwater macroinvertebrate fauna. *Marine and Freshwater Research*, *68*, 1788.

Carew, M. E., Pettigrove, V. J., Metzeling, L., & Hoffmann, A. A. (2013). Environmental monitoring using next generation sequencing: Rapid identification of macroinvertebrate bioindicator species. *Frontiers in zoology*, *10*, 45.

Curry, C. J., Gibson, J. F., Shokralla, S., Hajibabaei, M., & Baird, D. J. (2018). Identifying North American freshwater invertebrates using DNA barcodes: Are existing COI sequence libraries fit for purpose? *Freshwater Science*, *37*, 178–189.

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters*, *10*, 1–4.

Dudgeon, D., Arthrington, A. H., Gessner, M. O., Kawabata, Z.-I., Knowler, D. J., Lévêque, C., … Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation challenges. *Biological Reviews*, *81*, 163.

Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—Sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, *10*, e0130324.

Elbrecht, V., & Leese, F. (2017). Validation and development of freshwater invertebrate metabarcoding COI primers for environmental impact assessment. *Frontiers in Environmental Science*, *5*, 1–11.

Elbrecht, V., Peinert, B., & Leese, F. (2017). Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution*, *7*, 6918–6926.

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.-N., … Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, *4*, e1966.

Elbrecht, V., Vamos, E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, *8*, 1265–1275.

Emilson, C. E., Thompson, D. G., Venier, L. A., Porter, T. M., Swystun, T., Chartrand, D., … Hajibabaei, M. (2018). DNA metabarcoding and morphological macroinvertebrate metrics reveal the same changes in boreal watersheds across an environmental gradient. *Scientific Reports*, *8*(1), 4578. https://doi.org/10.1038/s41598-018-22978-3

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

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for

multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, 2, 6.

Faircloth, B. C., & Glenn, T. C. (2012). Not all sequence tags are created equal: Designing and validating sequence identification tags robust to indels. *PLoS ONE*, 7, e42543.

Gibson, J. F., Shokralla, S., Curry, C., Baird, D. J., Monk, W. A., King, I., & Hajibabaei, M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. *PLoS ONE*, 10, e0138432.

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

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, 6, 1–7.

Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konynenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology*, 12, 28.

Harris, J. K., Sahl, J. W., Castoe, T. A., Wagner, B. D., Pollock, D. D., & Spear, J. R. (2010). Comparison of normalization methods for construction of large, multiplex amplicon pools for next-generation sequencing. *Applied and Environmental Microbiology*, 76, 3863–3868.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Kelly, P., Pereira-Maxwell, F., Carnaby, S., & White, I. (2005). Confidence in polymerase chain reaction diagnosis can be improved by Bayesian estimation of post-test disease probability. *Journal of Clinical Epidemiology*, 58, 252–260.

Kitson, J. J. N., Hahn, C., Sands, R. J., Straw, N. A., Evans, D. M., & Lunt, D. H. (2018). Detecting host-parasitoid interactions in an invasive Lepidopteran using nested tagging DNA metabarcoding. *Molecular Ecology*, 2, e197.

Kopylova, E., Navas- Molina, J. A., Mercier, C., Xu, Z. Z., Mahé, F., He, Y., ... Knight, R. (2016). Open-source sequence clustering methods improve the state of the art. *mSystems*, 1, 1–16.

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, 79, 5112–5120.

Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7(1), 17668.

Lange, A., Jost, S., Heider, D., Bock, C., Budeus, B., Schilling, E., ... Hoffmann, D. (2015). AmpliconDuo: A split-sample filtering protocol for high-throughput amplicon sequencing of microbial communities. *PLoS ONE*, 10, e0141590.

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, A., Bruce, K., ... Weigand, A. M. (2018). Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: A perspective derived from the DNAqua-net cost action. In D. Bohan, A. Dumbrell, G. Woodward & M. Jackson (Eds.), *Next generation biomonitoring: Part 1. Advances in ecological research* (pp. 63–99). Oxford, UK: Elsevier.

Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, 5, e3006–e3027.

Meusnier, I., Singer, G. A., Landry, J. F., Hickey, D. A., Hebert, P. D., & Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, 9, 214.

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

Salipante, S. J., Kawashima, T., Rosenthal, C., Hoogestraat, D. R., Cummings, L. A., Sengupta, D. J., et al. (2014). Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Applied and Environmental Microbiology*, 80, 7583–7591. https://doi.org/10.1128/AEM.02206-14

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

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

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

Theissinger, K., Kästel, A., Elbrecht, V., Makkonen, J., Michiels, S., Schmidt, S., ... Brühl, C. (2018). Using DNA metabarcoding for assessing chironomid diversity and community change in mosquito controlled temporary wetlands. *Metabarcoding and Metagenomics*, 2, e21060.

Van der Loo, M. (2014). The Stringdist Package for Approximate String Matching. *The R Journal*, 6, 111–122.

Vierna, J., Doña, J., Vizcaíno, A., Serrano, D., & Jovani, R. (2017). PCR cycles above routine numbers do not compromise high-throughput DNA barcoding results. *Genome*, 60, 868–873.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., ... Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature*, 467, 555–561.

Wilson, C. C., Wozney, K. M., & Smith, C. M. (2015). Recognizing false positives: Synthetic oligonucleotide controls for environmental DNA surveillance. *Methods in Ecology and Evolution*, 7, 23–29.

Wu, L., Wen, C., Qin, Y., Yin, H., Tu, Q., Van Nostrand, J. D., ... Zhou, J. (2015). Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. *BMC Microbiology*, 15, 125.

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., ... Miya, M. (2016). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, 7, 40368.

Zepeda-Mendoza, M. L., Bohmann, K., Baez, A. C., & Gilbert, M. T. P. (2016). DAMe: A toolkit for the initial processing of datasets with PCR replicates of double-tagged amplicons for DNA metabarcoding analyses. *BMC Research Notes*, 9, 255.

SUPPORTING INFORMATION

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