Running Title (45 char max): Metabarcoding with fusion primer sets

Title: Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring

Authors: Vasco Elbrecht¹*, Dirk Steinke¹,²

Affiliations:
1) Centre for Biodiversity Genomics, University of Guelph, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada
2) Department of Integrative Biology, University of Guelph, 50 Stone Road East, Guelph, Ontario, N1G 2W1, Canada

*Corresponding author: Vasco Elbrecht (elbrecht@uoguelph.ca),

Abstract:
The viability of DNA metabarcoding for assessment of freshwater macrozoobenthos has been demonstrated over the past years. It matured to a stage where it can be applied to monitoring at a large scale, keeping pace with increased high throughput sequencing (HTS) capacity. However, workflows and sample tagging need to be optimized to accommodate for hundreds of samples within a single sequencing run. We here conceptualize 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. We hope that our metabarcoding workflow will be used as a practical guide for future large-scale biodiversity assessments involving freshwater invertebrates. However, we also want to point out that this is just one possible metabarcoding approach, and that we hope this article will stimulate discussion and publication of alternatives and extensions.

Key words: Biomonitoring, High throughput sequencing, Macrozoobenthos, Multiplexing, Fusion Primer, Metabarcoding workflow, Replication
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 are promising alternatives (Baird & Hajibabaei 2012) to morphological identification, which is often limited in resolution and dependent on taxonomic experience (Sweeney et al. 2011). In addition to reducing human bias, DNA based identifications can also lead to improved stream assessment (Stein et al. 2013). Over the past few years several studies demonstrated the feasibility of metabarcoding-based monitoring of freshwater macroinvertebrates (Hajibabaei et al. 2011; Carew et al. 2013; Gibson et al. 2015; Andújar et al. 2017). Despite some methodological limitations (presence/absence data, primer bias (Piñol et al. 2014; Elbrecht & Leese 2015)), assessment results are at least comparable if not superior to conventional morphology-based stream monitoring approaches (Gibson et al. 2015; Elbrecht et al. 2017b; Emilson et al. 2017). Some macroinvertebrate reference databases are already fairly comprehensive especially for common taxa (Carew et al. 2017; Curry et al. 2018). Furthermore, many aspects of the metabarcoding approach have been thoroughly validated (Hajibabaei et al. 2012; Carew et al. 2013; Elbrecht & Leese 2015; Gibson et al. 2015; Elbrecht & Leese 2017; Elbrecht, Peinert & Leese 2017a; Emilson et al. 2017; Andújar et al. 2017). 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 (Elbrecht et al. 2017b; Emilson et al. 2017; Andújar et al. 2017). 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 (Fig. 1), using the BF2+BR2 fusion primer system (Elbrecht & Leese 2017) which has been shown to work well with macroinvertebrate monitoring samples (Elbrecht et al. 2017b) but is 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, positive and negative controls already at the tissue homogenization stage, samples affected by laboratory issues can be reliably detected and if necessary excluded from subsequent analysis. Furthermore, by minimizing 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 utilizing controls and replicates in a manner that highlights samples affected by errors.

Sample collection, homogenization and DNA extraction

After samples are collected using a standardized protocol (Fig. 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 (Fig. 1B). While this increases the chance that some taxa and specimens will be overlooked (Haase et al. 2010), in most metabarcoding studies of freshwater macroinvertebrates specimens were separated from debris (Carew et al. 2013; Gibson et al. 2015; Elbrecht et al. 2017b; Emilson et al. 2017). This is often done as part of preceding morphological identifications, or out of the concern that homogenizing an entire sample might introduce PCR inhibitors and complicate standardization.

Although work intensive methods such as specimen flotation are currently being explored (Andújar et al. 2017), there is not enough evidence yet to decide if homogenization of full kick samples without separating invertebrate specimens from collected substrate is also feasible. Once specimens are separated from debris they can be dried (Fig. 1C) and homogenized (Fig. 1D). Alternatively, DNA of bulk samples has also been extracted directly from the preservation ethanol (Hajibabaei et al. 2012), through homogenization of the wet sample (Hajibabaei et al. 2011), or by lysing the complete sample (Braukmann et al. submitted). To ensure complete homogenization, it is recommended to grind dried bulk samples 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 et al. 2017b)). Two replicates per sample should be used for DNA extraction (Fig. 1E), both of which are metabarcoded to facilitate the detection of insufficient tissue homogenization. If homogenization 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, Fig. 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 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 (Fig. 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
detect tag switching. For more studies with a high contamination risk e.g. environmental DNA (eDNA), mock samples and even synthesized DNA (Wilson, Wozney & Smith 2015) can be could be used as positive control. However, generating and thoroughly characterizing 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 minimize variability introduced by insufficient homogenization it is recommended to thoroughly homogenize the positive control sample (e.g. with liquid nitrogen) to ensure it's 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 normalized 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 normalization might not be needed and DNA concentration can always be adjusted in response to low PCR success.

**Amplification and tagging: Two step PCR protocol**

After the DNA is extracted and normalized, the barcode marker can be amplified. For freshwater macrozoobenthos, the cytochrome oxidase subunit I (COI) gene is usually used, but some authors also recommended ribosomal markers (Deagle et al. 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 et al. 2017b; Elbrecht & Leese 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 utilizing universal primers, while the second PCR uses fusion primer versions of the same primer sets, which include an inline tag and Illumina sequencing tails (Fig. 2). Fusion primers can be used directly in a single PCR approach (Caporaso et al. 2012; Kozich et al. 2013; Fadrosh et al. 2014), but a two-step PCR setup is less susceptible to PCR inhibition (Schnell, Bohmann & Gilbert 2015). Additionally, fusion primers greatly reduce chances of tag switching (Elbrecht et al. 2017b), 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 (Wu et al. 2015; Elbrecht & Leese 2015). That being 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 $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 $100 ($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 (Fig. 3, see Fig S1 for full primer sequences and Tab 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 (Script S1).

Seven previously developed primer pairs were incorporated into the design process, but the overall base composition was kept similar where possible (Fig S2). The similarity between tags of each generated set was subsequently visualized (Fig S3), and the primer set with most divergent tags was chosen in order to reduce potential tag switching through sequencing errors. Tags in the selected set differed by at least 3 bp, with the exception of four fusion primers that had only a 2 bp insert.

We also calculated the Levenshtein distance utilizing the R package stringdist v0.9.4.6 (Van der Loo 2014) to ensure single insertions or deletions (indels) won’t lead to tag switching (Figure S4, Faircloth and 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 50 μl with a high quality standard Taq. It is our experience that proof reading Taq’s often struggle with degeneracy and long primer tails. For the first PCR (Fig. 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 (Fig. 1H). After the initial PCR 1 μl amplicon is used as template for the second PCR that individually tags each sample (Fig. 1I). The number of cycles needed in each PCR might have to be optimised depending on how strongly samples are inhibited. While the cycle
The 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 (Vierna et al. 2017; Krehenwinkel et al. 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 10-15 additional PCR cycles (e.g. if the band on the gel was only barely visible).

**Library Preparation and Pooling**

Amplicons of the second PCR can be directly used for sequencing after chromatographic quantification (Fig. 1J) and cleanup (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 cleanup. 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 (SequlPrep kit; Invitrogen) to normalize DNA concentrations (Harris et al. 2010). In some cases sample concentrations can 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 (Theissinger et al. 2018; Beermann et al. 2018). It should also be stressed that both the quantification and pooling step 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 et al. 2017b). As negative controls are difficult to quantify due to low concentration any adjustment to the concentration of other samples would lead to a strong overrepresentation. 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 cleanup is magnetic bead purification as it also allows for removal of amplicons that do not match the targeted marker length (Fig. 1L). Usually a left-sided size selection is sufficient as long as no strong double bands are present. Alternative cleanup methods (e.g. spin column based) will be needed if BSA was included as a PCR enhancer, as it can prevent re-suspension of magnetic beads (Elbrecht et al. 2017a). 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 (Fig. 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).

**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 et al. 2017; Elbrecht et al. 2017a). For macrozoobenthos bulk samples we recommend a sequencing depth of at least 100,000 sequences per replicate, but 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 2017) 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).

**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 (Fig. 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 et al. 2017b)). While it is out of the scope of this manuscript to describe the bioinformatic process in detail, we 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 homogenization, cross-contamination or PCR and sequencing errors (Lange et al. 2015; Zepeda-Mendoza et al. 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 (Fig. 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 et al. 2005). Some tag switching might be observed, but usually only at a very low abundance and therefore it is not of concern (Elbrecht et al. 2017b). We suggest to subtract 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 and higher 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 utilizing eDNA requires very strict and conservative processing to reduce detections of false positives (Yamamoto et al. 2016; Bista et al. 2017), 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 (Yamamoto et al. 2016; Bista et al. 2017; Elbrecht et al. 2017b), 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 et al. 2017 and 1.5% for Yamamoto et al. 2017).

**Discussion and conclusions**

Our proposed metabarcoding strategy is based on sufficiently validated laboratory methods, 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 (Fig S1) are well tested for macroinvertebrate communities (Elbrecht et al. 2017b; Elbrecht & Leese 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 (Fig. 1) utilizing a simplified fusion primer based sample tagging process.
The number of samples that can be multiplexed with our tagging system is optimized for the currently available Illumina platforms. However, the throughput of sequencers continues to increase with new sequencers and kits being introduced frequently. Already today a shorter COI fragment could be used to amplify DNA from macrozoobenthos bulk samples (Meusnier et al. 2008), which would allow for sequencing at ~50x increased throughput (e.g. HiSeq vs. 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 helps 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, samples identified based on morphology). Also a sequencing run with tissue from individual unique 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, like e.g. terrestrial arthropods, algae or eDNA studies.
Acknowledgements

We would like to thank Thomas Braukmann, Sean Prosser, Natalia Ivanova, Edith Vamos, Nina Röder, Arne Beermann and Matthias Geiger for helpful discussions and feedback, which improved earlier versions of this manuscript. Two reviewers provided helpful comments that greatly improved the quality of this manuscript. V.E. and D.S. are supported by funding through the Canada First Research Excellence Fund. This work represents a contribution to the ‘Food From Thought’ research program and the EU Action ‘DNAqua-Net’ (CA15219).

Author contributions

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

Supporting information

Fig. S1: Newly developed fusion primer sets (BF2+BR2), suitable for tagging 288 individual wells.

Fig. S2: Base composition of the inline tagging region.

Fig. S3: Hamming distance between tags for all fusion primers.

Fig. S4: Levenshstein distance between tags for all fusion primers.

Script S1: R script used to randomly generate inline barcodes for the given primer sets (includes visualization, as shown in Fig. S3 and Fig. 3).

Table S1: Table providing an overview of proposed tagging combinations (as shown in Fig. 3).

Literature Cited


Table 1: Sequencing depth per well with different Illumina sequencing platforms suitable for the BF2+BR2 fusion primers (k = 1,000 spots).

<table>
<thead>
<tr>
<th>Sequencer</th>
<th>MiSeq</th>
<th>HiSeq (1 of 2 lanes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing Kit</td>
<td>250 PE v2 Nano*</td>
<td>250 PE v2</td>
</tr>
<tr>
<td>Throughput (max)</td>
<td>1,000k</td>
<td>15,000k</td>
</tr>
<tr>
<td>Number of plates sequenced:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One (96 wells)</td>
<td>10.4k</td>
<td>156.2k</td>
</tr>
<tr>
<td>Two (192 wells)</td>
<td>5.2k</td>
<td>78.1k</td>
</tr>
<tr>
<td>Three (288 wells)</td>
<td>3.5k</td>
<td>52.1k</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>* ~2/3 the cost of the 250 PE v2 kit, too expensive.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Overview of the proposed metabarcoding workflow 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 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.
Figure 2: Overview of the two-step metabarcoding PCR protocol (using HotMaster Tag, QuantaBio, USA 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 cleanup) as template for the second PCR step utilizing fusion primers, which adds inline tags as well as Illumina sequencing adaptors. Note that the extension time is increased for the second PCR in order 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. 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.