**Development and** validation of DNA metabarcoding COI primers for aquatic invertebrates using the R package "PrimerMiner**"**

Vasco Elbrecht 1, 2\*, Florian Leese 1,3

Affiliations:

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

2) Department of Animal Ecology, Evolution and Biodiversity, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany

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

\*Corresponding author: vasco.elbrecht@uni-due.de, phone: +49.201-1834053

**Abstract**

1) DNA metabarcoding is a powerful tool to assess biodiversity by amplifying and sequencing a standardized gene marker region. However, typically used barcoding genes, such as the cytochrome c oxidase subunit I (COI) region for animals, are highly variable. Thus, different taxa in communities under study are often not amplified equally well and some might even remain undetected due to primer bias. To reduce these problems, optimized region- and/or ecosystem- specific metabarcoding primers are necessary.

2) We developed the R package PrimerMiner, which batch downloads DNA barcode gene sequences from BOLD and NCBI databases for specified target taxa and then applies sequence clustering to reduce biases introduced by differed number of available sequences per species. To design primers targeted for freshwater invertebrates, we downloaded COI data for the 15 most important invertebrate groups relevant for stream ecosystem assessment. Four primer sets with high base degeneracy were developed and their performance tested by sequencing ten mock community samples consisting each of 52 freshwater invertebrate taxa. Additionally, we evaluated the developed primers against other metabarcoding primers *in silico* using PrimerMiner.

3) Amplification and sequencing was successful for all ten mock community samples with the four different primer combinations. The developed primers varied in amplification efficiency and amount of taxa detected, but all primer sets detected more taxa than standard Folmer barcoding primers. Additionally, the BF / BR primers amplified taxa very consistently, the BF2+BR2 and BF2+BR1 primer combination even better than a previously tested ribosomal marker (16S). Except for the BF1+BR1 primer combination, all BF / BR primers detected all 42 insect taxa present in the mock samples. *In silico* evaluation of the developed primers showed that they are also likely to work very well on other non aquatic invertebrate samples.

4) With PrimerMiner, we here provide a useful tool to obtain relevant sequence data for targeted primer development and evaluation. Our sequence datasets generated with the newly developed metabarcoding primers demonstrate that the design of optimized primers with high base degeneracy is superior to classical markers and enable us to detect almost 100% of animal taxa present in a sample using the standard COI barcoding gene. Therefore, the PrimerMiner package and primers developed using this tool are useful beyond assessment of biodiversity in aquatic ecosystems.

**Key words:** Primer development, DNA metabarcoding, primer bias, ecosystem assessment, in silico PCR, data mining1) Introduction

DNA barcoding allows for the reliable identification of collected specimens without prior knowledge of species taxonomy. A prerequisite is the availability of reliable reference databases. For animals, the usefulness of the cytochrome c oxidase I (COI) gene for species identification has been widely demonstrated and extensive reference databases exist for many taxonomic groups (Larsen *et al.* 2011; Ratnasingham & Hebert 2013). However, identifying single specimens using DNA barcoding is still quite expensive, as DNA has to be extracted individually, the barcoding region amplified and then typically Sanger sequenced (Cameron *et al.* 2006; Stein *et al.* 2014). Recent advances in high throughput sequencing (HTS) make it now possible to extract DNA and sequence the barcoding region from bulk environmental samples often containing hundreds to thousands of specimens. This technique, coined DNA metabarcoding, has been already widely used to generate comprehensive taxa list for a wide range of ecosystems and environments (Taberlet *et al.* 2012). One main challenge of DNA metabarcoding is the often severe primer bias that prevents detection of 100% of the taxa present in a sample and limits quantification of biomass from read abundances (Piñol *et al.* 2014; Elbrecht & Leese 2015).

A universal barcoding primer pair, amplifying a fragment of suitable length for HTS is thus the most critical component when assessing environmental samples with metabarcoding. As a coding gene, the COI barcoding region shows high codon degeneracy throughout its sequence, making design of "truly" universal primers difficult (Deagle *et al.* 2014; Sharma & Kobayashi 2014). Several universal COI barcoding primers have been developed of which many are now used or could be used in metabarcoding studies (Figure 1, Folmer *et al.* 1994; Hebert *et al.* 2004; Meusnier *et al.* 2008; Van Houdt *et al.* 2010; Zeale *et al.* 2010; Shokralla *et al.* 2011; Leray *et al.* 2013; Geller *et al.* 2013; Gibson *et al.* 2014; Shokralla *et al.* 2015; Brandon-Mong *et al.* 2015). Despite being universal and often including several degenerate bases, these metabarcoding primers typically recover only 80-90% or even less of the taxa present in a sample (Leray *et al.* 2013; Elbrecht & Leese 2015; Brandon-Mong *et al.* 2015). Furthermore, many primers used in metabarcoding have never been thoroughly evaluated with respects to primer bias and the proportion of undetected taxa. Thus, the development and critical evaluation of universal primers is still a pressing issue.

Details on primer design and/or used sequence data are often not described extensively (e.g. (Hajibabaei *et al.* 2011; Shokralla *et al.* 2015)). Typically many reference barcode sequences for the taxonomic target groups are taken from NCBI or BOLD and aligned (Zeale *et al.* 2010; Leray *et al.* 2013; Gibson *et al.* 2014) or alternatively only mitochondrial genomes or a small subset of barcoding sequences are used (Geller *et al.* 2013; Deagle *et al.* 2014; Brandon-Mong *et al.* 2015). A key problem when downloading complete datasets is that some taxa are typically overrepresented with hundreds of sequences deposited (e.g. because many sequences are available from detailed phylogeographic studies). This can in principle be circumvented when using only mitochondrial genomes. However, typically such data sets are very limited as mitochondrial genomic sequences are still rare for many taxonomic groups. However, obtaining good quality reference data is essential in manual and software based primer development. While there are many programs available to aid primer development (e.g. Primer3, Untergasser *et al.* 2012, EcoPCR, Ficetola *et al.* 2010), the challenge of batch downloading and systematically preparing obtained sequence data for primer development has not been tackled until now. Therefore, we have developed the R package PrimerMiner. The software allows the user to selectively batch download and cluster sequences into Operational Taxonomic Units (OTU). Clustering is independent to reported taxonomy and reduces biases introduced by misidentified taxa, database redundancies and overrepresented taxa. PrimerMiner includes visualisation tools to manually search for suitable metabarcoding primers. Further new *in silico* primer evaluation tools are introduced with PrimerMiner, which take type and position of mismatches between primer and template into account.

To test the PrimerMiner approach, we designed four DNA metabarcoding primer sets, targeting 15 freshwater invertebrate taxa of central importance in bioassessment programs. All primer sets were evaluated using ten mock communities with 52 taxa each, which have been used for primer evaluation in previous studies (Elbrecht & Leese 2015; Elbrecht et al. 2016). Additionally, the developed primers were evaluated *in silico* against commonly used DNA barcoding and metabarcoding primers.

# 2) Material and Methods

## The PrimerMiner R package

PrimerMiner is a fully automated R based sequence downloader and processor that condenses sequence data from NCBI and BOLD into Operational Taxonomic Units (OTUs) (Figure 2). It can download sequence data for a selected gene marker and specified taxonomic groups and also extract the respective target gene sequence from mitochondrial genomes if available. Thus, PrimerMiner takes full advantage of available partial sequences and mitochondrial genomes, laying a good data basis for primer development. All sequences are then clustered with a custom threshold (default 3%) and the OTUs for each taxonomic group exported as a fasta file for subsequent alignments and automated primer design with specific software or manually. The clustering strategy adopted in PrimerMiner has several key advantages: 1) Overrepresented taxa and duplicated sequences are merged into few OTUs. 2) Taxonomic variation within species is retained (wobble bases) while rare haplotypes can be ignored when generating OTU consensus sequences. 3) Highly diverse species are automatically represented by two or more OTUs. 4) Clustering is taxonomy-independent and thus can deal with misidentified species as long as their order / family was identified correctly.

The latest version is available on GitHub with a quick video guide on YouTube (https://github.com/VascoElbrecht/PrimerMiner). An internet connection as well as Mac OSX or Linux operating system is required, as PrimerMiner relies on Vsearch for clustering (https://github.com/torognes/vsearch). The program is configured with a txt file, which is created by running "batch\_config()". Target orders or families for which sequences should be obtained have to be specified in a csv file. Thus inclusion of a subset of taxa from a lower taxonomic level is possible. For example, for a certain order, a subset of families can be downloaded (e.g. only aquatic Coleoptera families) by specifying these in the second table column. Downloading data from higher than family taxonomy can cause the download to fail if group names are not unique and is thus not recommended. By running "batch\_download()" matching barcode sequences are downloaded and processed.
By default, complete and partial COI sequences are download from the BOLD and NCBI databases. Additionally, the target marker is extracted from mitochondrial genomes if available on NCBI. Sequences are then dereplicated and clustered using Vsearch with 3% similarly threshold. Majority consensus sequences for each OTU are written into a fasta file for each group. All raw sequencing data as well as intermediate files and summary statistics are automatically saved. Subsequently, the generated OTU sequences for each group have to be aligned with e.g. Geneious (Kearse *et al.* 2012) and can then be used in other primer development tools or visualized for manual primer development using the "plot\_alignments()" command.

## Primer development using PrimerMiner

The PrimerMiner package v0.2 was used to download COI and cluster sequences for the 15 most assessment relevant freshwater invertebrate groups from NCBI and BOLD (Accessed February 2015, Taxa table S1). Sequences were aligned with MAFFT v7.017 (Katoh *et al.* 2002) as implemented in Geneious 8.1.7 (Kearse *et al.* 2012) and the alignment for each group was visualized with PrimerMiner. The alignment plot was used to identify suitable primer binding sites manually. Two forward (BF1, BF2) and two reverse primers (BR1, BR2) were designed with high base degeneracy. Fusion primers were generated by adding Illumina adapters and inline barcodes as described in (Elbrecht & Leese 2015) to increase sequence diversity and allow for a one step PCR protocol.

## Testing of DNA metabarcoding primers on mock communities

Amplification success of the BF / BR primers was evaluated using ten mock communities, each containing a set of 52 freshwater invertebrates used in previous studies (Elbrecht & Leese 2015). The identical DNA aliquot and one step PCR protocol as in (Elbrecht & Leese 2015) was used for all four primer combinations. As in the previous studies, each sample was uniquely tagged from both sides, but sometimes only 25 ng instead of 50 ng DNA was used in PCR (see Figure S1). For each primer combination all ten samples were run in the PCR setup, using one PCR replicate per sample. Ready-to-load products were magnet-bead purified (left sided, 0.8x SPRIselect, Beckman Coulter, Bread, CA, USA) and quantified using the Qubit HS Kit (Thermofisher Scientific, Carlsbad, CA, USA). For each primer combination, equimolar amounts of amplicons were pooled into one library (taking fragment length differences into account, Figure S1). The library was sequenced on one lane of a HiSeq 2500 (rapid run, 2x250 bp) with 5% PhiX spike-in, carried out by the DNA Sequencing Center of Brigham Young University, USA.

Bioinformatic processing of high throughput data was kept as similar as possible to previous studies (Elbrecht & Leese 2015; Elbrecht *et al.* 2016). In short, reads were demultiplexed (R script **- will be supplied in next version of preprint**) and paired end reads merged using Usearch v8.1.1831 -fastq\_mergepairs with -fastq\_merge\_maxee 1.0 (Edgar & Flyvbjerg 2015). Where necessary, reads were converted into reverse complement. For each primer combination all ten replicates were pooled, and sequences which were present only one single time in the dataset (singletons) removed prior to clustering with Usearch (cluster\_otus, 97% identity, strand plus, includes chimera removal) (Edgar 2013). Dereplicated reads for each of the 40 samples (including singletons) were compared against the respective OTU dataset, using usearch\_global with a minimum match of 97% and strand plus. Like in previous studies, low abundance OTUs without at least one sample above 0.003% sequences assigned, were considered unreliable and excluded from the dataset. Taxonomy of the remaining OTUs was identified and manually verified using the BOLD and NCBI database. To ensure that taxonomy was consistently assigned across primer combinations and in comparison to the reference COI study (Elbrecht & Leese 2015), the most abundant sequence for each OTU in each sample was extracted using an R script, and the haplotype of all individual specimens assembled if possible.

## *In silco* evaluation of primers

PrimerMiner has powerful *in silico* primer evaluation capabilities, allowing or evaluation of single primers and primer pairs on any given sequence alignment. Unlike ecoPCR (Ficetola *et al.* 2010), PrimerMiner factors in the position and type of each primer / sequence mismatch, which gives a more comprehensive picture, as amplification success is highly dependent on a good matching 3 ' primer end (Piñol *et al.* 2014) (**add more refs**). Using the command "primer\_evaluation()" PrimerMiner calculates individual penalty score for each template to primer mismatch, factoring in the position and type of mismatch and thus giving a more realistic evaluation of amplification efficiencies. Penalty scores for position and mismatch type are fully customisable, by providing your own penalty tables. Mismatch evaluations for each sequence are stored in a table, allowing full transparency and processing in other programs. With the function "combine\_2\_primers()" two primer pairs can be evaluated against each other using the generated tables with "primer\_evaluation()", giving a maximum threshold under a primer pair is considered working for amplification. All metabarcoding primers shown in Figure 1 were evaluated against 30 insect orders alignments following the taxonomy by (Misof *et al.* 2014). Data was downloaded and clustered with PrimerMiner v0.3b on April 2016.

# 3) Results

## Developed primers using PrimerMiner

Using the alignments of 15 freshwater assessment relevant groups (Figure S2) we designed four primer pairs (Figure S3). Sequence coverage was increased 249 (SD=395) times on average by including COI barcode sequences to the mitochondrial reads (Figure S2). The two forward and two reverse primers show high base degeneracy to amplify as many insect taxa as possible. Amplified regions range from 217 bp for internal barcodes up to 421 bp for combinations using a degenerated version of the HCO2198 primer. While samples in this study were tagged uniquely from both sides, the inline barcodes allow for tagging of up to 72 samples for each primer combination (see Figure S4 for recommended primer combinations).

All four primer combinations were tested on ten invertebrate mock samples on a Illumina HiSeq sequencer. PCR efficiency varied across primer combinations, with PCRs involving the BF2 primer showing good amplification whereas those with BF1 primer always showing decreased yields (Figure S5). Amplification efficiency with fusion primers was always substantially lower than the positive control (standard COI Folmer primers, without Illumina tail). Sequencing was successful for all samples, with obtaining very similar amounts of sequences for all replicates (on average 1.55 million reads per sample, SD = 0.2, Figure S1A). Cluster density on the lane was low (402 k/mm2) yielding only 48.74% of the expected sequencing output, yet with good sequence quality (Q30 ≥ 92.17%, raw data deposited on SRA: SRX1619153). The amplified read lengths had an influence on the amount of sequences retained in bioinformatic processing. Longer amplicons have less overlap when PE merged and are thus excluded more often due to expected errors > 1 (Figure S1B). Additionally primer combinations that used the P5\_BF12 primer lost more sequences than other combinations, as ~1/5 of the reads had poor Phred scores. Furthermore, there were issues with the BF1 and BF2 primer which showed insertions or deletions on the 3' end affecting sequence length by 1-2 bp across all replicates (Figure S6). Some primer combinations also amplified up to 1.35% shorter or longer fragments than expected (Figure S7).

## Amount of taxa recovered

All insect taxa present in the mock samples were detected with each primer combination (Table 1), with exception of the BF1 + BR1 combination that failed to amplify the Scirtidae (Coleoptera). All primers failed for some of the other metazoan taxa, with the BF1 + BR2 combination showing the least amount of undetected taxa. In comparison to the traditional Folmer primers (Folmer et al. 1994), all BF / BR freshwater primers showed a more consistent and equal read abundance across the mock samples (Figure 3). As in Elbrecht et al. (2016), the standard deviation from the expected abundance and precision for the primer pairs was estimated, which summarizes the variance in amplification. Primer combination BF1 + BR1 showed the highest inconsistencies in read abundance, while the BF2 + BR2 combination showed even higher precision than a previously tested 16S marker (Elbrecht et al 2016). The proportion of detected non-insect metazoan taxa varied between primer combinations, with the combination BF1+BR2 detecting all but one taxon.

**Table 1:** Number of species recovered with the newly developed primers and data on 16S and Folmer primers from previous tests (Elbrecht & Leese 2015; Elbrecht *et al.* 2016).



\* Standard deviation (SD) of logarithmic sequence abundance for all samples that worked (specimens with < 0.003% read abundance discarded)

\*\* Precision defined as the SD of the mean log distance to the expected abundance, calculated for each morphotaxon.

## *In silico* evaluation of primers

Figure 4 gives an overview of 11 forward and 12 reverse primers evaluated against OTUs of all insect orders. Reference data for binding sites of the standard Folmer primers HCO and LCO was very limited, and six out of 29 orders had below 100 sequences in total. Primer efficiencies were very similar across orders but varied between primers. However, primers incorporating wobble bases (BF1, BF2, BR1, BR2, jgHCO2198) or inosin (Ill\_B\_F, ArF5, Il\_C\_R, ArR5) performed better than primers with no or just few wobble bases. Figure S8 shows an evaluation of primer pairs, giving results consisted to evaluations of individual primers. It should be noted that some primers from the literature are not only poorly matching because they lack wobble bases, but can be affected by additional problems (see Figure S2, "critical mismatches"). For instance, near the 3’ ends, the EPT-long-univR has a completely unnecessary second inosine at a conserved position, while the Uni-MinibarF1 had a "T" at a position where more than half of the reference OTUs had an "A". Furthermore, the L499 primer targets a highly variable region. Finally, certain primers show mismatches to particular groups, e.g. the ZBJ-ArtF1c and BR1 primers do not match well to sequences of Bivalvia.4) Discussion

We used PrimerMiner to develop four primer sets for freshwater invertebrates based on OTU sequence alignments generated of mitochondrial and COI barcodes from NCBI and BOLD. By not only using only mitochondrial genomes but also including COI barcode data from BOLD and NCBI primer design was built upon a solid and balanced data basis. Clustering helped to avoid overrepresentation of taxa with many sequences available in data bases, making sure that each species is represented by only a few majority consensus OTU sequences. Due to the high variability throughout the COI gene alignments (Sharma & Kobayashi 2014) and complexity of the task, we here decided to search for primers manually, instead of using available software solutions. We deliberately decided to not factor in nucleotide variability present in only few groups (mostly non-insect Metazoa), to limit the degeneracy of the primers to a reasonable level.

We further decided (and recommend) to develop COI metabarcoding primers internal of the Folmer region, as sequence coverage is still quite limited on the Folmer primer binding sites (Figure 4). We consider 100 OTUs for a given order as a minimum coverage for a primer binding site to capture its variability and select necessary wobble bases. Due to the codon degeneracy, larger alignments do not necessarily give much additional information. Thus, for the HCO binding region it is often possible to obtain reliable information while the sequence depth of the LCO primers is often limited to mitochondrial genomes (<100 OTUs available). In conclusion, PrimerMiner is an efficient and valuable tool to obtain and visualize meaningful sequence data to design and evaluate universal metabarcoding primers, tailored to the taxonomic groups present in the studied ecosystem.

## Amplification success of mock communities

All primer sets amplified the ten mock communities successfully. By factoring in the different amplicon lengths in library pooling we obtained similar amount of reads for each sample. All degenerated COI primers showed superior detection rates (up to 100% of insects and 98% of all morphotaxa) and more consistent read abundances compared to the standard Folmer barcoding primers that lacked any base degeneracy (Folmer *et al.* 1994; Elbrecht & Leese 2015). The primer sets BF2 in combination with BR1/BR2 even showed better detection rates and higher precision than a previously used 16S primer, which was tested on the same communities (Elbrecht *et al.* 2016). Also *in silico* analysis of the BF / BR primers against all insect taxa on NCBI and BOLD confirmed their excellent detection rates, with mean success rates near 100%. (Deagle *et al.* 2014) strongly argued against the use of degenerated primers to be used in DNA metabarcoding, proposing the use of ribosomal markers with more conserved binding regions instead. However, we here clearly show that highly degenerated COI primers are not only feasible but also superior to ribosomal metabarcoding of animals when it comes to primer performance and available reference databases.

While our developed primers show very reliable amplification results, there are also problems associated with the primers itself and well as the applied metabarcoding protocol. First, while the use of fusion primers potentially decreases the chance of index switching and reduces needed laboratory work, it also reduces PCR efficiency substantially (Schnell *et al.* 2015). While primer combinations involving BF2 primers were less affected by this issue, it was more pronounced with the BF1 primer especially in combination with BR1. Further, concerns have been raised by biases associated with use of indexed primers (O’Donnell *et al.* 2016). While we could not observe any obvious effects in our current data set (most taxa are detected to equal proportions regardless of primer index), there was a drop in sequence quality when using the BF12 primer. Whether this is a systematic effect associated with the primer index or a problem in e.g. primer synthesis / quality cannot be determined from this data set. However, independent of this possible bias, it did not have any effects on the number of detected taxa. Additionally, 17% of reads from the BF2+BR2 primer combinations were discarded due to low expected error values, as the overlap was limited with 250 PE sequencing of a 421 bp region on the HiSeq system. Further, with highly degenerated primers, the specificity of the primers decreases (Deagle *et al.* 2014) potentially amplifying non target regions or unexpected lengths. This effect was often minimally, with few sequences deviation from the expected length (below <0.5 % for most primers sets), with these numbers being potentially inflated by PCR / Sequencing errors and pseudo genes (e.g. Nemuridae). However, more problematically the BF1 and BF2 primers were affected by shifting effects making up to 40% of the sequences 1-2 bp shorter or longer at the primer binding side. It is not particularly clear what causes this effect, which can be also to observed lesser degrees in in datasets from previous studies (Elbrecht & Leese 2015; Elbrecht *et al.* 2016). Potentially the high degeneracy of the forward primers in combination with low diversity nucleotides at the primer’s 3 ' end (e.g. C[cta]TT[tc]CC in BF2) makes this effect particularly pronounced. Thus we recommend designing primers with two unique nucleotides on the 3 ' end. The effect of this minimal shifting shortens the read length by 1-2 bp which has no effect on detection on taxa (OTUs will still match the same reference taxon, regardless of 1-2 bp being clipped form the sequence). However, when calculation OTU based biodiversity indices, the small shift might lead to a bias in these metrics due to inflated OTU numbers. This might be countered by increasing the OTU clustering threshold to e.g. 4%, but we advice to take OTU based diversity measures with caution using the BF / BR primer set.

## Primer success is determined by base degeneracy

*In silico* analysis of 23 potentially suitable primers for COI DNA metabarcoding sowed that high primer degeneracy leads to the best amplification of freshwater and insect taxa. This was also confirmed experimentally, with the tested macroinvertebrate mock communities showing high primer bias with standard Folmer primers (Elbrecht & Leese 2015), and very consisted amplification with higher detection rates with the primers developed in this study. While other primers from (Gibson *et al.* 2014) and (Shokralla *et al.* 2015) probably lead to equally good amplification rates as the BF/BR primers, a lack of degeneracy can lead to substantial primer bias. While these biases might not affect PCR on single organisms for DNA barcoding strongly, they will substantially skew detection rates of complex multispecies bulk samples, in the worst case leading to taxa remaining undetected (Piñol *et al.* 2014; Elbrecht & Leese 2015). This might already be the case, when primers have to little degeneracy like the mlCOIint primers by (Leray *et al.* 2013), which have a maximum degeneracy of two nucleotides at each position. The mlCOIint primers were tested with two mock communities, containing DNA from previously barcoded taxa (Leray & Knowlton 2015). Leray et al. 2015 reported that up to 35% of taxa remained undetected, which is consistent with the *in silico* primer evaluations in this study.

Probably even more problematic are primers, which show no base degeneracy at all. While the primer bias associated with the high variation of the COI gene have been well known (Clarke *et al.* 2014; Deagle *et al.* 2014; Sharma & Kobayashi 2014; Piñol *et al.* 2014; Elbrecht & Leese 2015), primers without base degeneracy like ZBJ-Art by (Zeale *et al.* 2010) are widely used and recommended e.g. for gut content analysis (Pompanon 2012???). This can be really problematic, as large proportions of biodiversity might be missed or underrepresented in studies using these primers. Even when primers have good success rates for barcoding of single specimens (Meusnier *et al.* 2008), they are likely to introduce huge primer bias in metabarcoding studies. Thus careful evaluation of primers to the specific groups of interest in the planned metabarcoding study is curtail. PrimerMiner provides helpful tools to obtain and evaluate group specific sequence data needed for theses evaluations. Further, the efficiency of popular primer sets should additionally tested using mock communities, to detect specific biases introduced by the primers or the specific metabarcoding protocol.

## Recommended approaches for assessment of insects and freshwater taxa

The success of every DNA metabarcoding project depends on well designed primers, which amplify the target communities as consistent as possible. Amplification bias depends on primer binding regions, which can be more conserved in ribosomal genes than in COI. Thus 18 and 16S markers have been proposed as suitable alternatives, despite lacking comprehensive reference databases (Clarke *et al.* 2014; Deagle *et al.* 2014; Elbrecht *et al.* 2016). Given the *in silico* evaluations and better performance of the BF2 + BR1 / BR2 primer sets, it can be settled that ribosomal markers are not necessary for reliable DNA metabarcoding on animals. Thus, to only remaining challenge is to find the ideal COI metabarcoding marker, suitable for your groups of interest. PrimerMiner can be a helpful tool to evaluate existing markers and if needed build new ones. Also we encourage to try and evaluate combining primers from different primer sets and test them *in silico*.

When using DNA metabarcoding approaches for ecosystem assessment, protocols from the literature should be critically evaluated. We recommend using the illumina HiSeq sequencer in rapid run mode (250 bp PE reads) and include replications to reduce changes fro tag switching and exclude false OTUs from the dataset. While we have previously encouraged the use of fusion primer due to their ease of use (single step PCR, (Elbrecht & Leese 2015)), we have to acknowledge that they decrease PCR efficiency, and thus two step PCRs might be better suited for environmental samples which often contain PCR inhibitors.

Additionally, metagenomic approaches using enrichment for mitochondrial genomes could be suitable for assessment of ecosystems, which potentially less bias as the PCR amplification step can be omitted (Liu *et al.* 2015). However, as briefly discussed in (Elbrecht *et al.* 2016), metagenomics methods have to be further validated and mitochondrial reference genome libraries completed (Dowle *et al.* 2015). Thus, the selection of a specific metabarcoding or metagonomics approach depends on future developments, available resources and expertise in the laboratories. However, it is clear that if one decides to apply DNA metabacoding, primers have to be carefully evaluated or even newly developed to optimally amplify the targeted groups of the specific project.

## Conclusions

With PrimerMiner, we have developed a useful R package for primer development and evaluation, which we here used to design new DNA metabarcoding primers targeting freshwater invertebrates. Our *in silico* evaluations as well as mock communities metabarcoding experiments clearly indicated that with highly degenerated COI primers almost 100% of the taxa were not only detected, but also amplified with highly similar read numbers. Thus, we argue that COI is the marker of choice to use in animal metabarcoding, dismissing other markers such as ribosomal markers as a suitable alternative due to the poor reference data for these. We additionally encourage a more thorough *in silico* and *in vivo* evaluation of existing primers, as many are not suitable for DNA metabarcoding due to low base degeneracy, potentially high primer bias or critical design flaws.

**Acknowledgements:** FL and VE are supported by a grant of the Kurt Eberhard Bode foundation to FL. We thank Bianca Peinert for her help with primer development, and Edward Wilcox and ScienceExchange for HiSeq sequencing.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

# Figures

**Figure 1**: Selection of potential COI primer sets for DNA metabarcoding of insects, targeting the Folmer region. Primer pairs are shown based on typically used / suggested combinations used in the literature, but also other combinations are possible. Table S1 gives an overview of the exact primer sequences and sources.

**Figure 2**: Overview of the principle behind the PrimerMiner package for sequence downloading and clustering. Both mitochondrial genomes as well as partial gene sequences are downloaded and clustered, to make utilise the maximum of available sequence information while minimising biases introduced by overrepresented taxa in the sequence data. Primer trimming is necessary if database sequences have not been properly trimmed.

**Figure 3**: Comparison of the COI Folmer primer performance and the four tested primer combinations newly developed. All primer combinations were tested with the same ten bulk samples each containing 52 morphologically distinct macroinvertebrate taxa. The 52 taxa are shown on the x-axis with the relative number of reads obtained for each morphotaxon by black dots on the logarithmic y-axis (mean read abundance indicated by red circles), for each respective primer combination. Sequence abundance was normalized across the ten replicates and the amount of tissue used in each DNA extraction. Only OTUs with a minimum read abundance of 0.003% in at least one of the ten samples were included in analyses. Number of samples for which a morphotaxon was not detected is indicated by orange and red numbers in each plot. A thick vertical line in light red indicates if a morphotaxon was not detected.

**Figure 4**: **Preliminary data, error penalties subject to changes / kind of mismatch not jet implemented!** Overview of *in silico* evaluation of primer performance using PrimerMiner with OTU data from 29 insect orders. Primer performance is shown for each group in pie charts (red = failure, green = working, grey = missing data / gaps). Every primer sequence match with a mismatch penalty score of above 50 is considered a failure. Every order with at least 100 OTUs is used for calculation of the average and the box plot showing the mean penalty scores for each group. Good primers have a yellow background.References

Brandon-Mong, G.J., Gan, H.M., Sing, K.W., Lee, P.S., Lim, P.E. & Wilson, J.J. (2015). DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. *Bulletin of Entomological Research*, **105**, 717–727.

Cameron, S., Rubinoff, D. & Will, K. (2006). Who Will Actually Use DNA Barcoding and What Will It Cost? *Systematic Biology*, **55**, 844–847.

Clarke, L.J., Soubrier, J., Weyrich, L.S. & Cooper, A. (2014). Environmental metabarcodes for insects: in silicoPCR reveals potential for taxonomic bias. *Molecular ecology resources*, **14**, 1160–1170.

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**, 20140562–20140562.

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

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

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

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 (M. Hajibabaei, Ed.). *PloS one*, **10**, e0130324–16.

Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.-N., Coissac, E., Boyer, F. & Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, **4**, e1966–12.

Ficetola, G.F., Coissac, E., Zundel, S., Riaz, T., Shehzad, W., Bessière, J., Taberlet, P. & Pompanon, F. (2010). An in silico approach for the evaluation of DNA barcodes. *BMC genomics*, **11**, 434.

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

Geller, J., Meyer, C., Parker, M. & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome coxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular ecology resources*, **13**, 851–861.

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

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*.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H. & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 14812–14817.

Katoh, K., Misawa, K., Kuma, K.-I. & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research*, **30**, 3059–3066.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Larsen, S., Pace, G. & Ormerod, S. (2011). Experimental effects of sediment deposition on the structure and function of macroinvertebrate assemblages in temperate streams. *River Research and Applications*, **27**, 257–267.

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

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T. & Machida, R.J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in zoology*, **10**, 1–1.

Liu, S., Wang, X., Xie, L., Tan, M., Li, Z., Su, X., Zhang, H., Misof, B., Kjer, K.M., Tang, M., Niehuis, O., Jiang, H. & Zhou, X. (2015). Mitochondrial capture enriches mito-DNA 100 fold, enabling PCR-free mitogenomics biodiversity analysis. *Molecular ecology resources*.

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.

Misof, B., Liu, S., Meusemann, K. & Peters, R.S. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *…*.

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

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*, 1–12.

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

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.

Sharma, P. & Kobayashi, T. (2014). Are ‘universal’ DNA primers really universal? *Journal of applied genetics*, **55**, 485–496.

Shokralla, S., Porter, T.M., Gibson, J.F., Dobosz, R., Janzen, D.H., Hallwachs, W., Golding, G.B. & Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports*, **5**, 9687.

Shokralla, S., Zhou, X., Janzen, D.H., Hallwachs, W., Landry, J.-F., Jacobus, L.M. & Hajibabaei, M. (2011). Pyrosequencing for Mini-Barcoding of Fresh and Old Museum Specimens (S.-O. Kolokotronis, Ed.). *PloS one*, **6**, e21252.

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

Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L.H. (2012). Environmental DNA. *Molecular Ecology*, **21**, 1789–1793.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. & Rozen, S.G. (2012). Primer3--new capabilities and interfaces. *Nucleic acids research*, **40**, e115–e115.

Van Houdt, J.K.J., Breman, F.C., Virgilio, M. & De Meyer, M. (2010). Recovering full DNA barcodes from natural history collections of Tephritid fruitflies (Tephritidae, Diptera) using mini barcodes. *Molecular ecology resources*, **10**, 459–465.

Zeale, M.R.K., Butlin, R.K., Barker, G.L.A., Lees, D.C. & Jones, G. (2010). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular ecology resources*, **11**, 236–244.