

A broadly applicable COI primer pair and an efficient single-tube amplicon library preparation protocol for metabarcoding

Oskar Rennstam Rubbmark  | Daniela Sint  | Nina Horngacher | Michael Traugott 

Mountain Agriculture Research Unit, Institute of Ecology, University of Innsbruck, Innsbruck, Austria

Correspondence

Oskar Rennstam Rubbmark, Mountain Agriculture Research Unit, Institute of Ecology, University of Innsbruck, Innsbruck, Austria.

Email: oskar.rubbmark@hotmail.com

Funding information

Austrian Science Fund, Grant/Award Number: P26144; BMWF, Grant/Award Number: SPA05/122; Mountain Agriculture Research Unit; regional government of Tyrol

Abstract

The nucleotide variation in the cytochrome c oxidase subunit I (COI) gene makes it ideal for assigning sequences to species. However, this variability also makes it difficult to design truly universal primers. Here, we present the forward primer “Sauron-S878,” specifically designed to facilitate library preparation for metabarcoding. This primer is modified to improve the coverage of terrestrial species compared to the primer mCOLintF, optimized for aquatic systems, which raised the *in silico* coverage from 74.4% to 98.3% of available NCBI sequences (perfect match in 3' region, up to three mismatches in remaining primer). When paired with the reverse primer “jgHCO2198” (fragment length ~313 bp), these primers amplified 98.4% of 255 tested DNA extracts from various taxa, which are better than many other common COI barcoding primers. Furthermore, a single-tube protocol was developed, wherein these primers amplify the target gene, and attach MID and Illumina sequencing adapters in one reaction. This eliminates the need for re-amplification or enzymatic ligation during library preparation while keeping the flexibility to modularly combine primers and MID. Using the single-tube approach, three replicates of three mock samples were sequenced on a MiSeq platform with no adverse effects compared to commercial Nextera indexing kits. From this run, 75% of all included taxa could be recovered, with no considerable bias among taxonomic groups. Despite the fact that 98.4% of the extracts were confirmed to amplify *in vitro*, this number was lower than expected. A reason for this discrepancy was a clear link between the relative concentration of a specific DNA type in the template and the number of returned reads for this DNA. We would argue that such a bias may be especially problematic in metabarcoding where samples usually contain trace DNA in unknown amounts. However, how this affects the completeness of metabarcoding results has yet been poorly investigated.

KEYWORDS

diet analysis, DNA barcoding, environmental DNA, high-throughput sequencing, next-generation sequencing, primer bias

1 | INTRODUCTION

Metabarcoding is an easy to use and powerful method that increasingly is being employed to detect the presence of species in applications ranging from the analysis of community bulk samples (Ji et al., 2013; Yu et al., 2012) to biodiversity assessments from environmental DNA (Bohmann et al., 2014; Taberlet, Coissac, Hajibabaei, & Rieseberg, 2012; Thomsen & Willerslev, 2015) and studies of trophic interactions (De Barba et al., 2014; Deagle, Kirkwood, & Jarman, 2009; Pompanon et al., 2012; Valentini et al. 2009). It combines DNA-based identification of species (barcoding) with next-generation sequencing (NGS or high-throughput sequencing—HTS) by using so-called universal primers, usually targeting a specific group of interest in order to mass amplify DNA from collected samples containing mixes of DNA (Taberlet et al., 2012). Metabarcoding has considerable advantages over more traditional approaches, where taxonomic assignment is done morphologically. For example, environmental samples can be collected in a way that minimizes disturbances to sensitive ecosystems compared to more traditional sampling methods (De Barba et al. 2010). In addition, by using existing sequence databases for species identification, hard to come by taxonomic expertise can be reduced.

By convention, the most commonly used gene for barcoding of Metazoan diversity has so far been the mitochondrial cytochrome c oxidase subunit one (COI) gene. The main reason for this is that, even though other genes have been shown to work better to identify plants (*rbcL*, *matK*; CBOL Plant Working Group, 2009), fungi (ITS; Schoch et al., 2012) and bacteria (16S; Tringe & Hugenholtz, 2008), COI has usually been suitable in identifying most animals to species level (Hebert, Cywinska, & Ball, 2003). Because of this, it was selected as the target gene for the barcode of life initiative (BOLD), and so far, the number of animal species sequenced for this gene fragment (~2.3 million sequences from ~280,000 species in GenBank) is much greater than for other common barcoding genes such as 16S (~380,000 sequences from ~90,000 species) or 18S (~170,000 sequences from ~70,000 species). A reason for this is that these alternative genes generally offer lower taxonomic resolution, which provides a strong argument for why COI is a good candidate for metabarcoding of animals. Particularly, as even though alternative barcoding regions may be more suited for primer design, these will restrict scientists to treating individual taxa as observed operational taxonomic units (OTUs; Ji et al., 2013). This can have negative effects on the quality of results, especially when reference sequences are not available or when species cannot be distinguished based on the delivered sequence information (e.g., 18S). This will hamper species identification and make important characteristics such as species traits difficult or impossible to assign.

The suitability of COI for metabarcoding is thus high; however, it is also being questioned (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014) due to the fact that it has been difficult to design truly universal primers for this gene (Geller, Meyer, Parker, & Hawk, 2013). The reason for this is that as a coding gene, it exhibits considerable variation in every 3rd base, which means that

highly conserved regions are lacking (Hebert et al., 2003). In fact, it has been shown that most tested COI primers that claim to be universal are only marginally so and often fail to amplify many taxa (Clarke, Soubrier, Weyrich, & Cooper, 2014; Deagle et al., 2014; Elbrecht & Leese, 2015; Leray et al., 2013). This means that among species that could be identified and confirmed present in a sample, there may be an unknown range of false absences due to methodological error. For interpreting results, this then becomes problematic. On one hand, this may not be of concern when comparing structural differences between samples (beta diversity), but could become very problematic for comparing differences in alpha diversity (Clooney et al., 2016). So far, however, the majority of the testing done to identify primer bias has been based on small sets of species or sequences from mainly aquatic and invertebrate taxa. Consequently, even for many of the more commonly used metabarcoding primers, there is still no comprehensive knowledge on their taxonomic coverage and for which taxa they work well or poorly.

For metabarcoding, it is furthermore especially important that this testing is done, not only by comparing the presence or absence of an amplification *in vitro*, but also by taking into account the differences in match and/or amplification efficiency between primers and DNA sequences. This is needed because, even when primers amplify the DNA of a certain species, if the fit of primers is different between taxa, the better fitting taxa will be amplified preferentially in a competitive reaction (Bru, Martin-Laurent, & Philippot, 2008; Green, Venkatramanan, & Naqib, 2015). In mixed samples, this will be further problematic as such biases will interact with, for example, unequal amounts of biological material (e.g., tissue) or different copy numbers of the targeted gene depending on tissue type or species (Pompanon et al., 2012) and further affect the probability to detect a species. While there are attempts to develop PCR-free approaches that are also suitable for metabarcoding (Creer et al., 2016; Denver, Brown, Howe, Peetz, & Zasada, 2016; Paula et al., 2016), currently there are still too many limitations to replace the so far used target sequencing in the near future.

Still, these different sources of bias are troubling, and solutions have, for example, included the design of primers that amplify less variable barcoding genes (e.g., “ribosomal markers” Clarke et al., 2014; Deagle et al., 2014). However, by using barcoding regions with less variability, taxonomic resolution will decrease and therefore it is as mentioned, desirable to attempt to reduce these biases in order to retain the use of COI. Fortunately, a number of recent studies have shown that while COI is variable, with some of the degenerate primers available such as *mCOLintF* (Leray et al., 2013) and *tgHCO2198* (Geller et al., 2013), it is nevertheless possible to get at least for some groups a comparable taxonomic coverage to alternative barcoding regions (Clarke, Beard, Swadling, & Deagle, 2017; Elbrecht et al., 2016). The variation within COI does, however, mean that a large amount of degeneracy is needed. This is because even single mismatches close to the 3' end of primers can significantly reduce the amplification efficiency of a given taxon in a PCR (Green et al., 2015; Lefever, Pattyn, Hellemans, & Vandesompele, 2013; Stadhouders

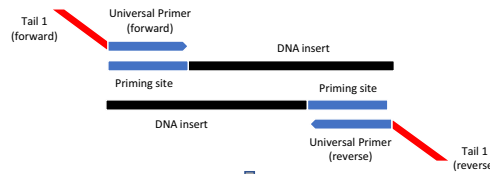
et al., 2010). This will cause a gradient of amplification efficiency between DNA types in competitive PCRs (Bru et al., 2008; Green et al., 2015) and consequently biased sequencing results. Furthermore, within degenerate primer pools, each primer sequence has a different melting temperature and thus amplifies at a different efficiency in PCRs depending on cycling conditions in addition to the increased risk of mis-priming (Leray & Knowlton, 2017).

This means that even with completely degenerate primer pools, taxonomic biases can only be reduced but not circumvented. There is, however, the possibility to minimize such problems by incorporating a universal tail at the 5' end of the primers which is not complementary to the template (Green et al., 2015). This approach has been used for conventional barcoding, where the incorporation of, for example, m13 tails in primers can reduce the sequencing of

Box 1. Flowchart of the single-tube library preparation procedure

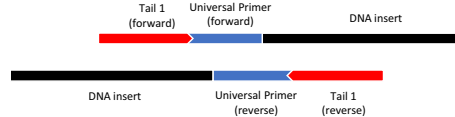
First step of PCR

- Step I cycles of the PCR binds universal primers
- The tails, that are not complementary to templates, remain unbound
- The high annealing temperature during these cycles prevents the step II primers (which are complementary to the tails) from binding



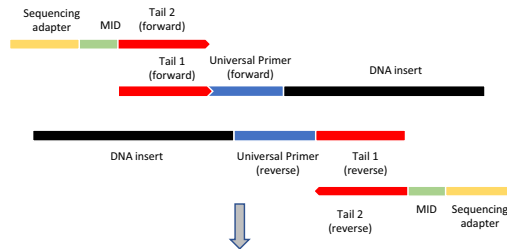
15 step I PCR cycles (high annealing temperature)

- After step I cycles the majority of the templates now has an incorporated universal (forward or reverse) tail as well as a copy of respective universal primer



Second step of PCR

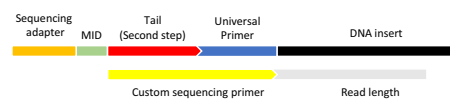
- Step II cycles (conducted at a lower annealing temperature)
- Tail 2 primers are complementary to the tails that were incorporated during step I and use these as priming sites to attach sequencing adapters and MID



10 step II PCR cycles (low annealing temperature)

- After the full set of PCR cycles are conducted, the majority of the PCR product will be the amplified template with attached tail, MID and sequencing adapter (A)
- By using universal primers as priming sites for custom sequencing primers (A) the effective read length is extended (by the length of the universal primer)
- This differ from commercial kits (B) where tails are used as priming sites which lead to universal primers being sequenced

A



B

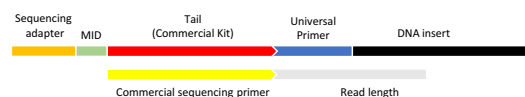
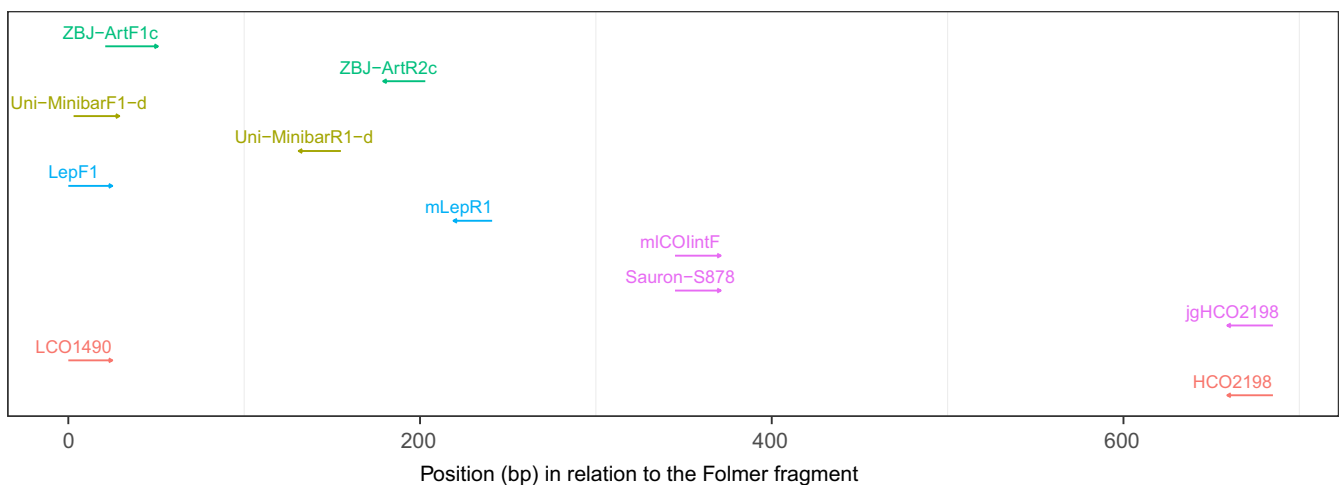


TABLE 1 Primers evaluated and designed for this study

Primer	Sequence	Type	References
Uni-MinibarF1-d	TCYACTAATCATAAAGATATTGGYAC	Universal Forward	Jordaens et al. (2013)
Uni-MinibarR1-d	AAAATTATAATAAARGCRTGRGC	Universal Reverse	Jordaens et al. (2013)
ZBJ-ArtF1c	AGATATTGGAACWTTATATTTTATTTTGG	Universal Forward	Zeale et al. (2011)
ZBJ-ArtR2c	WACTAATCAATTWCCAAATCCTCC	Universal Reverse	Zeale et al. (2011)
LepF1	ATTCAACCAATCATAAAGATATTGG	Universal Forward	Hebert, Penton, Burns, Janzen, and Hallwachs (2004)
mLepR1	CCTGTTCCAGCTCCATTTT	Universal Reverse	Hajibabaei, Janzen, Burns, Hallwachs, and Hebert (2006)
mCOLintF	GGWACWGGWTGAACWGTWTAYCCYCC	Universal Forward	Leray et al. (2013)
lgHCO2198	TAIACYTCIGGRTGICCRAARAAYCA	Universal Reverse	Geller et al. (2013)
Sauron-S878	GGDRCWGGWTGAACWGTWTAYCCNCC	Universal Forward	This study
Sauron-Tail-S879	CacctgcttctaaatGGDRCWGGWTGAACWGTWTAYCCNCC	Sauron-S878+ tail	This study
lgHCO2198-Tail-A867	CacttcgactctttacTANACYTCNGGRTGNCCRAARAAYCA	JgHCO2198-A867+tail	This study
i7-Tail-A867	CAAGCAGAAGACGGCATAACGAGAT[TAAGCGA]cacttcgactctttac	MiSeq i7+[MID]+ tail	This study
i5-Tail-S879	AATGATACGGCGACCACCGAGATCTACAC[CTCTCTAT]cacctgcttctaaat	MiSeq i5+[MID]+ tail	This study
SEQ-S879	CACCTGCTTCTAAATGGDRCWGGWTGAACWGTWTAYCCNCC	Sequencing primer (read 1, position 18)	This study
IND-A867	TGRTTYTTYGGNCAAYCCNGARGTNTAGTAAAGAGTCGAAGTG	Indexing primer (position 19)	This study
SEQ-A867	CACTTCGACTCTTTACTANACYTCNGGRTGNCCRAARAAYCA	Sequencing primer (read 2, position 20)	This study

Note. Lowercase letters in the Sauron-Tail-S879 and lgHCO2198-tail-A867 primers show the location of universal tails. For the step 2 primers, i5-Tail-S879 and i7-Tail-A867, MIDs are within brackets, followed by MiSeq i7 and i5 adapters.

**FIGURE 1** Position of each of the tested primers in relation to the position of the Folmer primers (LCO1490 and HCO2198). Direction (3' end) of each primer is indicated by arrows, and colors indicate primer pairs tested together

nonspecifically bound primers and boost sequence quality (Ivanova et al. 2007). Using such tailed primers, the PCR can be divided into two steps, where the first few cycles of the PCR (step I) allow a mix of primers of different fit to anneal to matching templates, while the universal tail remains unbound. Later, in the exponential buildup of PCR product (step II), the tail incorporated into the amplicons generated during the first few cycles, now serves as an identical priming site across all templates.

Besides reducing the potential for unequal DNA amplification, these tails have the additional benefit that they facilitate an easy attachment of sequencing adapters and multiplex identifiers (MIDs) needed for NGS. For example, for amplicon sequencing on the MiSeq or HiSeq platforms (Illumina, San Diego, CA, USA), this is commonly done by incorporating a Nextera (Illumina)-specific tail into locus-specific primers. This tail is later used by commercially available Nextera indexing kits for tagmentation of samples during re-amplification of adapters. As this is performed in two separate PCRs, labor costs and contamination risks caused by additional handling increase. However, as an alternative to this approach, it may be possible to perform both amplification and re-amplification steps in a single PCR, instead of two separate PCRs. This could be done in a PCR comprising two types of cycles (from here on referred to as “single-tube PCR,” Box 1). This requires that tails that are significantly shorter than the Nextera tails are used, in order to allow tails to have lower annealing temperatures than the locus-specific primers. Using these tails and by adapting cycling conditions as described by Clarke et al. (2014), it should theoretically be possible to perform both amplification and re-amplification steps in a single PCR. This single-tube PCR approach would, if effective, have considerable advantages over traditional NGS library preparation procedures. Most importantly, it could not only significantly reduce the biases from differential amplification of templates during PCR, but also reduce contamination risks and labor as there would be no need for repeated PCRs and handling.

There are three aims for this study:

- (1) To characterize how commonly used barcoding primers differ in amplification success across all taxonomic animal groups where sequences are available in public databases. Both in order to shed light on, in particular terrestrial, taxa that may have been missed in previous studies investigating published primers, and to aid scientists in selecting appropriate primers for future studies.
- (2) To test whether it is possible to increase the fit of the best of these primers and thereby reduce biases caused by a lack in conserved priming sites within COI. Specifically, we aim to test whether the amplification of DNA is sufficiently stable across taxonomic groups to allow the COI gene, with its large number of published sequences, to be used for metabarcoding purposes.
- (3) To equip designed primers with a universal tail and optimize a single-tube PCR protocol to achieve sequencing results comparable to commercially available kits such as the Nextera indexing kit. The rationale behind this being that such a method would not only

counteract biases associated with mixed primer pools but also enable a cost- and labor-effective way of incorporating MIDTs and NGS-specific adapters with a minimized risk of cross-contamination.

2 | MATERIAL AND METHODS

2.1 | Generation of a reference sequence alignment of the COI barcoding region

Using the search criteria “metazoa AND COI” (including alternative dictions for COI), we retrieved all of the sequences available for the barcoding region of COI (~1.3M sequences; May 2015) from NCBI (www.ncbi.nlm.nih.gov/genbank). In order to find conserved regions and compare priming sites located within this gene, it was necessary to align all of these sequences. As creating such a large alignment of sequences would be very computationally demanding, we first reduced this material to include only one (the longest) representative of each unique species. The reduction to only one representative sequence per species was done to eliminate the problem that uneven numbers of representative sequences are available for different taxa. If all sequences would have been kept for the analysis, there would have been the risk that the evaluation of primer fit would be biased toward species with high numbers of representative sequences. This gave a total of ~170,000 unique sequences from throughout the animal kingdom which were aligned using MAFFT 7.271 (Katoh & Standley, 2013). From within this alignment, the “DNA barcoding fragment” was identified and annotated.

2.2 | In silico evaluation of published primers

First, we conducted a literature search to investigate which universal primers have so far been used for metabarcoding purposes and selected those that to date had been used most frequently (Table 1). It should here be noted that most of these primers were not designed with complete universality in mind, even though they often have been used in later studies as if they were. Second, we located the position on which each of these primers fit within the alignment of the reference sequences (Figure 1). Then, for each priming site and primer, we implemented a pattern matching algorithm using Biostrings (Pages, Abouyou, Gentleman, & DebRoy, 2016) to investigate how the number of matching bases between each primer and individual sequences varies. As amplification depends on both the amount of mismatch between primers and sequences, as well as where this mismatch is located (Stadhouders et al., 2010), we here set up two matching criteria: (a) We allowed either 0, 1, 2, or 3 mismatches within the four bases closest to the 3' end of the primer and (b) nested within each of these, an additional criterion of either 0, 1, 2, 3, or 4 mismatches within the remaining bases of the primer. For each primer and mismatch combination, we then determined the proportion of sequences that met each criterion from within

the reference alignment. This allowed us to observe how the fit of each primer varies between a perfect match to a maximum of 7 mismatching bases (three of which would be located in the first four bases of the 3' end), when amplification is not expected to occur. Furthermore, we chose to evaluate each primer individually to not be restricted to reference sequences where both priming sites are available. This greatly increased the number of reference sequences analyzed for each primer as DNA sequences often did not cover both priming sites for the general primer pairs under evaluation.

2.3 | In vitro evaluation of published and newly designed primers

A large set of DNA extracts from a wide range of animal taxa is available from current as well as previous projects within our laboratory. These extracts were supplemented with additional material to cover a total of 255 DNA extracts from taxonomic groups ranging from nematodes to vertebrates (Table S1). All sample material was supplied and identified by experts, and, if not already DNA extracted, samples were extracted using a Biosprint96[®] robotic platform with the Biosprint 96 DNA Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's recommendations, except that TES buffer (0.1 M TRIS, 10 mM EDTA, 2% SDS; pH 8) was used for lysis and 1× TE buffer for elution. Furthermore, to verify that all reference extracts contained amplifiable DNA, each sample was tested in a PCR with universal or phylum-specific primers.

To evaluate taxonomic coverage for selected and newly designed (see below) primers (Table 1) under laboratory conditions, a series of PCRs were conducted on each reference DNA extract. PCR conditions were standardized for all primer combinations and performed in a reaction mix containing 1.5 µl DNA extract, 0.5 µl bovine serum albumin (BSA; 10 mg/ml), 1 µl of each primer (10 µM), 1 µl dNTP (2 mM), 2 µl reaction buffer with MgCl₂ (NEB, Ipswich, US) with an additional 0.48 µl MgCl₂ (25 mM; NEB) to obtain a final concentration of 3 mM, 0.05 µl of oneTaq (5 U/µl; NEB), and PCR grade water to adjust the volume to 10 µl. Reactions were performed in a Nexus Mastercycler with cycling conditions set to 2 min at 94°C, 35 cycles of 20 s at 94°C, 30 s at 50°C, 60 s at 68°C, and the final elongation for 3 min at 68°C. Amplification of all PCR products was verified using the QIAxel (Qiagen) capillary electrophoresis system with the software BioCalculator 3.2 (Qiagen, Method: AL320) where a DNA fragment of the expected length and with signal strength of ≥0.07 relative fluorescent units was defined as a successful amplification.

2.4 | Primer design

Informed by in silico results, we used the consensus sequence of the most variable bases in the least variable priming sites to design a new forward primer called Sauron-S878. In order to increase the fit across all taxonomic groups, additional degeneracy was incorporated into the primer mCOLintF (Leray et al., 2013) which was optimized for aquatic systems and is a modification of the primer C1-J-1859 published by Simon et al. (1994). While similar to this primer,

Sauron-S878 was modified to have an improved fit also for terrestrial species and is intended to be used together with the reverse primer jgHCO2198. JgHCO2198 is a highly degenerated version of the commonly used barcoding primer HCO2198 (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) which in this study had an overall good performance during both in silico and in vitro testing. The combination of Sauron-S878 and jgHCO2198 will amplify a fragment of ~313 bp (Figure 1), which is suitable for metabarcoding purposes using available NGS platforms (Pompanon et al., 2012).

As part of the objective of this study was to enable the use of these primers to reliably amplify COI for NGS purposes, we adapted both Sauron-S878 and jgHCO2198 with individually designed universal 5' end tails (referred to as "tail 1") in the primers Sauron-Tail-S879 and jgHCO2198-Tail-A867 (Table 1). These primers can be used to incorporate additional sequences, such as MIDAs, that allow samples to be separated after sequencing, as well as platform-specific adapters, into the 5' ends of amplicons. Here, this was done for the MiSeq platform (Illumina) in the complementary linker sequences i5-Tail-S879 and i7-Tail-A867 (Table 1; referred to as "tail 2"). These universal tails together with the locus-specific primers also double as a template for a set of custom sequencing primers designed for the MiSeq platform that was shown to be effective in a MiSeq sequencing run (read 1 [position #18]—SEQ-S879, Index [position #19]—IND-A867, and read 2 [position #20]—SEQ-A867; Table 1). One of the advantages with these custom sequencing primers is that by including the universal primers as templates, these are prevented from being sequenced. This also results in a read length that is extended by the length of the primer and thus allows a greater overlap (improved quality) when pairing reads during processing of results compared to the Nextera indexing kit, where the sequencing primer binds solely to the long Nextera linker.

To simplify library preparation, we developed a new protocol for attaching these in one PCR as proposed by Clarke et al. (2014). In this two-step single-tube PCR approach (flowchart and graphical presentation of included steps in Box 1), the tail 1's were designed to have a lower annealing temperature than the used locus-specific primers, as well as a poor fit for all reference sequences. This was done to prevent them from annealing to templates during the first step cycles. During these cycles, when annealing temperature is kept high, amplification occurs using universal primers, leading to the unbound tail 1's being correctly incorporated into the PCR product. During later second step cycles, both forward and reverse universal tail 1's are used as priming sites for tail 2's. During this second step, annealing temperature is lowered in order to allow the tail 2's to bind to tail 1 templates. This activates the step 2 primer and allows amplification to occur on the step one tail, which, after the first step, will be identical across all amplicons.

As the tails may alter the characteristics of original primers, tail-incorporated primers (Sauron-Tail-S879, jgHCO2198-Tail-A867) were tested in parallel with selected and designed universal primers during in vitro testing to ensure that this tail did not negatively affect amplification efficiency for any DNA extracts. PCR conditions were optimized by varying primer concentrations, annealing temperature,

and cycle numbers between the first and second step of the single-tube PCR until the majority of the final PCR product was a DNA fragment of the desired length containing platform adapters, MIDAs as well as template DNA. We again confirmed that no DNA extracts failed to amplify using this approach where amplification was previously successful and subjected a subset of these extracts to Sanger sequencing to confirm the incorporation of platform adapters and MIDAs.

After optimization, the PCR set-up that produced the strongest products of the targeted fragment with a minimum of unwanted amplification of primer artifacts was a reaction mix of 2 μ l DNA extract, 0.5 μ g of BSA (10 mg/ml), 0.15 μ l of each 1st step primer (10 μ M; Sauron-Tail-S879, jgHCO2198-Tail-A867), 1 μ l of each 2nd step primer (10 μ M; i5-Tail-S879, i7-Tail-A867), 5 μ l of Multiplex PCR Kit reaction mix (Qiagen), and PCR grade water to adjust the volume to 10 μ l. Cycling conditions were set to 15 min at 95°C, 15 1st step cycles of 30 s at 94°C, 90 s at 55°C, 60 s at 72°C, and 20 2nd step cycles of 30 s at 94°C, 90 s at 45°C, 60 s at 72°C, and a final elongation of 10 min at 72°C.

2.5 | Evaluation of newly designed primers and single-tube PCR approach for NGS

For the final evaluation, we assembled three pooled mock communities from among the 253 DNA extracts (there was no sufficient volume remaining after in vitro testing of two DNA extracts that had to be excluded) of 84–85 species each (Table S1). To do this, the DNA concentration of each of the reference extracts was first quantified (mean of three individual measurements of different DNA aliquots) using a NanoDrop 1000 (Thermo Scientific,

Wilmington, DE, USA). After this, DNA extracts were picked such that each mock sample included a unique mix of extracts representing an equal coverage of taxonomic groups and DNA extract concentrations (Table S1). Each of these mock samples was then further subdivided into (a) an unbalanced pool, where an equal volume of each DNA extract was pooled, and (b) a balanced pool where all DNA extracts with a DNA concentration higher than 5 μ g/ μ l were diluted to 5 μ g/ μ l and then mixed with the remaining extracts (<5 μ g/ μ l DNA) at equal volumes. As the NanoDrop cannot measure low DNA concentrations precisely, the DNA extracts with a concentration below 5 μ g/ μ l were left unaltered, to avoid pretending a higher precision than could be obtained. Each mock sample was replicated three times during the following tests to account for variability in PCRs and NGS. This gave a total of 18 individually tested mock communities (three sets of DNA extracts \times 2 balance levels \times 3 replicates), each uniquely tagged with MIDAs using the developed single-tube PCR approach (further on referred to as single-tube mock communities).

2.6 | Library preparation

Each of the 18 single-tube mock communities was prepared in 50 μ l reactions and combined into a ready to sequence library, by first amplifying these using the developed primers and the single-tube PCR approach (for details on PCR conditions and protocols, see Data S1). Furthermore, to compare how the single-tube approach performed in comparison with commercial kits such as the Nextera indexing kit (Illumina), an additional library was prepared that included each three replicates of the three mock communities with balanced concentrations. This second library included a version of each of the LepF1/

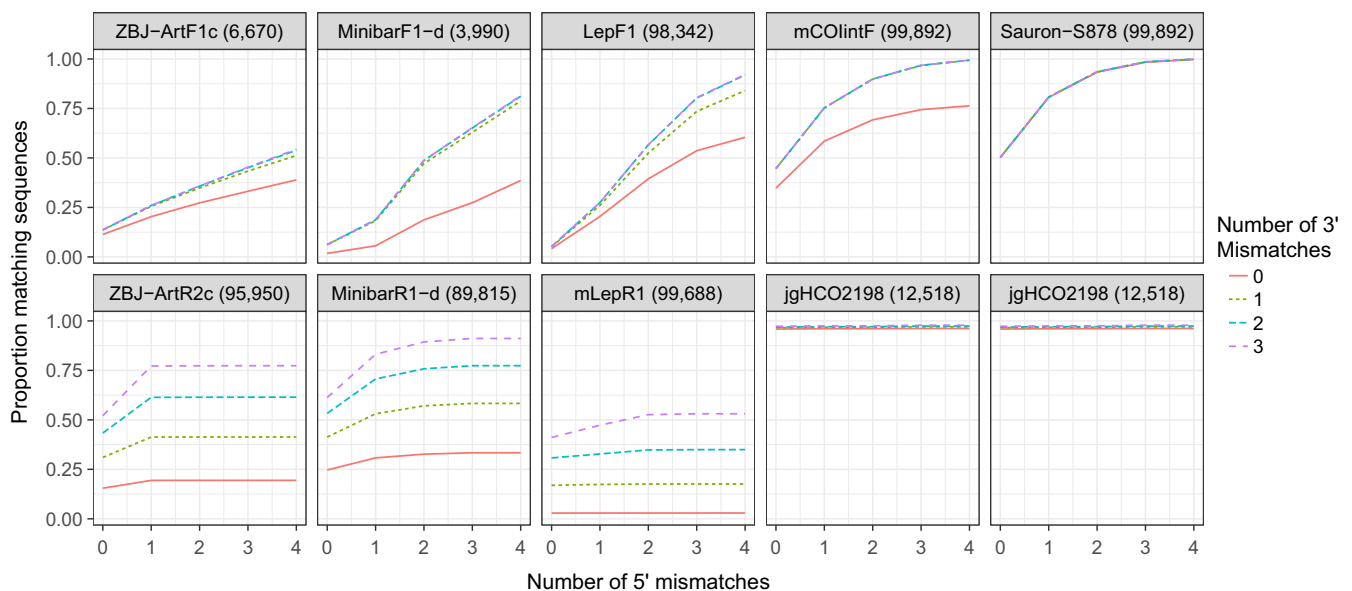


FIGURE 2 Overall proportion of NCBI sequences matching the 3' region (first four bases; line type) and 5' region (rest of the primer; x-axis) of general COI barcoding primers. Forward and reverse primers are on top of each other, and the number of NCBI sequences on which each primer was tested is shown in parentheses beside the primer name. The newly designed forward primer (Sauron-S878) is on the right-hand side and combined with jgHCO2198, as was the primer mCOLintF. Note that for Sauron-S878, mCOLintF, and jgHCO2198, there is almost no variability in fit for the 3' end and lines are therefore overlapping

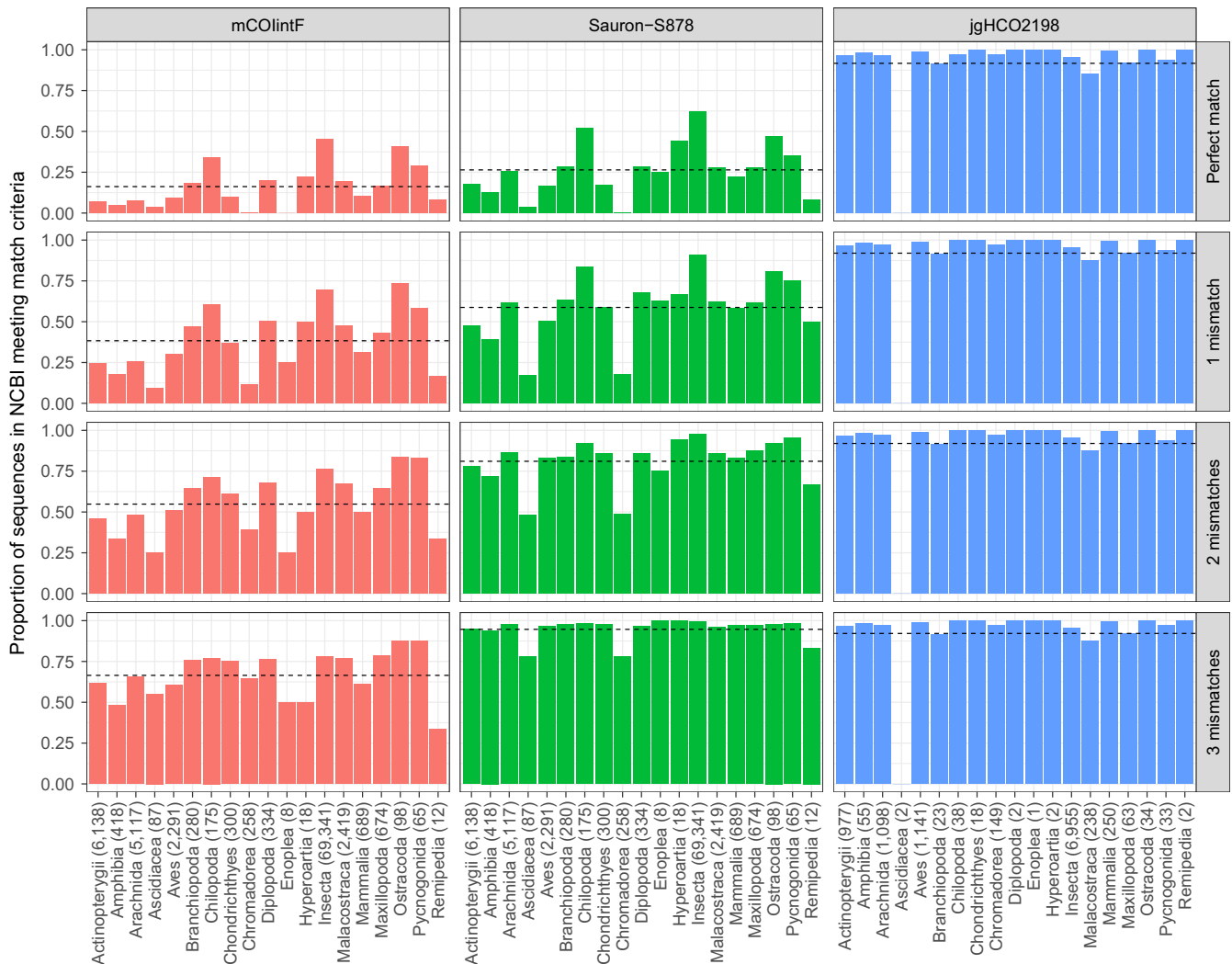


FIGURE 3 In silico testing. Proportion of NCBI sequences matching (full match in the first four bases of the 3' end and 0–3 mismatches in the remaining 5' region; figure rows) general COI primers (figure columns) for different classes of animals with the number of NCBI sequences covering the priming site in parentheses. The newly designed forward primer (Sauron-S878) is in the central column flanked by mCOLintF (left) and the reverse primer jgHCO2198 (right). Horizontal dashed lines show mean coverage of all NCBI sequences

mLepR1, mCOLintF/jgHCO2198, and Sauron-S878/jgHCO2198 primer combinations with a Nextera specific tail incorporated at the 5' end of each primer according to the manufacturer's recommendation (Table 1). Preparation of this library included a first amplification using Nextera tail-incorporated locus-specific primers, followed by a first clean up of unwanted primers and artifacts. After this, an additional re-amplification was performed to incorporate sequencing adapters and MID (for details on PCR conditions and protocols see Data S1).

For each library preparation method, amplicons from individual samples with incorporated MID and sequencing adapters were cleaned using SPRIselect magnetic beads (Beckman Coulter, Brea, CA, USA) according to the manufacturer's recommendation for left side size selection with a ratio of beads to PCR product volume of 0.8. After clean up, the remaining DNA was quantified using the QIAxcel (Qiagen) capillary electrophoresis system with the software ScreenGel v1.4 (Qiagen). Thereafter, all samples were pooled at

equimolar concentrations into ready to sequence libraries to ensure an even sequencing depth across samples. Ready-to-sequence libraries were submitted along with sequencing primers for paired-end sequencing on an Illumina MiSeq platform using the MiSeq Reagent Kit Nano v2 (300 bp; Illumina) at the Biomedical Sequencing Facility of the CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences and the Medical University of Vienna.

2.7 | Data analyses

Raw sequencing reads were demultiplexed, quality-checked, trimmed, and combined into paired-end reads using Usearch (Edgar, 2010). To remove all singleton sequences and reads shorter than 300 bp, these reads were then dereplicated using Usearch. After this, remaining reads were clustered into OTUs based on a 97% sequence similarity using Usearch. From among clusters, the centroid sequence was selected as a representative and a taxonomic ID was

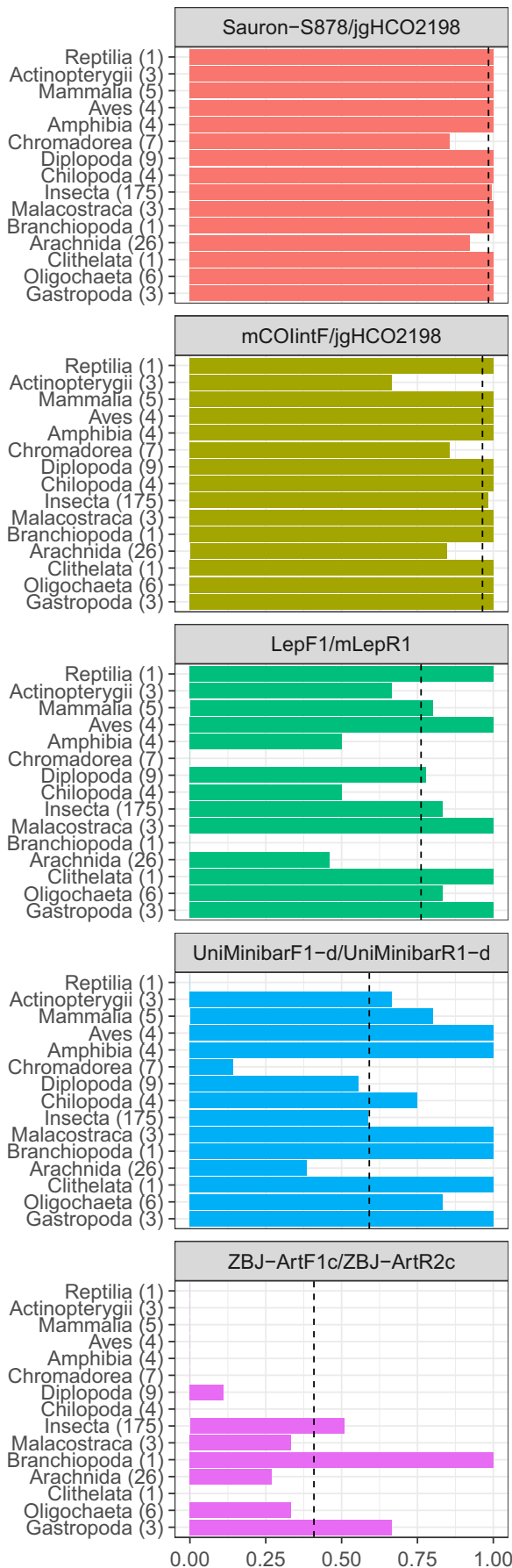


FIGURE 4 In vitro testing. Proportion of DNA extracts amplified from each taxonomic group (number of tested individuals in parentheses) with commonly used “universal” primer pairs ranked by high (top) to low (bottom) amplification success. The newly designed primer Sauron-S878 is here located at the top as the primer with the highest amplification success. Dashed vertical lines show mean detection success

assigned using blastn (Altschul, Gish, Miller, Myers, & Lipman, 1990) based on the NCBI Nucleotide database (Benson, Karsch-Mizrachi, Lipman, Ostell, & Wheeler, 2006) with a minimum ID threshold set to 90% and a word length of 28 bases. From received hits, the most likely identity was then selected based on E-score (with a maximum threshold of e^{-10}), match length, and on known information about the expected identity of the included taxa.

Because of the reduced overlap caused by the sequencing of the universal primers from Nextera libraries with the commercial sequencing primers (Box 1), as well as slightly different phred scores between runs (single-tube library, $27.3_{\text{mean}}, \pm 2.6_{\text{SD}}$; Nextera, $35.7_{\text{mean}}, \pm 1.9_{\text{SD}}$), pairing of reads from the Nextera sequencing run needed to be more restrictive. In order to make results comparable between the two libraries, we first analyzed the single-tube results for each replicate of the three balanced concentration mock samples using the same setting as for the Nextera library preparation method. Then, after reads from the two library preparation methods had been compared, the single-tube library reads were reprocessed by allowing more mismatches during pairing due to the increased overlap in these reads in order to increase the number of paired reads that passed quality filtering. The reason for doing this was here to make it easier to interpret whether any potential differences in libraries were due to the artifacts of the preparation method or due to a difference in overlap between forward and reverse reads between methods.

2.8 | Statistical analysis

All analyses were based on the presence or absence of in silico or in vitro amplification, and statistical tests were performed using generalized linear models fitted with binomial error distributions. For each model, diagnostic plots were examined to confirm that model assumptions were met (Zuur et al. 2010). NGS analyses were based on the presence or absence of detections, results from mock communities were combined as each mock community contained a different set of taxa, and replicates were used to test whether the successful detection of an extract varies between taxonomic classes. To model the probability of detecting an included DNA extract given the concentration of the DNA of this taxon in relationship to the concentration of other extracts, and how this probability is affected by the number of sequences attained (sequencing depth), values were predicted from binomial models fitted with measured data. All statistical tests were performed in R version 3.3.2 (R core team, 2016).

3 | RESULTS

3.1 | In silico

In silico evaluation of priming site variability showed that the newly designed forward primer Sauron-S878, in combination with the reverse primer jgHCO2198, performed better than all other primer pairs ($p < 0.001$). In fact, most of the tested metabarcoding primers exhibited considerable variation in both number and position of mismatches between primers and sequences. Among published primers, all but the reverse primer jgHCO2198 had a variable fit across taxonomic groups in the 3' end (Figure 2) where a good fit is necessary for amplification to take place. Under restrictive conditions (no mismatches allowed in the 3' end of the primer and three mismatches between the remainder of the primer and template), the best forward primer ($p < 0.001$) was our new primer Sauron-S878 that fit on 98.3% of NCBI sequences (Figure 3). The next best fitting forward primer mCOLintF fit only on 74.4% of the same sequences due to a mismatch at the 3' end. Taxonomic groups that were fitting poorly with mCOLintF were, for example, Ascidiacea (65.5%), but also common groups such as Arachnida (65.8%), nematodes (64.7%), and many vertebrates: Mammalia (61.0%),

Aves (60.3%), Amphibia (48.3%), and Actinopterygii (61.6%). For each of these groups, the new primer Sauron-S878 considerably increased coverage to 78.2% for Ascidiacea, 97.6% for Arachnida, 77.9% for nematodes, 97.4% for Mammalia, 96.7% for Aves, 94.0% for Amphibia, and 94.8% for Actinopterygii.

Among the reverse primers, only jgHCO2198 had an overall good fit of 96.2% and only performed poorly for Ascidiacea which were missed completely (although it should be noted that a reference sequence for the jgHCO2198 binding site was only available for two species within this group). With a more relaxed fit criterion (seven mismatching bases of which three are located in the first four bases of the 3' end), when the least well-fitting of the primers will not amplify and a considerable bias is expected, the best fitting of the published primers continued to be the primers mCOLintF and jgHCO2198. Furthermore, even if individual forward or reverse primers among other primers sometimes performed well, it should be noted that under these conditions, the least well-fitting primer of each primer pair (forward/reverse) only found a match on 81.4% (MinibarF1-d), 54.4% (ZBJ-ArtF1c), and 53.1% (mLepR1) of sequences.

		Single-tube Sauron-S878/jgHCO2198			Nextera LepF1/mLepR1			Nextera mCOLintF/jgHCO2198			Nextera Sauron-S878/jgHCO2198		
Unplaced	Chordata Mammalia	0.96	0.31	1.05				0.58	0.77	2.54	0.37	1.85	0.62
Unplaced	Arthropoda Insecta	40.19	58.31	66.49				69.88	71.68	11.02	81.16	8.33	79.06
Nematoda	Chromadorea Tylenchida (3)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0	0.93	0.21
Nematoda	Chromadorea Rhabditida (4)	0.48	0.31	0.26	0	0	0	0.29	0	0	0.19	0.93	0.21
Mollusca	Gastropoda Eupulmonata (3)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Annelida	Cilielata Arhynchobdellida (1)	0.96	0.94	0.52	0	0	0	0.58	0.51	1.69	0.37	1.85	0.41
Annelida	Cilielata Splanolaxida (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Chordata	Reptilia Squamata (1)	0.48	0	0	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Chordata	Mammalia Rodentia (1)	0	0	0	0	0	0	0	0	0	0	0	0
Chordata	Mammalia Eulipotyphla (1)	0	0	0	0	0	0	0	0	0	0	0	0
Chordata	Mammalia Chiroptera (3)	0.96	0.94	0.79	0	0	0	0.58	0.51	1.69	0.56	1.85	0.41
Chordata	Aves Passeriformes (2)	0.96	0.63	0.52	0	0	0	0.58	0.51	1.69	0.37	1.85	0.41
Chordata	Aves Galliformes (2)	0.96	0.63	0.52	0	0	0	0.58	0.51	1.69	0.37	1.85	0.41
Chordata	Amphibia Caudata (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Chordata	Amphibia Anura (3)	1.44	0.94	0.79	0	0	0	0.88	0.77	2.54	0.56	2.78	0.62
Chordata	Actinopterygii Salmoniformes (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Chordata	Actinopterygii Perciformes (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Chordata	Actinopterygii Cypriniformes (1)	0	0.31	0	0	0	0	0	0	0	0	0.93	0
Arthropoda	Crustacea Isopoda (2)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Malacostraca Decapoda (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Insecta Trichoptera (2)	0.48	0.31	0	0	0	0	0	0	0	0	0	0
Arthropoda	Insecta Thysanoptera (4)	0.48	0.31	0	0	0	0	0	0.26	0	0	0.93	0.21
Arthropoda	Insecta Raphidioptera (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Insecta Plecoptera (1)	0.48	0	0	0	0	0	0	0	0.85	0	0	0
Arthropoda	Insecta Orthoptera (5)	0.96	0.63	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Arthropoda	Insecta Odonata (4)	0.96	0.94	0.52	0	0	0	0.88	0.77	2.54	0.56	2.78	0.62
Arthropoda	Insecta Neuroptera (1)	0.48	0.31	0	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Arthropoda	Insecta Lepidoptera (7)	3.35	2.82	1.83	16.67	0	0	1.17	1.02	3.39	0.75	3.7	0.82
Arthropoda	Insecta Hymenoptera (35)	7.66	5.96	5.5	33.33	0	0	5.56	4.59	14.41	2.99	13.89	3.08
Arthropoda	Insecta Homoptera (6)	1.91	1.25	1.05	0	0	0	0.88	0.77	2.54	0.56	2.78	0.62
Arthropoda	Insecta Hemiptera (6)	0.96	0.94	0.52	0	0	0	0.58	0.51	1.69	0.37	1.85	0.41
Arthropoda	Insecta Ephemeroptera (1)	0.48	0.31	0.26	16.67	0	0	0.29	0.26	0.85	0.19	0	0.21
Arthropoda	Insecta Diptera (38)	9.09	6.58	5.5	0	0	0	4.09	4.34	10.17	2.61	13.89	2.67
Arthropoda	Insecta Dermaptera (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Arthropoda	Insecta Collembola (3)	0.96	0.63	0.52	0	0	0	0.58	0.51	1.69	0.37	1.85	0.41
Arthropoda	Insecta Coleoptera (61)	13.88	9.09	7.85	33.33	100	100	5.56	5.36	20.34	3.36	16.52	4.11
Arthropoda	Diplopoda Polydesmida (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Diplopoda Julida (6)	0.96	0.63	0.52	0	0	0	0.58	0.51	1.69	0.37	0.93	0.41
Arthropoda	Diplopoda Glomerida (1)	0.48	0	0	0	0	0	0	0	0	0	0	0.21
Arthropoda	Diplopoda sp (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Chilopoda Lithobiomorpha (2)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Arthropoda	Chilopoda Geophilomorpha (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Chilopoda sp (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Branchiopoda Diplostraca (1)	0.48	0.31	0.26	0	0	0	0.29	0.26	0.85	0.19	0.93	0.21
Arthropoda	Arachnida Trombiculiformes (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Arachnida Sarcoptiformes (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Arachnida Oribatidiformes (1)	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	Arachnida Opiliones (2)	0	0	0	0	0	0	0	0.26	0	0	0	0
Arthropoda	Arachnida Araneae (21)	4.78	3.13	2.36	0	0	0	2.34	2.55	6.78	1.87	5.56	1.64
Arthropoda	Arachnida Acari (1)	0	0	0	0	0	0	0	0	0	0	0	0

FIGURE 5 Percentage of retrieved samples from taxonomic orders from balanced, replicated mock communities (for simplification all three mock communities are pooled) by primer pair and library preparation method. Color is proportional to the relative number of retrieved sequences

FIGURE 6 (a) Percentage of retrieved sequences per sample from taxonomic orders (number of included samples in parentheses) from balanced, replicated mock samples (for simplification all three mock communities are pooled). Color is proportional to the relative number of retrieved sequences. For ease of visualization, samples have been grouped taxonomically, for a complete list of detected taxa, see Table S1 (Supporting Information). (b) Variability in the proportion of extracts identified with NGS from mock communities (y-axis) per taxonomic group (x-axis). Error bars represent 95% confidence intervals, and sample sizes are shown within parentheses

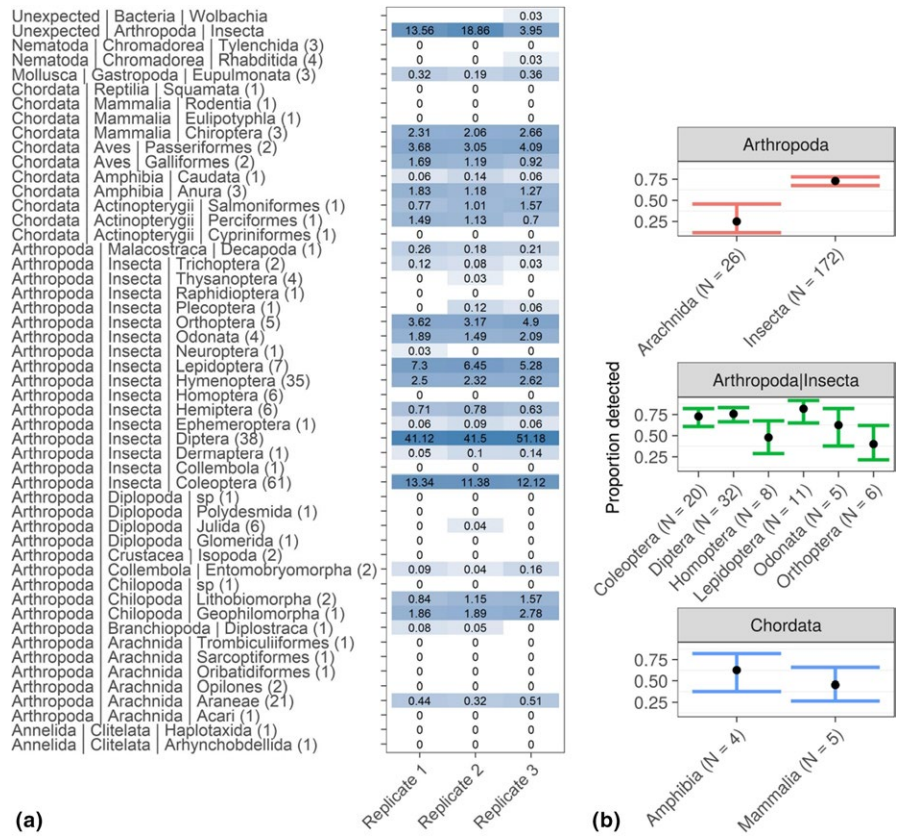
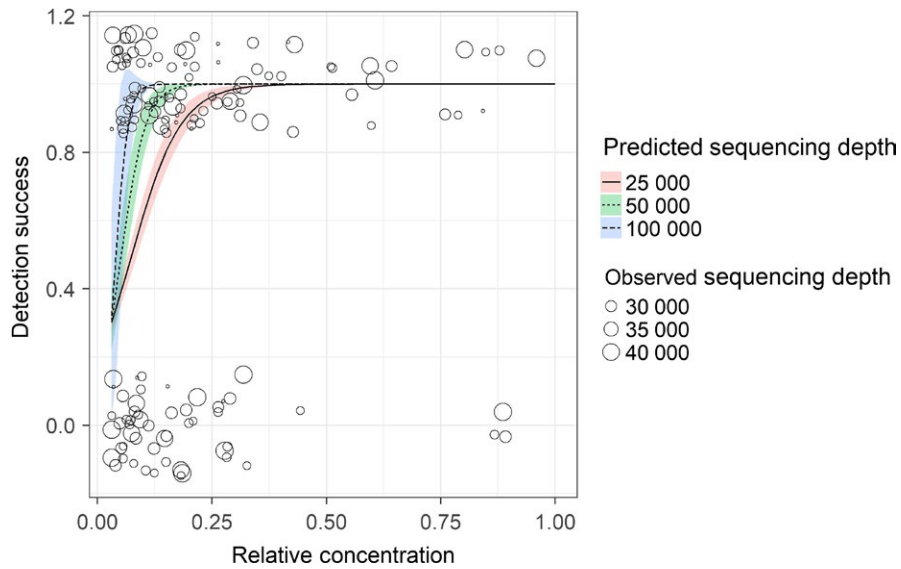


FIGURE 7 Proportion of extracts identified with NGS from mock communities (y-axis) modeled against the proportion of sample DNA concentration in relation to total DNA concentration (x-axis). Colors show the influence of different levels of sequencing depth and depict the predicted standard error. Observed data are here shown by circles proportional in size to sequencing depth



3.2 | In vitro

In in vitro testing, the newly designed primer Sauron-S878 combined with the reverse primer jgHCO2198 had an amplification success of 98.4%, and continued to perform significantly better than other primer pairs (Figure 4; $p < 0.001$), with the exception of mCOLintF/jgHCO2198 that also performed well with an amplification success of 96.4%. Of the other primer combinations, LepF1/mLepR1 amplified 76.2%, MinibarF1-d/MinibarR1-d

59.1%, and ZBJ-ArtF1c/ZBJ-ArtR2c 40.9% of the tested 255 taxa. While mCOLintF/jgHC2198 did have an overall good amplification efficiency, it did perform poorly for Arachnids (84.6%), where common groups such as Linyphiidae did not always amplify, for Actinopterygii (66.6%), and for nematodes (85.7%). While neither of these groups were statistically different from other groups amplified by the same primer, we can here show that by replacing the primer mCOLintF by the new primer Sauron-S878, the amplification success of these groups could be increased, to 92.3% for Arachnids,

100% for Actinopterygii, and only the amplification of nematodes remained unchanged at 85.7%.

3.3 | NGS results

From the library prepared using the developed single-tube PCR approach and custom sequencing primers, a total number of 607,528 raw paired-end reads were generated from a small run on the MiSeq platform. After preprocessing, these ranged from 27,507 to 55,094 reads distributed between balanced and unbalanced, replicated mock samples with a mean read number of 36,637. To test how the developed single-tube PCR approach and custom sequencing primers perform against the commercially available Nextera indexing kit (Illumina), a second small run was conducted on the same MiSeq platform. From this run, 546,099 raw paired-end reads were generated that, after preprocessing, ranged from 10,035 to 30,729 reads distributed between balanced, replicated mock samples and primers with a mean read number of 20,226. These reads were subsequently compared to the 277,778 raw paired-end reads generated from the balanced mock samples included in the first MiSeq run (single-tube PCR approach). After processing and matching of the sequences produced by the Nextera library to the NCBI Nucleotide database, the proportion of individual samples in each mock community that could be identified using each primer combination (Figure 5) was considerably lower for the LepF1/mLepR1 primer combination (2%, $p < 0.001$), than for mCOLintF/jgHCO2198 (42%), and Sauron-S878/jgHCO2198 (40%). From the Nextera library, samples amplified by the latter two primer combinations showed no differences. However, compared to the Nextera preparation method, the single-tube PCR approach in combination with the Sauron-S878/jgHCO2198 primers showed an increased recovery rate of included taxa of 50% ($p < 0.05$). Furthermore, it was also noticeable that in most replicates, the single-tube method produced less sequences from arthropods that were not included in mock samples than the Nextera method (unplaced arthropods, Figure 5). This suggests that the extra overlap between reads with the single-tube method may decrease read errors even though these still occur. This was also indicated when the sequence database was reduced to include only those species included in mock samples. This unreported test is biased as reads are forced to align to sequences, but nevertheless allowed the proportion of taxa recovered to be increased to ~68% with the single-tube method.

After reads from the two library preparation methods had been compared, we found that because of the increased overlap between the single-tube library reads, these could be better processed by allowing more mismatches during pairing of reads. Thus, we reanalyzed this library and again matched processed reads against the NCBI Nucleotide database. The proportion of individual samples in each mock community that could be identified among the samples included at balanced concentrations now varied between 87% and 67%, and in total 75% of the DNA extracts. Among the DNA extracts that did not produce sequences

were several underrepresented taxonomic orders (Figure 6a; for a detailed list of identified taxa, see Table S1). However, after correcting for false discovery rate, a significant bias could only be found among well-represented groups (where there were more than four samples for statistical relevance; Figure 6b) for arachnids ($p < 0.001$). It should be noted that among the included DNA extracts that produced no sequences, all missed arachnids had a low DNA concentration ($< 2 \mu\text{g}/\mu\text{l}$).

From unbalanced mock samples, we could show that these drop-outs may alternatively be attributable to an insufficient sequencing depth. Particularly, we could see that the lower the proportional DNA concentration of a species was within a sample, the more sequences needed to be read to reliably recover this species from within meta-samples. We modeled the likelihood of detecting individual sample sequences within each mixed sample, against the proportion of a sample's DNA concentration in relationship to the total DNA concentration (before PCR), and showed that the likelihood of detection of low concentration samples drops considerably, particularly at lower sequencing depth (Figure 7, $p < 0.01$, $R^2_{\text{McFadden}} = 0.25$). This relationship was additionally reflected in the DNA extracts with a concentration of $< 5 \mu\text{g}/\mu\text{l}$ that were detected at a lower rate ($p < 0.001$) in balanced mock communities (32%, $SE = 0.11$) than higher concentration extracts (69%, $SE = 0.14$) included at equimolar proportions.

4 | DISCUSSION

In this study, we showed that, while the concerns about using COI in metabarcoding studies have been growing (Clarke et al., 2014; Deagle et al., 2014; Leray et al., 2013), new primers such as Sauron-S878, mCOLintF, and jgHCO2198 make it possible to, even though species sometimes are missed, reliably amplify COI from a wide range of taxa. We did so by extensively testing these primers *in vitro* on DNA and *in silico* on sequences from a large set of taxa. In addition, primers were adapted and tested with a universal tail that, within a new single-tube PCR laboratory protocol, allows NGS-specific adapters to be incorporated in a single PCR where previously additional laboratory procedures were needed. As additional benefit, the effective read length and thus sequence overlap of paired-end reads can be extended with this approach compared to the commercial Nextera indexing kit. Using this methodology, we show how NGS libraries can be prepared as efficiently as with commercial kits but with a reduction in PCR-related biases from primer mismatches, handling as well as a reduction in the work normally needed to attach sequencing adapters and MIDs.

4.1 | Primer comparison

We demonstrated that many of the so far published primers do exhibit a large amount of bias across taxonomic groups during both *in silico* and *in vitro* testing which should be considered when selecting primers and interpreting results. Because of the variability in COI, the primers

may miss some groups completely and have poor amplification efficiency in other taxa. The best results during *in vitro* and *in silico* testing were obtained with the primer pair Sauron-S878/jgHCO2198 which matched and amplified over 98% of the evaluated taxa. The next best option was the combination mCOLintF/jgHCO2198 which performed also well during the *in vitro* testing, but matched only 74.4% of the taxa in NCBI due to a poor fit on important terrestrial taxa such as spiders, but also many vertebrates. By incorporating additional degeneracy into the primer Sauron-S878 in order to better cover terrestrial groups such as spiders, we also improved the fit for a large range of other taxa where a mismatch situation was occurring at the 3rd base from the 3' end. This improved fit suggests that in a competitive reaction, fewer taxa are likely to be amplified with low efficiency due to poor primer fit, reducing the overall primer bias.

All other primers failed to amplify large numbers of the tested DNA extracts: for example, ZBJ-ArtF1c/ZBJ-ArtR2c failed to amplify over half of the tested DNA extracts. Such biases are problematic, particularly if one is unaware of them. On the other hand, under certain circumstances, this can also be beneficial. For example, in the trophic studies on vertebrates, these primers mostly have been used for, one is usually not interested in amplifying consumer DNA. In such cases, primers that do not (or poorly) amplify consumer DNA might be advantageous (Deagle et al., 2009), although this means accepting a high degree of bias in the detection of prey taxa, especially as the primers ZBJ-ArtF1c/ZBJ-ArtR2c not only missed vertebrate DNA but also many taxa of arthropods.

4.2 | Single-tube protocol and NGS

To maximize taxonomic coverage, Sauron-S878 is a highly degenerate primer being in the end a mixture of 768 different primer combinations. To compensate for the potential problems caused by differing annealing temperatures between primers known to occur with such a high degree of degeneracy as in both Sauron-S878 and jgHCO2198, we incorporated universal tails into these primers (Green et al., 2015). The rationale here was that primer bias is likely to be reduced by conducting the majority of the exponential PCR cycles with these tails as identical templates across amplicons. We here, in addition, developed an approach that uses this tail to split a single PCR into two steps where (a) a first step of amplification is conducted using the universal primer pair at a higher annealing temperature to obtain higher specificity for the priming site and (b) a second step amplification is conducted using the universal tails as templates at lowered annealing temperature. This approach not only has the potential to reduce biases associated with high degrees of degeneracy, but can be used to cost-effectively include NGS adapters into amplicons for library preparation.

The developed single-tube PCR approach was used in combination with a set of custom sequencing primers and was confirmed to work well compared to the standard Illumina primers used with the Nextera library preparation method on a MiSeq run. An additional benefit of the single-tube PCR approach is that, unlike with the Nextera preparation method, this method uses the locus-specific

primers as part of the template for sequencing primers, meaning that these are not sequenced. This allows a longer insert length, which facilitates pairing of reads as the overlap between forward and reverse reads is extended by the length of each primer. For example, in our case, the extended read length achieved with the single-tube approach allowed considerably more taxa to be recovered from within the tested samples, as fewer reads had to be discarded due to poor quality scores.

To test for biases between taxonomic groups, we used the single-tube PCR approach to prepare a library of three replicated mock samples constructed from the DNA extracts used during primer testing. We did not detect any significant bias between the included taxonomic groups, even though not all of the included DNA extracts did produce sequences. It should, however, be noted that it had been confirmed that these DNA extracts did amplify with the primers using the developed protocol during *in vitro* testing. From the subset of the mock samples where initial extract concentrations had not been balanced for concentrations, we could see a likely explanation for this: The more uneven DNA concentrations are within a sample, the more sequences need to be read to reliably recover lower concentration samples from within mixed samples. More specifically, the probability of detecting a species will depend on both the relative concentration of the DNA of a given species in relationship to other species in the same sample as well as sequencing depth. This bias in combination with potential primer biases may lead to more or less random numbers of reads obtained for different taxa from an individual NGS run. Other studies have here come to similar conclusions, and it has repeatedly been stressed that when interpreting the results from metabarcoding studies, the detection of taxa should preferably be treated as the presence or absence as the number of returned sequences may be misleading (Elbrecht & Leese, 2015; Ficetola et al., 2015; Yu et al., 2012). Here, we would like to suggest that this argument may be overly simplified and suggest that even the presence/absence data may carry a considerable bias if the proportion of DNA types is uneven between mixed samples and sequencing depth is low as sampling may not be complete. Consequently, as the relative DNA concentration in most cases is unknown and not only affected by the amount of tissue from which the DNA is extracted, but also by the level of DNA degradation, tissue type and copy number (Deagle, Eveson, & Jarman, 2006; Thomas, Deagle, Eveson, Harsch, & Trites, 2015), care should be taken when deciding how many samples to pool in one run. The comparably low sequencing depth (10,000–30,000 quality-checked reads for each sample with 85 taxa to be identified) in this study was chosen, to mimic the suggestion to sequence hundreds or thousands of samples in one run making use of a nested tagging approach (Kitson et al., 2015). The idea behind this is that few thousand sequences per sample would be sufficient in obtaining the desired information and everything else would be lost investment. However, we could clearly show that with such a reduced sequencing depth, the risk to miss rare taxa increases dramatically. The basic assumption that NGS will describe all of the sequences that are in a sample is simply not true, but what is more likely to occur is that one describes the

most abundant DNA types in a sample unless sequencing depth is sufficiently high. As the abundance of DNA types in a sample is not necessarily related to the most abundant species in an environment, or diet, there is definitely a need for caution in interpreting metabarcoding results.

5 | CONCLUSIONS

The Sauron-S878 and jgHCO2198 primers enable researchers to continue employing COI for metabarcoding purposes without considerable primer biases in order to make use of the extensive available sequence databases. The developed protocols can reduce potential amplification biases produced by the (necessary) high degree of degeneration, such as preferential amplification depending on annealing temperature. Furthermore, the single-tube PCR approach for library preparation minimizes the contamination risk due to repeated handling and re-amplification. For metabarcoding results, however, we would argue that there are still additional methodological biases present that so far have not been well addressed but may strongly affect the completeness of results. Because of this, scientists need to consider carefully how much sequencing depth is needed for describing DNA from mixed samples. Mainly because the relative proportions of DNA from different species in such samples are usually not known but drastically affect detection probability, especially for DNA contained in low proportions.

Finally, it should be mentioned that most of the biases discussed in this paper are exponentially magnified by the current need to conduct PCRs before sequencing. Without exponential amplification in PCR, differences in detection probability are likely to be smaller. However, even though PCR induced biases will eventually be circumvented by PCR-free approaches (Creer et al., 2016; Denver et al., 2016; Paula et al., 2016), more abundant DNA will still be preferentially described from mixed samples.

ACKNOWLEDGMENTS

The work presented in this paper was funded by the project "Effects of fertilization type on biocontrol of pests" (Austrian Science Fund (FWF) P26144). We also acknowledge additional financial support from the BMFW (Sparkling Science SPA05/122) and through grants of the Mountain Agriculture Research Unit and the regional government of Tyrol, as well as a PhD scholarship provided by the University of Innsbruck to OR.

AUTHOR CONTRIBUTIONS

MT and DS obtained funding and together with OR conceived/designed the study. DS and NH performed laboratory work. OR conducted all analyzes of data, interpreted results, and compiled tables and figures. OR wrote the first draft of the manuscript, and MT and DS contributed to finalizing the paper. All authors gave final approval for publication.

DATA ACCESSIBILITY

The data generated for this study have been deposited in the Dryad Repository, <https://doi.org/10.5061/dryad.5jq5t36>.

ORCID

Oskar Rennstam Rubbmark  <http://orcid.org/0000-0003-3628-2348>

Daniela Sint  <http://orcid.org/0000-0003-3387-4284>

Michael Traugott  <http://orcid.org/0000-0001-9719-5059>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2006). GenBank. *Nucleic Acids Research*, 34(Suppl 1), D16–D20. <https://doi.org/10.1093/nar/gkj157>
- Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., ... de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, 29(6), 358–367. <https://doi.org/10.1016/j.tree.2014.04.003>
- Bru, D., Martin-Laurent, F., & Philippot, L. (2008). Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Applied and Environmental Microbiology*, 74(5), 1660–1663. <https://doi.org/10.1128/AEM.02403-07>
- CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 106(31), 12794–12797. <https://doi.org/10.1073/pnas.0905845106>
- Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, 7(3), 873–883. <https://doi.org/10.1002/ece3.2667>
- Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, 14(6), 1160–1170. <https://doi.org/10.1111/1755-0998.12265>
- Clooney, A. G., Fouhy, F., Sleator, R. D., O'Driscoll, A., Stanton, C., Cotter, P. D., & Claesson, M. J. (2016). Comparing apples and oranges? Next generation sequencing and its impact on microbiome analysis. *PLoS ONE*, 11(2), e0148028. <https://doi.org/10.1371/journal.pone.0148028>
- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., ... Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, 7(9), 1008–1018. <https://doi.org/10.1111/2041-210X.12574>
- De Barba, M., Waits, L. P., Genovesi, P., Randi, E., Chirichella, R., & Cetto, E. (2010). Comparing opportunistic and systematic sampling methods for non-invasive genetic monitoring of a small translocated brown bear population. *Journal of Applied Ecology*, 47(1), 172–181.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, 14(2), 306–323. <https://doi.org/10.1111/1755-0998.12188>
- Deagle, B. E., Eveson, J. P., & Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples—a case study on DNA in faeces. *Frontiers in Zoology*, 3(1), 11. <https://doi.org/10.1186/1742-9994-3-11>

- 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*(9), 20140562. <https://doi.org/10.1098/rsbl.2014.0562>
- Deagle, B. E., Kirkwood, R., & Jarman, S. N. (2009). Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, *18*(9), 2022–2038. <https://doi.org/10.1111/j.1365-294X.2009.04158.x>
- Denver, D. R., Brown, A. M., Howe, D. K., Peetz, A. B., & Zasada, I. A. (2016). Genome skimming: A rapid approach to gaining diverse biological insights into multicellular pathogens. *PLoS Pathogens*, *12*(8), e1005713. <https://doi.org/10.1371/journal.ppat.1005713>
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- 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*(7), e0130324. <https://doi.org/10.1371/journal.pone.0130324>
- 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. <https://doi.org/10.7717/peerj.1966>
- Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., ... Rayé, G. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*(3), 543–556. <https://doi.org/10.1111/1755-0998.12338>
- 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 c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, *13*(5), 851–861. <https://doi.org/10.1111/1755-0998.12138>
- Green, S. J., Venkatraman, R., & Naqib, A. (2015). Deconstructing the polymerase chain reaction: Understanding and correcting bias associated with primer degeneracies and primer-template mismatches. *PLoS ONE*, *10*(5), e0128122. <https://doi.org/10.1371/journal.pone.0128122>
- Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. (2006). DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(4), 968–971. <https://doi.org/10.1073/pnas.0510466103>
- Hebert, P. D., Cywinska, A., & Ball, S. L. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, *270*(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hebert, P. D., 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*(41), 14812–14817. <https://doi.org/10.1073/pnas.0406166101>
- Ivanova, N. V., Zemplak, T. S., Hanner, R. H., & Hebert, P. D. (2007). Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, *7*(4), 544–548.
- Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... Larsen, T. H. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, *16*(10), 1245–1257. <https://doi.org/10.1111/ele.12162>
- Jordaens, K., Sonet, G., Richet, R., Dupont, E., Braet, Y., & Desmyter, S. (2013). Identification of forensically important Sarcophaga species (Diptera: Sarcophagidae) using the mitochondrial COI gene. *International Journal of Legal Medicine*, *127*(2), 491–504.
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kitson, J. J. N., Hahn, C., Sands, R. J., Straw, N. A., Evans, D. M., & Lunt, D. H. (2015). Nested metabarcoding: A robust tool for studying species interactions in ecology and evolution. *BioRxiv*, 035071.
- Lefever, S., Pattyn, F., Hellemans, J., & Vandesompele, J. (2013). Single-nucleotide polymorphisms and other mismatches reduce performance of quantitative PCR assays. *Clinical Chemistry*, *59*, 1470–1480. <https://doi.org/10.1373/clinchem.2013.203653>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... 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. <https://doi.org/10.1186/1742-9994-10-34>
- Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, *5*, e3006.
- Pages, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2016). Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.38.4.
- Paula, D. P., Linard, B., Crampton-Platt, A., Srivathsan, A., Timmermans, M. J., Sujii, E. R., ... Volger, A. P. (2016). Uncovering trophic interactions in arthropod predators through DNA shotgun-sequencing of gut contents. *PLoS ONE*, *11*(9), e0161841. <https://doi.org/10.1371/journal.pone.0161841>
- Pompanon, F., Deagle, B. E., Symondson, W. O., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, *21*(8), 1931–1950. <https://doi.org/10.1111/j.1365-294x.2011.05403.x>
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., ... Miller, A. N. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, *109*(16), 6241–6246. <https://doi.org/10.1073/pnas.1117018109>
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., & Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene-sequences and a compilation of conserved polymerase chain-reaction primers. *Annals of the Entomological Society of America*, *87*, 651–701. <https://doi.org/10.1093/aesa/87.6.651>
- Stadhouders, R., Pas, S. D., Anber, J., Voermans, J., Mes, T. H., & Schutten, M. (2010). The effect of primer-template mismatches on the detection and quantification of nucleic acids using the 5' nuclease assay. *The Journal of Molecular Diagnostics*, *12*(1), 109–117. <https://doi.org/10.2353/jmoldx.2010.090035>
- Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology*, *21*(8), 1789–1793. <https://doi.org/10.1111/j.1365-294x.2012.05542.x>
- Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., & Trites, A. W. (2015). Quantitative DNA metabarcoding: Improved estimates of species proportional biomass using correction factors derived from control material. *Molecular Ecology Resources*, *16*, 714–726.
- Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA—an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, *183*, 4–18. <https://doi.org/10.1016/j.biocon.2014.11.019>
- Tringe, S. G., & Hugenholtz, P. (2008). A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology*, *11*(5), 442–446. <https://doi.org/10.1016/j.mib.2008.09.011>

- Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution*, 24(2), 110–117.
- Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623. <https://doi.org/10.1111/j.2041-210X.2012.00198.x>
- Zeale, M. R., Butlin, R. K., Barker, G. L., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular ecology resources*, 11(2), 236–244.
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in ecology and evolution*, 1(1), 3–14.

SUPPORTING INFORMATION

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

How to cite this article: Rennstam Rubbmark O, Sint D, Horngacher N, Traugott M. A broadly applicable COI primer pair and an efficient single-tube amplicon library preparation protocol for metabarcoding. *Ecol Evol*. 2018;8:12335–12350. <https://doi.org/10.1002/ece3.4520>