

# Metabarcoding for stomach-content analyses of Pygmy devil ray (*Mobula kuhlii* cf. *eregoodootenkee*): Comparing tissue and ethanol preservative-derived DNA

Matteo Barbato<sup>1</sup>  | Toby Kovacs<sup>1</sup> | Melinda A. Coleman<sup>2</sup>  | Matt K. Broadhurst<sup>3</sup>  | Mark de Bruyn<sup>1</sup> 

<sup>1</sup>School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia

<sup>2</sup>National Marine Science Centre, NSW Department of Primary Industries, Coffs Harbour, New South Wales, Australia

<sup>3</sup>Fisheries Conservation Technology Unit, NSW Department of Primary Industries, National Marine Science Centre, Coffs Harbour, New South Wales, Australia

## Correspondence

Matteo Barbato and Mark de Bruyn, The University of Sydney, School of Life and Environmental Sciences, Sydney, NSW, Australia.

Emails: matte.barbato@gmail.com; mark.debruyn@sydney.edu.au

## Abstract

The application of high-throughput sequencing to retrieve multi-taxon DNA from different substrates such as water, soil, and stomach contents has enabled species identification without prior knowledge of taxon compositions. Here we used three minibarcodes designed to target mitochondrial COI in plankton, 16S in fish, and 16S in crustaceans, to compare ethanol- and tissue-derived DNA extraction methodologies for metabarcoding. The stomach contents of pygmy devilrays (*Mobula kuhlii* cf. *eregoodootenkee*) were used to test whether ethanol-derived DNA would provide a suitable substrate for metabarcoding. The DNA barcoding assays indicated that tissue-derived operational taxonomic units (OTUs) were greater compared to those from extractions performed directly on the ethanol preservative. Tissue-derived DNA extraction is therefore recommended for broader taxonomic coverage. Metabarcoding applications should consider including the following: (i) multiple barcodes, both taxon specific (e.g., 12S or 16S) and more universal (e.g., COI or 18S) to overcome bias and taxon misidentification and (ii) PCR inhibitor removal steps that will likely enhance amplification yields. However, where tissue is limited or no longer available, but the ethanol-preservative medium is still available, metabarcoding directly from ethanol does recover the majority of common OTUs, suggesting the ethanol-retrieval method could be applicable for dietary studies. Metabarcoding directly from preservative ethanol may also be useful where tissue samples are limited or highly valued; bulk samples are collected, such as for rapid species inventories; or mixed-voucher sampling is conducted (e.g., for plankton, insects, and crustaceans).

## KEYWORDS

DNA extraction, DNA metabarcoding, Illumina sequencing, marine predator, stomach-content

## 1 | INTRODUCTION

High-throughput sequencing platforms enable generation of accurate and cost-effective multispecies genetic assays (Mardis, 2008). Deoxyribonucleic acid (DNA) barcoding is a methodology that can provide precise and semi-automatable species identification through the design of forward–reverse primer sets for highly conserved regions of mitochondrial (mt) DNA (Hebert et al., 2003). Combining DNA barcoding and high-throughput sequencing has delivered a promising tool—DNA metabarcoding—to detect biodiversity in DNA extracted from materials including water (Stat et al., 2017), air (Kraaijeveld et al., 2015), soil (Drummond et al., 2015), fecal (Berry et al., 2017), and stomach-content samples (Berry et al., 2015).

In the last decade, DNA metabarcoding of fecal matter and stomach contents has been developed, with accurate taxonomic resolution of dietary biodiversity, in attempts to infer trophic interactions among both terrestrial (Bohmann et al., 2011; Clare, Fraser, Braid, Fenton, & Hebert, 2009) and aquatic organisms (Berry et al., 2015). A fundamental step in this kind of study is DNA extraction, usually from tissue or fecal samples. However, Shokralla, Singer, and Hajibabaei (2010) proposed a novel DNA extraction methodology; successfully Sanger sequencing a universal COI barcode directly from ethanol-derived DNA extracts. The ethanol used for sample preservation was hypothesized to be an adequate DNA carrier, thus through low temperature evaporation of ethanol and subsequent re-suspension of the DNA pellet, the target organism's mtDNA was polymerase chain reaction (PCR) amplified and Sanger sequenced.

More recently, the same ethanol DNA extraction protocol was tested for utility in retrieving trace DNA from freshwater benthic larval communities (Hajibabaei, Spall, Shokralla, & Konynenburg, 2012) using both Sanger and 454 sequencing. Only individuals present at very low abundance (i.e., 1 individual) were not ascertained via these DNA barcoding assays. In contrast, Robertson, Minich, Bowman, and Morin (2013) compared DNA extracted from cetacean tissue of Short-beaked common dolphin (*Delphinus delphis*), Long-beaked common dolphin (*D. capensis*), Dall's porpoise (*Phocoenoides dalli*), Killer whale (*Orcinus orca*), Humpback whale (*Megaptera novaeangliae*), and False killer whale (*Pseudorca crassidens*). These samples were preserved for 3 to 20 years in two liquid preservatives, ethanol and 20% salt saturated dimethyl sulfoxide (DMSO). The authors were unable to detect either mtDNA or nuclear DNA from ethanol samples, while only mtDNA was amplified successfully from DMSO and Sanger sequenced.

Here, we propose that for stomach-content analyses, DNA extracted directly from preservative ethanol could provide an accurate representation of prey diversity in a more cost-effective and efficient manner than DNA extraction from tissue. Moreover, DNA metabarcoding may alleviate the impediment of identifying all prey DNA contained in stomach remains, if liquefied remains can be efficiently filtered in an unbiased manner. However, extracting DNA directly from stomach contents is not straightforward, because mixed-taxon stomach samples can be several

kilograms in weight, particularly in apex—(e.g., great white shark (*Carcharodon carcharias*)) and mesopredators (e.g., blacktip shark (*Carcharhinus limbatus*), *Thunnus* spp.) (Pompanon et al., 2012). It can therefore be difficult to representatively subsample small amounts of (sometimes digested) tissue to recover the full diversity of ingested prey. Nonetheless, it is important to note that digestion rate plays a crucial role in the detection of tissue-dependent extractions, sometimes causing an underrepresentation of those taxa where tissue is rapidly digested (Sousa et al., 2016). In this study, we compared the ethanol DNA extraction method (ETH) with a tissue extraction method (TIS) on the stomach-content samples of Pygmy devilrays (*Mobula kuhlii* cf. *eregoodootenkee*, Müller & Henle 1841) (White et al., 2017), using three mtDNA minibarcodes: teleost (16S), crustaceans (16S), and plankton (COI). Despite recent debate over the taxonomy of *M. kuhlii* cf. *eregoodootenkee* (White et al., 2017), the conservation status of *M. eregoodootenkee* and *M. kuhlii* are, respectively, near threatened and data deficient (IUCN 2018). Knowledge regarding habitat type is also inaccurate since the former is described as an oceanic species, and *M. kuhlii* as neritic (IUCN 2018). However, these mobulids share the same distribution across the Indo-West Pacific Ocean. Dietary habits of both species have been scarcely described in both species, but it is generally accepted that planktonic crustaceans and possibly small fishes and cephalopods are the main prey items (Couturier et al., 2012). The rare recovery of these rays allowed us to test a novel methodology for diet research in these taxa and to fill a knowledge gap for this understudied mobulid. Since stomach-content studies preclude a priori knowledge of the target species, we used operational taxonomic unit (OTU) analysis to test our hypothesis of no differences in recovered predator–prey relationships via metabarcoding between ethanol-preserved and tissue-derived DNA.

## 2 | MATERIAL AND METHODS

### 2.1 | Sample collection

Between January and May 2017, 31 specimens of pygmy devil rays were identified and collected by the New South Wales Department of Primary Industries (NSW DPI) from five bather protection nets deployed off northern New South Wales, Australia. After recovery, the individual specimens were transported by vehicle at ambient temperature to NSW DPI Center in Ballina (NSW, Australia) and immediately frozen at  $-20^{\circ}\text{C}$  to preserve the specimens, prior to bulk assessment of all animals, where the animals were collectively defrosted and necropsied following Broadhurst, Laglbauer, Burgess, and Coleman (2018). From visual inspection of stomach contents, it was possible to distinguish small fragments of prey remains, some of which were identified as sandy sprat (*Hyperlophus vittatus*). The stomach contents of 27 of the 31 specimens, which were relatively undegraded, were preserved in 50 ml of 100% ethanol to avoid further degradation of the samples.

## 2.2 | DNA extraction

In September 2017, 1.5 ml aliquots of ethanol were taken from each of 10 pygmy devil ray stomach-content samples without disturbing the settled tissue matter. For each sample, five technical replicates were taken to increase the final amount of DNA. These samples, along with five negative controls, were heated at 56°C until the ethanol had completely evaporated (Shokralla et al., 2010). Microbiology grade DNase-free water was added, and the samples were left at ambient temperature overnight. The samples were then vortexed to encourage resuspension, and technical replicates were combined. Tissue was also taken from the same 10 samples for comparison. We attempted to subsample the tissue as representatively as possible, particularly subsampling larger tissue pieces in an attempt to avoid swamping the sample with a single taxon. The DNA was extracted using a Qiagen DNeasy blood and tissue kit (Qiagen, Germany) following the manufacturer's instructions (Qiagen, 2006), with 10 min incubation after addition of elution buffer, instead of 1 min, to ensure adequate elution. The DNA extraction was carried out in a pre-PCR laboratory to minimize the risk of contamination. Filter tips were used, and negative controls were included in all stages of the laboratory workflow. Clean room protocols were followed, with extensive bleaching of the work areas and UV-treatment of equipment, wherever possible.

## 2.3 | Positive controls

Positive controls are a vital inclusion in any metabarcoding PCR because they allow distinction between problems regarding PCR conditions and sample DNA extractions, and can also be used to troubleshoot and calibrate downstream metabarcoding issues (Deiner et al., 2017). Here, positive controls were sampled for DNA on a subsequent day to the stomach-content samples to avoid cross-contamination and comprised the following: white banana prawn (*Penaeus merguensis*, de Man 1888), brown tiger prawn (*Penaeus esculentus*, Haswell 1879), deepwater flathead (*Platycephalus conatus*, Waite & McCulloch 1915), and eastern school whiting (*Sillago flindersi*, McKay 1985). These specimens were chosen because they were likely to amplify using the primers used in this study and were accessible from a local fish market. Sections of tissue were removed and the same DNeasy tissue extraction kit was used. Positive controls were

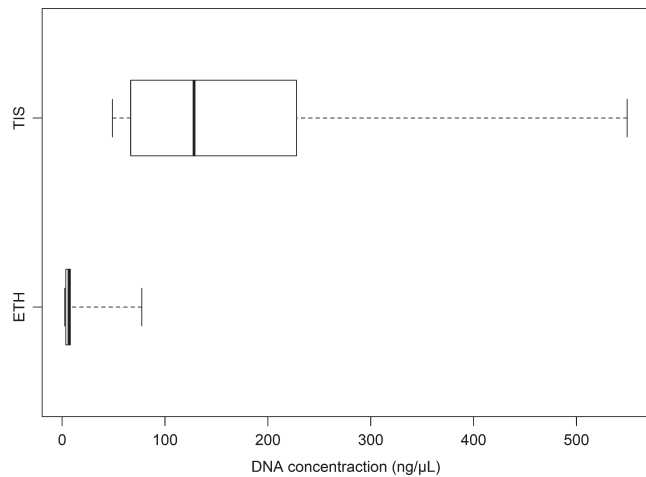
tested using routine universal primers HCO2198 and LCO1490 for prawns (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994), and Fish F1 and Fish R2 for fish (Ward, Zemlak, Innes, Last, & Hebert, 2005) to confirm the extraction protocols and PCR assays were working.

## 2.4 | Metabarcoding assay

Three previously designed and tested group-specific barcode primers were selected for teleosts, crustaceans, and plankton, targeting 16S mtDNA (teleost and crustacean) and the cytochrome c oxidase subunit I gene (plankton) (Table 1). The DNA extract concentrations were measured using a NanoDrop. The ratio of 260 and 280 nm absorbance was recorded to examine protein inhibition in the DNA extracts (Zarzoso-Lacoste et al., 2013). The PCR was performed following the protocol recommended by Taberlet, Bonin, Zinger, and Coissac (2018) using 10 µl of AmpliTaq Gold 360 DNA Master Mix (Thermo Fisher Scientific, USA), 2 µl of forward and reverse primer, 0.16 µl bovine serum albumin to decrease the risk of PCR inhibition, 2 µl of genomic DNA, and 4 µl of DNA-free water. After activation at 96°C for 10 min, each PCR run consisted of 35 cycles: 30 s at 96°C denaturation, 30 s at 50°C hybridization, and 1 min at 72°C elongation (Taberlet et al., 2018). The PCR products were tested on a 2% Agarose gel to confirm the successful amplification of the target amplicon of correct size, while subsequent PCR assays were undertaken using each primer with the addition of the appropriate Illumina overhang adaptors (forward overhang 5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG [sequence-specific primer] 3' and reverse overhang 5'GTCTCGTGGGCTCGGAG ATGTGTATAAGAGACAG- [sequence-specific primer] 3') using the same PCR conditions as above. Two PCR replicates for each primer set were conducted for each sample. The PCR products were sent to the Ramaciotti Centre for Genomics at The University of New South Wales (Sydney, Australia) for PCR cleanup using AMPure XP beads and Index PCR. This process involves attaching Illumina unique MID (Multiplex IDentifiers) to the amplicons of each sample, using primers that bind to the Illumina overhang adaptors. After two repetitions of PCR clean-up, the libraries were quantified, normalized, and pooled before paired-end sequencing was performed on an Illumina MiSeq platform. Because single step (indexing) PCR has been shown to significantly bias sequence abundance, we chose to use the Illumina recommended two-step procedure for metabarcoding (O'Donnell, Kelly, Lowell, & Port, 2016).

**TABLE 1** Minibarcode assay, Primer sequence, and PCR conditions

| Barcode | Target taxa | Primer             | Sequence                         | Amplicon length (bp) | T <sub>m</sub> (°C) | Reference            |
|---------|-------------|--------------------|----------------------------------|----------------------|---------------------|----------------------|
| 16S     | Fish        | Fish 16S Forward/d | 5' GACCCTATGGAGCTTTAGAC 3'       | ~200                 | 54                  | Berry et al. (2017)  |
|         |             | 16S reverse/d      | 5' CGCTGTTATCCCTADRGTAACT 3'     |                      |                     | Deagle et al. (2007) |
| COI     | Plankton    | Minibar-Mod-F      | 5' TCCACTAATCACAAAGAYATYGGYAC 3' | ~127                 | 52                  | Berry et al. (2015)  |
|         |             | Minibar-Mod-R      | 5' AGAAAATCATAATRAANGCRTGNGC 3'  |                      |                     |                      |
| 16S     | Crustacean  | Crust16S_F(short)  | 5' GGGACGATAAGACCCTATA 3'        | ~170                 | 51                  | Berry et al. (2017)  |
|         |             | Crust16S_R(short)  | 5' ATTACGCTGTTATCCCTAAAG 3'      |                      |                     |                      |



**FIGURE 1** DNA concentration (NANODROP) after extraction ( $t$ -value = 3.47;  $p$ -value = 0.007). Bars show the range of DNA concentration

## 2.5 | Bioinformatic pipeline

All Illumina reads went through quality filtering comprising two main steps. The first involved pairing, merging, and trimming off the forward–reverse primers using Geneious software (Kearse et al. 2012). Reads were discarded if primers were not present, not exactly matching the primer nucleotide lengths, or if reads exceeded the mismatch number of ambiguities according to the primer sequences. The second quality-filtering step was executed through USEARCH. Reads were removed if containing  $\geq 1$  ambiguities, maximum error above 0.5, were less than 50 bp in size or exceeding the expected amplicon length (Table 1). Sequences were dereplicated into groups of unique sequences. Singleton groups were discarded (Edgar, 2010).

Following USEARCH OTU pipeline recommendations, we conducted an OTU analysis that consisted of clustering sequences with 97% similarity through the execution of the UPARSE algorithm (Edgar 2010). Moreover, this process removed any sequencing errors, PCR artefacts, chimeras, and low-abundance clusters  $< 0.75\%$  of the total number of unique sequences identified within the sample. The OTUs from ETH and TIS extractions were inspected to identify common OTUs using the USEARCH *search\_exact* and *search\_global* function, respectively, with 100 or 80% identity thresholds. All ETH and TIS reads were separately mapped using the USEARCH command *otutab* that allowed visualizing the number of reads for each OTU within each sample.

Subsequently, BLASTn (Basic Local Alignment Search Tool, National Center for Biotechnology Information's (NCBI) GenBank nucleotide database, Altschul, Gish, Miller, Myers, & Lipman, 1990) was performed for searching OTUs with default parameters and a reward of 1 through the use of the high-performance computer facility Artemis at The University of Sydney. Outputs were imported into MEGAN6 (MEtaGenome ANalyzer; Huson et al., 2016) to inspect taxonomic identification using the LCA parameter set as a minimum bit score of 150.0 and the top 5% matches. Only fully matching queries were considered in the assignment of taxa. Additional

information on identified taxa was obtained from the Atlas of Living Australia (ATLAS) (2018). In all cases, when OTUs were examined further for taxonomy, they made biological (potential pygmy devilray prey items) and geographical (Australian east coast marine distribution) sense. In the case of the 16S fish assay, the main prey item identified by metabarcoding was also the dominant prey item in visual inspections of gut contents (sandy sprat, *Hyperlophus vittatus*). Particular attention was paid to exogenous contaminations, absent in the controls but present in at least one sample. Specifically, if the distribution of a suspicious taxon, examined by ATLAS, was not present in the area of study or it was impossible to be pygmy devilray prey that OTU was discarded.

## 2.6 | Statistical analysis

R Studio was utilized for all statistical analysis (RStudio Team, 2015). Data normality was evaluated by Shapiro (Shapiro and Wilk 1965) and Levene (Levene, 1960) tests on DNA concentration and absorbance ratio data. Both were normally distributed (Shapiro test  $p$ -value  $< 0.05$  and Levene test  $p$  value  $> 0.05$ ) so a paired  $t$ -test was executed. The normality of common OTU reads was evaluated in the same manner. As these values were not normally distributed (Shapiro test  $p$  value  $> 0.05$ ), the non-parametric Wilcoxon test (Wilcoxon, 1945) was executed in order to inspect the similarity of the read distribution of common OTUs between the TIS and ETH methods for the three minibarcodes (16S fish, 16S crustacean and COI plankton), using R Studio.

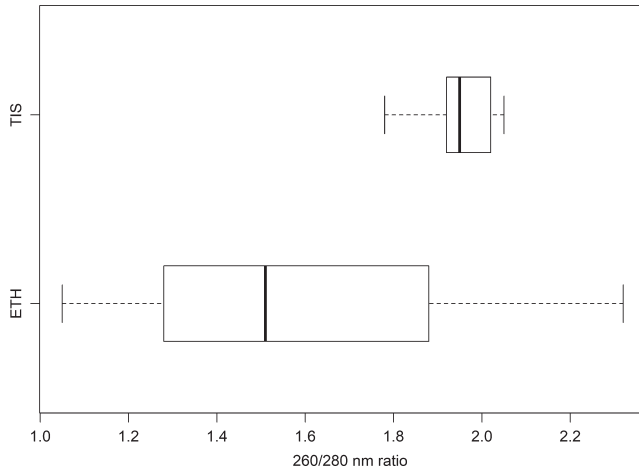
## 3 | RESULTS

### 3.1 | DNA extraction

The initial DNA concentrations derived from TIS were significantly greater than those for the ETH ( $t$ -value = 3.47;  $p$ -value = 0.007; Figure 1). Based on the 260/280 absorbance ratio (Figure 2), TIS DNA measurements ranged between 1.78 and 2.00, showing low protein inhibition whereas ETH-retrieved DNA had significantly greater inhibition levels (1.05–2.50) ( $t$ -value = 2.56;  $p$ -value = 0.03).

### 3.2 | Metabarcoding results and Operational Taxonomic Unit analysis

Executing BLASTn helped to identify exogenous contaminations and non-target taxa such as Amniota and Bacteria, which were confidently excluded from further analyses. The negative controls had low levels of endogenous contamination, and those OTUs were removed from both sample sets. The positive controls accurately identified the target species (white and brown banana prawns, deep-water flathead, and eastern school whiting), as well as other market fish (tuna, salmon, and conger eel for fish 16S), illustrating the universal fish primer efficiency and the sensitivity of the metabarcoding approach. The positive controls were purchased from a fish market with multiple species alongside one another, reflecting the



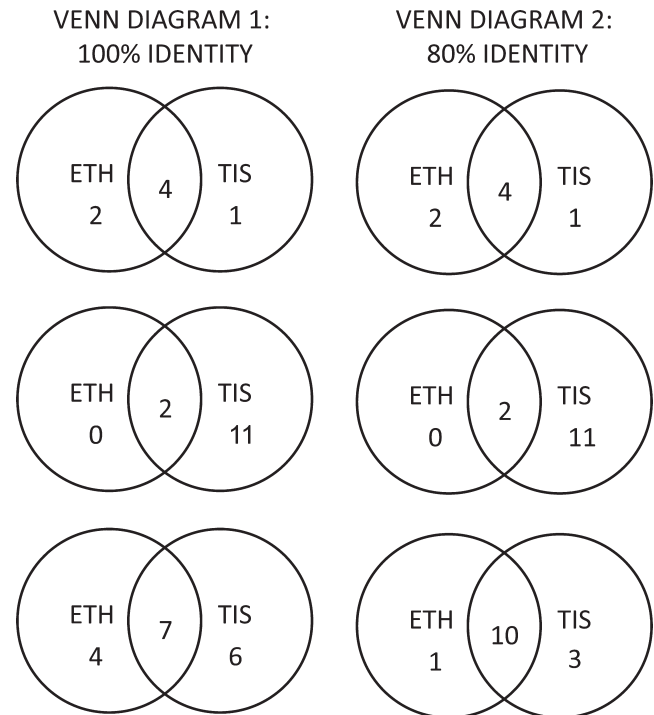
**FIGURE 2** Range of inhibition according to 260/280 nm ratio ( $t$ -value = 2.56;  $p$ -value = 0.03). Here values of 1.8 to 2.0 indicate little inhibition

sensitivity of the metabarcoding approach in detecting these likely contaminants.

The OTU analysis detected 11 and 13 OTUs in plankton COI, 6 and 5 OTUs in the fish 16S assay, and 2 and 13 OTUs in the crustacean 16S assay, for ETH and TIS, respectively. The distribution of reads for common OTUs across samples was similar between ETH and TIS (Table 3). The TIS method retained the greatest number of OTUs in plankton COI and crustacean 16S (Figure 3). In contrast, for the fish 16S assay, ETH showed one more OTU than TIS. In addition, we tested an 80% identity threshold to explore whether a lower similarity could give the same number of common OTUs. Our reasoning was that “false” OTUs could be generated when sequence lengths differed, or when high levels of intraspecific genetic diversity were present, or through sequencing error or other such biases. Although the 80% cut-off contributed to a larger number of common OTUs than for the 97% cut-off, it still indicated that TIS recovered more OTUs than the ETH (Figure 3).

## 4 | DISCUSSION

While DNA metabarcoding is not exempt from technical uncertainties such as PCR-bias and sequencing errors—which should not be underestimated (Bohmann et al., 2011; Berry et al. 2017; Albaina et al., 2016)—this DNA metabarcoding methodology can provide an improved analysis of feeding habits compared to visual inspection, in line with other studies (Table 2). However, this methodology alleviates but does not resolve the underrepresentation issue regarding prey tissues that have different digestion rates. To the best of our knowledge, however, no studies have tested ethanol, used to preserve stomach-content samples, as a suitable DNA carrier in the application of metabarcoding for dietary analysis. Our results show that it is possible to extract, amplify, and detect DNA by metabarcoding directly from ethanol in line with previous genetic applications such as biomonitoring surveys (Hajibabaei, Shokralla, Zhou, Singer,



**FIGURE 3** OTUs matched using USEARCH function search\_exact and search\_global. From top to bottom, 16S Fish, 16S crustacean and COI Plankton. Ethanol preservative-derived OTUs (ETH) and tissue-derived OTUs (TIS)

& Baird, 2011; Hajibabaei et al., 2012) and DNA presence–absence using Sanger sequencing (Shokralla et al., 2010). Nevertheless, when comparing OTU richness between the two methods, ethanol does not appear to recover the full range of prey. Although it was possible to detect common OTUs in both the ETH and TIS, and overall the read distributions of common OTUs were statistically equivalent for both methods ( $p$  value >0.05, Table 3), the TIS recovers more unique OTUs for plankton COI and crustacean 16S. Thus, ethanol-derived DNA may not inform on the entire spectrum of feeding habits of a species and could potentially be misleading when rare or low biomass prey are consumed.

In this study, we used a standard DNA extraction kit for the TIS samples, foregoing any attempts to remove inhibition (besides using BSA in the PCR assays) from the samples such as qPCR and serial dilutions, to preclude biasing the downstream analyses. Even so, the TIS extractions were less inhibited than the ETH indicated by the 260/280 absorbance ratio that reflects protein inhibition of samples (Figure 2). In a study of DNA extraction performance in stomach samples, Zarzoso-Lacoste et al., 2013) confirmed a better amplification yield with low DNA concentrations (8.2–71 ng/ml) and low protein inhibition (the 260/280 ratio being between 1.80 and 2.00), than high DNA concentrations (70.9–408 ng/ml) with absorbance ratios below 1.80, measured by Nanodrop.

The performance of the three minibarcodes was generally successful, and a multiple barcode approach can improve the outcome of metabarcoding studies (Berry et al., 2017). However, COI has been debated to not be particularly suited for short amplicon-based

**TABLE 2** Methodology summary of metabarcoding studies on marine organism diets using stomach-content samples

| Study   | Organism   | DNA extraction kit  | Inhibition removal               | qPCR   | Sequencer                |
|---|--|---|----------------------------------|--|--------------------------|
| Albaina et al. (2016)   | European sardine ( <i>Sardina pilchardus</i> ) and European sprat ( <i>Sprattus sprattus</i> )   | Salt extraction protocol  | Addition of bovine serum albumin | As proxy of prey abundance and sensitivity of 18S v9 assay | Pair-end Illumina MiSeq  |
| Berry et al. (2015)   | Common jack mackerel ( <i>Trachurus declivis</i> ), gemfish ( <i>Rexea solandri</i> ), tiger flathead ( <i>Platycephalus richardsoni</i> ), reef ocean perch ( <i>Helicolenus percoides</i> ), jackass morwong ( <i>Nemadactylus macropterus</i> ), Pink ling ( <i>Genypterus blacodes</i> ), blue warehou ( <i>Seriolella brama</i> ) and John dory ( <i>Zeus faber</i> ) | Qiagen DNeasy blood and tissue kit primer                                       | No                               | Yes, to measure concentrations of starting DNA template    | 454 GS Junior (Roche)    |
| Fernández-Álvarez, Machordom, García-Jiménez, Salinas-Zavala, & Villanueva (2018) | Flying squids (family Ommastrephidae)  | QIAamp DNA Investigator Kit   | Included in the extraction kit   | no   | Pair-end MiSeq Illumina  |
| Harms-Tuohy et al. (2016)   | Lionfish ( <i>Pterois volitans</i> )   | Qiagen DNeasy Blood and Tissue kit  | No                               | No   | Illumina MiSeq platform  |
| Leray et al. (2013)   | Cardinal fish ( <i>Nectamia savayensis</i> ), soldierfish ( <i>Myripristis berndti</i> ) and squirrelfish ( <i>Sargocentron microstoma</i> )   | Qiagen DNeasy Blood and Tissue kit  | MOBIO PowerClean DNA cleanup kit | No   | Roche 454 FLX sequencing |
| Jakubavičiūtė-Bergström, Eklöf, Haanel, & Bourlat (2017)                          | Three-spined stickleback ( <i>Gasterosteus aculeatus</i> )   | UltraClean <sup>®</sup> Tissue and Cells DNA Isolation Kit (MOBIO Laboratories) | Included in extraction kit       | No   | Pair-end Illumina Mi seq |
| Siegenthaler, Wangenstein, Benvenuto, Campos, & Mariani (2018)                    | Brown shrimp ( <i>Crangon crangon</i> )  | PowerSoil <sup>®</sup> DNA Isolation Kit (MoBio laboratories)                   | Included in extraction kit       | No   | Pair-end Illumina 2x250  |
| Sousa et al. (2016)   | Sun fish ( <i>Mola Mola</i> )  | JetQuick DNA Kit (GENOMED)  | No                               | No   | Roche 454 GS FLX         |
| Yoon et al. (2017)  | Antarctic toothfish ( <i>Dissostichus mawsoni</i> )  | AccuPrep <sup>®</sup> genomic DNA extraction kit                                | Included in extraction kit       | No   | Pair-end Illumina Miseq  |

| Sample | 16S Fish |         | 16S Crus |         | COI Plan |             |
|--------|----------|---------|----------|---------|----------|-------------|
|        | V value  | p value | V value  | p value | V value  | p value     |
| 1      | 3        | 0.37    | 0        | 1       | 3        | 1           |
| 2      | 0        | 0.18    | 0        | 0.5     | 4        | <b>0.04</b> |
| 3      | 6.5      | 0.71    | 0        | 1       | 0        | 0.1         |
| 4      | 0        | 0.37    | 0        | 0.5     | 5        | 1           |
| 5      | 1        | 1       | 0        | 1       | 1.5      | 1           |
| 6      | 3        | 0.37    | 0        | 0.5     | 5        | 1           |
| 7      | 1        | 1       | 1        | 1       | 0        | 0.37        |
| 8      | 3        | 1       | 0        | 0.5     | 6        | 0.4         |
| 9      | 0        | 0.37    | 0        | 0.5     | 0        | 0.37        |
| 10     | 2        | 0.37    | 0        | 1       | 11       | 1           |

**TABLE 3** The distribution of metabarcoding reads of the common operational taxonomic units (OTUs) between ETH and TIS for the three minibarcodes (16S Fish, 16S crus, and COI Plankton) was evaluated by using wilcoxon test (Wilcoxon 1945). When the *p*-value is lower than 0.05 (shown in BOLD), the two methods show different read distributions

applications because it does not contain sufficiently conserved regions, despite the availability of a species-specific reference database (Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). To overcome the need for a universal barcode, some researchers have replaced COI with an 18S nuclear DNA barcode in diet analysis (Albaina et al., 2016) and environmental DNA surveys (Stat et al., 2017). In fact, considering that bacterial sequences (also found in the raw data analysis and removed from OTU analysis), unidentified “no hits” and other eukaryotic taxa were detected by the COI barcode, this reflects its universal detection ability (Somervuo et al., 2017). For our results, the high overall number of COI OTUs could be due to COI barcode limitations, that is a large number of unidentified “no hits”.

Clearly, the period of digestion, and inhibition due to bacteria, digestive enzymes etc., are two crucial variables in dietary assays (Zarzo-Lacoste, Corse, & Vidal, 2013). The longer prey DNA goes through digestive phases, the more inhibited it is likely to be, effectively influencing its quantification and amplification (Taberlet et al., 2018). Nonetheless, gut-derived prey DNA is thought to be of better quality than scat-derived prey DNA since digestion is at its initial phase. As one example, Kamenova, Mayer, Coissac, Plantegenest, and Traugott (2017) noted gut DNA was detectable for longer than scat DNA after periods of feeding and starvation in the predatory carabid beetle (*Pterostichus melanarius*). In fish stomachs, DNA persistence has been recorded ranging between 16 and 24 hr post digestion using mid-throughput sequencing (Carreon-Martinez et al., 2011). However, because the stomach likely contains bacteria and co-extracted substances, DNA can still certainly be affected by inhibition impairing PCR amplification (Zarzo-Lacoste et al., 2013). Various commercial DNA extraction kits enable the reduction or removal of inhibition. Most published studies describing marine organism diets using stomach-content samples either included an analysis step, or extraction kit that reduced inhibition (Table 2). An example of dealing with PCR inhibitors in dietary studies involves scat samples, which are less invasive and far easier to collect than stomach samples (Berry et al., 2017; Hardy et al., 2017). Scat DNA was often found to be highly inhibited and degraded. Quantitative

PCR (qPCR) is a key step that facilitates determining an appropriate dilution to reduce inhibition levels for downstream analysis, and for estimating whether sufficient target DNA is obtained from the DNA extraction. In fact, coupling qPCR and high-throughput sequencing is thought to be the best option for in-depth diet analysis when using scat samples (Murray et al., 2011). However, studies on stomach samples of marine organisms did not always include qPCR before NGS sequencing (Table 2). Although we did not use the qPCR approach here, it is clear that qPCR provides clear benefits for metabarcoding studies.

The study organisms' (Pygmy devil ray) diet has not been well studied. Nevertheless, dietary research on mobulids has described the group as mostly planktivorous, feeding on various zooplankton and small fish (Couturier et al., 2012). Although the techniques in use, that is stable isotope and fatty acid profiling, and visual identification are generally accepted, most existing methods have failed to resolve higher taxonomy in prey identification (Pompanon et al., 2012; Berry et al., 2015). Since the collection of scat samples from many marine organisms are particularly challenging, dietary studies rely on stomach samples, which in turn are dependent on fishery and fishery-independent sampling or strandings. Moreover, the ethanol-retrieved DNA methodology could be applied where stomach contents are in a liquefied state or tissue is not available. It is clear that the DNA metabarcoding approach has already provided new insights into the dietary habits of several marine organisms (Table 2). This approach has been implemented in generalist feeders resulting in improved resolution compared to visual inspection. For instance, a dietary study of the Lionfish (*Pterois volitans*) contributed to identifying 39 species from digested liquiform samples, whose identification would be difficult using traditional methods (Harms-Tuohy, Schizas, & Appeldoorn, 2016). Interestingly, stomach-content analysis of the brown shrimp (*Crangon crangon*), a generalist/scavenger feeder, has been used recently as a natural sampler to monitor estuarine biodiversity through DNA metabarcoding (Siegenthaler, Wangenstein, Soto et al., 2018). Therefore, future dietary studies should consider the importance of fine-tuning a methodology to allow accurate recovery and identification of the full range of prey.

## 5 | CONCLUSIONS

Studying trophic dynamics via high-throughput sequencing requires an efficient low-cost methodology that accurately recovers, identifies, and characterizes all ingested prey—in terms of both taxonomic diversity and ideally, levels of biomass (Taberlet et al., 2018). The methodology implemented for dietary studies of marine organisms varies according to the study taxon and the sample condition (Table 2). Even though this study investigated only one animal species, we suggest that DNA extractions from stomach-content tissue are more reliable than extractions directly from preservative ethanol but comparing these two methods in other species is required in order to verify whether this finding is universal. Ethanol-based methods can still facilitate a rapid overview of diet, and such methods may be applicable for liquefied stomach contents, using ethanol for resuspension when tissue samples are not available. The method might also be applied to bulk mixed taxonomic vouchers (e.g., bulk insect/plankton/crustacean collections preserved in ethanol)—a method routinely used in rapid biodiversity inventories.

Future research warrants testing ethanol extraction coupled with inhibition removal kits to determine if ethanol performs equally well as tissue. However, this may entail an increase of material costs and laboratory work. Moreover, as qPCR has an important role in ensuring robustness of the methodology, its effectiveness should be evaluated when DNA extraction kits include an inhibition removal step. The application of multiple DNA metabarcoding assays is certainly a promising approach for taxonomic studies on the feeding habits of marine organisms. However, this type of study will always be constrained by spatial and temporal coverage. Integrating the metabarcoding approach with isotope analysis and fatty acid profiling would improve resolution for long-term dietary studies.

We conclude that when stomach-content samples are available, researchers should consider which methodology is appropriate according to the sample status (e.g., level of degradation) and current knowledge of the host diet (e.g., generalist or specialist feeder). However, assessing diet only via the visual identification of ingested prey may underestimate taxonomic diversity. It is essential to integrate new high-throughput approaches into dietary studies to inform fishery management and marine conservation; implicit among which is identifying appropriate, low-cost metabarcoding methodologies to accurately assess taxonomic composition.

## ACKNOWLEDGMENTS

We thank Joseph Di Battista and Michael Stat for general metabarcoding discussions and guidance with our bioinformatic pipeline. MdB acknowledges start-up funds from The University of Sydney and the USyd School of Life and Environmental Sciences. MKB acknowledges funding from NSW DPI. We are thankful to three anonymous referees for their helpful comments to improve the manuscript.

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTIONS

M.d.B., M.A.C., and M.K.B. conceived the study. M.K.B. and M.A.C. collected the samples. M.B., T.K., and M.d.B. designed and conducted the experiments. M.B. analyzed the data. M.B. and M.d.B. wrote the paper. All authors contributed intellectually to the interpretation of the results and writing of the manuscript.

## DATA ACCESSIBILITY

<https://doi.org/10.5061/dryad.288vb66>

## ORCID

Matteo Barbato  <https://orcid.org/0000-0001-5368-2090>

Melinda A. Coleman  <https://orcid.org/0000-0003-2623-633X>

Matt K. Broadhurst  <https://orcid.org/0000-0003-0184-7249>

Mark de Bruyn  <https://orcid.org/0000-0003-1528-9604>

## REFERENCES

- ALA (2018). Atlas of Living Australia. Retrieved from <http://www.ala.org.au>.
- Albaina, A., Aguirre, M., Abad, D., Santos, M., & Estonba, A. (2016). 18S rRNA V9 metabarcoding for diet characterization: A critical evaluation with two sympatric zooplanktivorous fish species. *Ecology and Evolution*, 6(6), 1809–1824. <https://doi.org/10.1002/ece3.1986>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Berry, O., Bulman, C., Bunce, M., Coghlan, M., Murray, D., & Ward, R. (2015). Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. *Marine Ecology Progress Series*, 540, 167–181. <https://doi.org/10.3354/meps11524>
- Berry, T. E., Osterrieder, S. K., Murray, D. C., Coghlan, M. L., Richardson, A. J., Grealy, A. K., ... Bunce, M. (2017). DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (*Neophoca cinerea*). *Ecology and Evolution*, 7(14), 5435–5453.
- Bohmann, K., Monadjem, A., Lehmkuhl Noer, C., Rasmussen, M., Zeale, M. R. K., Clare, E., & Gilbert, M. T. P. (2011). Molecular diet analysis of two African free-tailed bats (Molossidae) using high throughput sequencing. *PLoS ONE*, 6(6), e21441. <https://doi.org/10.1371/journal.pone.0021441>
- Broadhurst, M. K., Laglbauer, B. J. L., Burgess, K. B., & Coleman, M. A. (2018). Reproductive biology and range extension for *Mobula kuhlii* cf. *eregoodootenkee*. *Endangered Species Research*, 35, 71–80. <https://doi.org/10.3354/esr00876>
- Carreon-Martinez, L., Johnson, T. B., Ludsins, S. A., & Heath, D. D. (2011). Utilization of stomach content DNA to determine diet diversity in piscivorous fishes. *Journal of Fish Biology*, 78(4), 1170–1182. <https://doi.org/10.1111/j.1095-8649.2011.02925.x>
- Clare, E. L., Fraser, E. E., Braid, H. E., Fenton, M. B., & Hebert, P. D. N. (2009). Species on the menu of a generalist predator, the eastern red



- bat (*Lasiurus borealis*): Using a molecular approach to detect arthropod prey. *Molecular Ecology*, 18, 2532–2542.
- Couturier, L. I. E., Marshall, A. D., Jaine, F. R. A., Kashiwagi, T., Pierce, S. J., Townsend, K. A., ... Richardson, A. J. (2012). Biology, ecology and conservation of the Moblulidae. *Journal of Fish Biology*, 80(5), 1075–1119. <https://doi.org/10.1111/j.1095-8649.2012.03264.x>
- Deagle, B. E., Gales, N. J., Evans, K., Jarman, S. N., Robinson, S., Trebilco, R., & Hindell, M. A. (2007). Studying seabird diet through genetic analysis of faeces: A case study on macaroni penguins (*Eudyptes chrysolophus*). *PLoS ONE*, 2(9).
- 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>
- Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ... Pfrender, M. E. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26(21), 5872–5895. <https://doi.org/10.1111/mec.14350>
- Drummond, A. J., Newcomb, R. D., Buckley, T. R., Xie, D., Dopheide, A., Potter, B. C., & Grosser, S. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. *GigaScience*, 4, 1. <https://doi.org/10.1186/s13742-015-0086-1>
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26, 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Fernández-Álvarez, F. Á., Machordom, A., García-Jiménez, R., Salinas-Zavala, C. A., & Villanueva, R. (2018). Predatory flying squids are detritivores during their early planktonic life. *Scientific Reports*, 8(1), 3440. <https://doi.org/10.1038/s41598-018-21501-y>
- 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.
- Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE*, 6(4), e17497. <https://doi.org/10.1371/journal.pone.0017497>
- Hajibabaei, M., Spall, J. L., Shokralla, S., & van Konyenburg, S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology*, 12(1), 28. <https://doi.org/10.1186/1472-6785-12-28>
- Hardy, N., Berry, T., Kelahe, B. P., Goldsworthy, S. D., Bunce, M., Coleman, M. A., ... Figueira, W. (2017). Assessing the trophic ecology of top predators across a recolonisation frontier using DNA metabarcoding. *Marine Ecology Progress Series*, 573, 237–254.
- Harms-Tuohy, C. A., Schizas, N. V., & Appeldoorn, R. S. (2016). Use of DNA metabarcoding for stomach content analysis in the invasive lionfish *Pterois volitans* in Puerto Rico. *Marine Ecology Progress Series*, 558, 181–191. <https://doi.org/10.3354/meps11738>
- Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, 270, S96–S99.
- Huson, D. H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., ... Tappu, R. (2016). MEGAN community edition – interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Computational Biology*, 12(6), e1004957.
- IUCN. (2018). The IUCN Red List of Threatened Species. Version, 2017. Retrieved from [www.iucnredlist.org](http://www.iucnredlist.org)
- Jakubavičiūtė, E., Bergström, U., Eklöf, J. S., Haanel, Q., & Bourlat, S. J. (2017). DNA metabarcoding reveals diverse diet of the three-spined stickleback in a coastal ecosystem. *PLoS ONE*, 12(10), e0186929. <https://doi.org/10.1371/journal.pone.0186929>
- Kamenova, S., Mayer, R., Coissac, E., Plantegenest, M., & Traugott, M. (2017). Comparing three types of dietary samples for prey DNA decay in an insect generalist predator. *bioRxiv*, 098806.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... Drummond, A. (2012, June 12). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), 1647–1649.
- Kraaijeveld, K., Weger, L. A., Ventayol Garcia, M., Buermans, H., Frank, J., Hiemstra, P. S., & Dunnen, J. T. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular Ecology Resources*, 15, 8–16. <https://doi.org/10.1111/1755-0998.12288>
- 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), 34. <https://doi.org/10.1186/1742-9994-10-34>
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics: Essays in Honor of Harold Hotelling*. (pp. 278–292). Palo Alto, CA: Stanford University Press.
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24(3), 133–141. <https://doi.org/10.1016/j.tig.2007.12.007>
- Murray, D. C., Bunce, M., Cannell, B. L., Oliver, R., Houston, J., White, N. E., ... Haile, J. (2011). DNA-based faecal dietary analysis: A comparison of qPCR and high throughput sequencing approaches. *PLoS ONE*, 6(10), e25776. <https://doi.org/10.1371/journal.pone.0025776>
- 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. *PLoS ONE*, 11(3), e0148698. <https://doi.org/10.1371/journal.pone.0148698>
- 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>
- Qiagen, A. G. (2006). *DNeasy® blood and tissue handbook*. Hombrechtikon, Switzerland: Qiagen AG.
- Robertson, K. M., Minich, J., Bowman, A. J., & Morin, P. A. (2013). A thin soup: Extraction and amplification of DNA from DMSO and ethanol used as preservative for cetacean tissue samples. *Conservation Genetics Resources*, 5(4), 929–933. <https://doi.org/10.1007/s12686-013-9934-4>
- RStudio Team. (2015). *RStudio: Integrated Development for R*. Boston, MA: RStudio Inc. <http://www.rstudio.com/>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shokralla, S., Singer, G. A., & Hajibabaei, M. (2010). Direct PCR amplification and sequencing of specimens' DNA from preservative ethanol. *BioTechniques*, 48(3), 305–306. <https://doi.org/10.2144/000113362>
- Siegenthaler, A., Wangenstein, O. S., Benvenuto, C., Campos, J., & Mariani, S. (2018). DNA metabarcoding unveils multi-scale trophic variation in a widespread coastal opportunist. *Molecular Ecology. Online Early*, <https://doi.org/10.1111/mec.1488>
- Siegenthaler, A., Wangenstein, O. S., Soto, A. Z., Benvenuto, C., Corrigan, L., & Mariani, S. (2018). Metabarcoding of shrimp stomach content: Harnessing a natural sampler for fish biodiversity monitoring. *Molecular Ecology Resources. Online Early*, <https://doi.org/10.1111/1755-0998.12956>
- Somervuo, P., Douglas, W. Y., Xu, C. C., Ji, Y., Hultman, J., Wirta, H., & Ovaskainen, O. (2017). Quantifying uncertainty of taxonomic placement in DNA barcoding and metabarcoding. *Methods in Ecology and Evolution*, 8(4), 398–407. <https://doi.org/10.1111/2041-210X.12721>
- Sousa, L. L., Xavier, R., Costa, V., Humphries, N. E., Trueman, C., Rosa, R., ... Queiroz, N. (2016). DNA barcoding identifies a cosmopolitan

- diet in the ocean sunfish. *Scientific Reports*, 6, 28762. <https://doi.org/10.1038/srep28762>
- Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., ... Bunce, M. (2017). Ecosystem biomonitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, 7(1), 12240. <https://doi.org/10.1038/s41598-017-12501-5>
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA for Biodiversity Research and Monitoring*. Oxford, UK: Oxford University Press.
- Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, 360, 1847–1857.
- White, W. T., Corrigan, S., Yang, L., Henderson, A. C., Bazinet, A. L., Swofford, D. L., & Naylor, G. J. (2017). Phylogeny of the manta and devilrays (Chondrichthyes: Mobulidae), with an updated taxonomic arrangement for the family. *Zoological Journal of the Linnean Society*, 182(1), 50–75. <https://doi.org/10.1093/zoolinnea/zlx018>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80–83. <https://doi.org/10.2307/3001968>
- Yoon, T. H., Kang, H. E., Lee, S. R., Lee, J. B., Baeck, G. W., Park, H., & Kim, H. W. (2017). Metabarcoding analysis of the stomach contents of the Antarctic Toothfish (*Dissostichus mawsoni*) collected in the Antarctic Ocean. *PeerJ*, 5, e3977.
- Zarzoso-Lacoste, D., Corse, E., & Vidal, E. (2013). Improving PCR detection of prey in molecular diet studies: Importance of group-specific primer set selection and extraction protocol performances. *Molecular Ecology Resources*, 13(1), 117–127. <https://doi.org/10.1111/1755-0998.12029>

**How to cite this article:** Barbato M, Kovacs T, Coleman MA, Broadhurst MK, de Bruyn M. Metabarcoding for stomach-content analyses of Pygmy devil ray (*Mobula kuhlii* cf. *eregoodootenkee*): Comparing tissue and ethanol preservative-derived DNA. *Ecol Evol*. 2019;9:2678–2687. <https://doi.org/10.1002/ece3.4934>