A peer-reviewed version of this preprint was published in PeerJ on 20 September 2018.

View the peer-reviewed version (peerj.com/articles/5641), which is the preferred citable publication unless you specifically need to cite this preprint.

van der Reis AL, Laroche O, Jeffs AG, Lavery SD. 2018. Preliminary analysis of New Zealand scampi (*Metanephrops challengeri*) diet using metabarcoding. PeerJ 6:e5641 https://doi.org/10.7717/peerj.5641

Preliminary analysis of New Zealand scampi (*Metanephrops challengeri*) diet using metabarcoding

Aimee L van der Reis [Corresp., 1], **Olivier Laroche** [2], **Andrew G Jeffs** [1, 2], **Shane D Lavery** [1, 2]

[1] Institute of Marine Science, University of Auckland, Auckland, New Zealand

[2] School of Biological Sciences, University of Auckland, Auckland, New Zealand

Corresponding Author: Aimee L van der Reis
Email address: avan398@aucklanduni.ac.nz

Deep sea lobsters are highly valued for seafood and provide the basis of important commercial fisheries in many parts of the world. Despite their economic significance, relatively little is known about their natural diets. Microscopic analyses of foregut content in some species have suffered from low taxonomic resolution, with many of the dietary items difficult to reliably identify as their tissue is easily digested. DNA metabarcoding has the potential to provide greater taxonomic resolution of the diet of the New Zealand scampi (*Metanephrops challengeri*) through the identification of gut contents, but a number of methodological concerns need to be overcome first to ensure optimum DNA metabarcoding results.

In this study, a range of methodological parameters were tested to determine the optimum protocols for DNA metabarcoding, and provide a first view of *M. challengeri* diet. Several PCR protocols were tested, using two universal primer pairs targeting the 18S rRNA and COI genes, on DNA extracted from both frozen and ethanol preserved samples for both foregut and hindgut digesta.

The selection of appropriate DNA polymerases, buffers and methods for reducing PCR inhibitors (including the use of BSA) were found to be critical. Amplification from frozen or ethanol preserved gut contents appeared similarly dependable, but metabarcoding outcomes indicated that the ethanol samples produced better results from the COI gene. The COI gene was found to be more effective than 18S rRNA gene for identifying large eukaryotic taxa from the digesta, however, it was less successfully amplified. The 18S rRNA gene was more easily amplified, but identified mostly smaller marine organisms such as plankton and parasites. This preliminary analysis of the diet of *M. challengeri* identified a range of species (13,541 reads identified as diet), which included the ghost shark (*Hydrolagus novaezealandiae*), silver warehou (*Seriolella punctate*), tall sea pen (*Funiculina quadrangularis*) and the salp (*Ihlea racovitza*), suggesting that they have a varied diet, with a high reliance on scavenging a diverse range of pelagic and benthic species from the seafloor.

Preliminary analyses of New Zealand scampi (*Metanephrops challengeri*) diet using metabarcoding

Aimee L. van der Reis[1], Olivier Laroche[2], Andrew G. Jeffs[1, 2], Shane D. Lavery[1, 2]

[1]Institute of Marine Science, University of Auckland, Auckland, New Zealand

[2]School of Biological Sciences, University of Auckland, Auckland, New Zealand

Corresponding Author:

Aimee L. van der Reis[1]

Email address: avan398@aucklanduni.ac.nz

Abstract

Deep sea lobsters are highly valued for seafood and provide the basis of important commercial fisheries in many parts of the world. Despite their economic significance, relatively little is known about their natural diets. Microscopic analyses of foregut content in some species have suffered from low taxonomic resolution, with many of the dietary items difficult to reliably identify as their tissue is easily digested. DNA metabarcoding has the potential to provide greater taxonomic resolution of the diet of the New Zealand scampi (*Metanephrops challengeri*) through the identification of gut contents, but a number of methodological concerns need to be overcome first to ensure optimum DNA metabarcoding results.

In this study, a range of methodological parameters were tested to determine the optimum protocols for DNA metabarcoding, and provide a first view of *M. challengeri* diet. Several PCR protocols were tested, using two universal primer pairs targeting the 18S rRNA and COI genes, on DNA extracted from both frozen and ethanol preserved samples for both foregut and hindgut digesta.

The selection of appropriate DNA polymerases, buffers and methods for reducing PCR inhibitors (including the use of BSA) were found to be critical. Amplification from frozen or ethanol preserved gut contents appeared similarly dependable, but metabarcoding outcomes indicated that the ethanol samples produced better results from the COI gene. The COI gene was found to be more effective than 18S rRNA gene for identifying large eukaryotic taxa from the digesta, however, it was less successfully amplified. The 18S rRNA gene was more easily amplified, but identified mostly smaller marine organisms such as plankton and parasites. This preliminary analysis of the diet of *M. challengeri* identified a range of species (13,541 reads identified as diet), which included the ghost shark (*Hydrolagus novaezealandiae*), silver warehou (*Seriolella*

punctate), tall sea pen (*Funiculina quadrangularis*) and the salp (*Ihlea racovitza*), suggesting that they have a varied diet, with a high reliance on scavenging a diverse range of pelagic and benthic species from the seafloor.

Introduction

Commercial fisheries for deep sea lobster species, those typically captured below 50 m depth, are occurring in many parts of the world, currently producing total annual landings of around 50,000 t (Jeffs 2010; FAO 2016). Many of the targeted species are of high value, with wholesale export prices of over \$40 kg^{-1} common for some species, such as the New Zealand (NZ) scampi *Metanephrops challengeri* (Balss 1914; Seafood New Zealand 2017). Despite the economic importance of deep sea lobsters, the knowledge of their feeding ecology and diet is limited. Improved information on diet in deep sea lobsters has potential uses in identifying effective baits for trap fishing, development of suitable feeds for aquaculture, as well as understanding differences in growth rates in natural populations to help improve the management of the fisheries.

Species in two genera of deep-sea lobster, *Metanephrops* and *Nephrops*, are widely targeted by commercial fisheries, share similar ecology and are genetically closely related (Tshudy, Chan & Sorhannus 2007). The Norway lobster, *Nephrops norvegicus,* has a varied diet, which is necessary to achieve optimal growth (Cristo & Cartes 1998; Mente 2010). Similar varied diets have been suggested for some *Metanephrops* species (excluding *M. challengeri*), and include fish, crustaceans, polychaetes and amphipods, although these have only been identified at a crude taxonomic level using microscopic analysis of foregut contents (Choi et al. 2008; Sahlmann, Chan & Chan 2011; Wahle et al. 2012; Bell, Tuck & Dobby 2013). A significant problem with microscopic analysis of gut contents is poor taxonomic resolution, due to difficulties in identification of partly-digested specimens, which typically require an expert for reliable taxonomic identification (Dunn et al. 2010; Pompanon et al. 2012; Zhan et al. 2013; Berry et al. 2015; Young et al. 2015; Crisol-Martínez et al. 2016; Harms-Tuohy, Schizas & Appeldoorn

2016; Sousa et al. 2016). Also, soft-bodied animals are frequently suspected of being under-represented in such analyses as they are highly digestible (Bell, Tuck & Dobby 2013).

A more promising tool for diet analysis, DNA metabarcoding, combines universal DNA primers with high-throughput (next-generation) sequencing to identify a variety of species from a mixture of gut content DNA (Kress et al. 2015). A single universal primer pair has the ability to amplify a diverse range of species by targeting a single gene region that has been conserved among phylogenetically distinct taxa (e.g., the cytochrome oxidase I [COI] region in the mitochondrial DNA or the 18S ribosomal RNA [rRNA] region in the nuclear DNA; Aylagas et al. 2016). This molecular method has been successfully applied in diet studies of a variety of lobster larvae and marine fish species (O'Rorke et al. 2012; O'Rorke et al. 2014; Berry et al. 2015; Harms-Tuohy, Schizas & Appeldoorn 2016). One of the major advantages of this method of gut analysis is that a high degree of taxonomic resolution can be achieved, with the identification of individual species frequently possible. Also, the digested or liquid gut content, which would normally be of no value for visual identification of morphology, can provide additional dietary information that would otherwise have been missed. The potential difficulties of this approach include the sequencing and labour costs, the sensitivity of detection, and the taxonomic coverage of the reference sequence databases at the time of study. These difficulties will likely fade as research progresses in this field (Cowart et al. 2015; Srivathsan et al. 2016).

With sequencing cost and sensitivity being the major restrictions, it is in the best interest of researchers to first address methodological issues before a full analysis is conducted, in order to ensure the best possible results are obtained. One of the potential difficulties in metabarcoding is polymerase chain reaction (PCR) inhibition, which has often been found when using template genomic DNA extracted from biological material with a high proportion of organic (i.e., bile

salts) or inorganic (i.e., calcium ions) compounds, or from body fluids or some difficult organs (Rossen et al. 1992; Kreader 1996; Rådström et al. 2004; Farell & Alexandre 2012; Schrader et al. 2012). Another matter to address is determining which DNA polymerases are more susceptible to the effects of specific PCR inhibitors present that impede DNA amplification (Rådström et al. 2004). A third potential problem is the taxonomic coverage and resolution possible from the different reference sequences in the databases available, with some databases targeting different genes or different organisms. Other sampling factors may also be important in determining the success of metabarcoding for diet analysis. These include the preservation method (e.g., frozen or in ethanol) and the location of the gut contents to be analysed (from the foregut or hindgut), with the potential for these two locations to provide different results due either to the state of DNA degradation or the differences in timing of intake and digestive processing of dietary items.

The natural diet of *M. challengeri* is largely unknown, despite forming the basis of an important commercial fishery in New Zealand, and currently being of some interest for aquaculture development. The goal of this study is to provide a preliminary examination of the natural diet of this species using metabarcoding methods, by firstly addressing the optimisation of selected methodological factors, including issues of tissue choice and preservation, target gene, PCR inhibition, PCR reagents and reference database coverage.

Materials & Methods

Sample Collection and Locations

Samples of *M. challengeri* were collected in September of 2016 from the seafloor of the Chatham Rise, New Zealand, at depths ranging from 200 – 500 m from the R.V. Kaharoa by towing benthic trawl nets with a 45 mm cod-end. Upon landing the trawl on the deck, the *M. challengeri* that were intact were euthanized by immersing them in chilled 95% ethanol as soon as possible to avoid any degradation. The samples were transferred to the University of Auckland's laboratories (Auckland, New Zealand) for further examination. A single trawl location (start -42.9952°, 177.2367° at a depth of 322 m; end -43.0210°, 177.1800° at a depth of 252 m) and the individuals collected there were selected for the purpose of this study. For comparison, frozen individuals of commercially harvested *M. challengeri* were supplied by Sanford Ltd (Auckland, New Zealand) that were collected from the seafloor using a benthic trawl on the continental shelf off New Zealand's Auckland Islands (the trawl took place within a quadrant: corner A -49.0°, 168.5°; corner B -51.5°, 168.5°; corner C -52.0°, 164.5°; corner D -49.0°, 164.5°) and snap frozen immediately after collection.

A special permit (#549) for *M. challengeri* collection was provided by New Zealand's Ministry of Primary Industries. The specimens for this study were collected in accordance with approvals under New Zealand's Animal Welfare Act 1991 approved by the Animal Ethics Committee of the Nelson, Marlborough Institute of Technology (AEC2014-CAW-02).

Specimen Dissection, Gut Content Removal and DNA Extraction

The ethanol preserved individuals were placed on a paper towel to remove the excess ethanol, and the frozen individuals were thawed at room temperature until each individual's gut contents could be removed separately. Sterile dissection kits were used for each collection of *M.*

challengeri (frozen and ethanol) and the dissection tools were thoroughly cleaned between every individual in each collection; as well as between every individual's hindgut and foregut digesta removal. The gut digesta remained separated throughout the extraction process for each individual, and were placed into separate 1.5 mL microcentrifuge tubes, prior to DNA extraction.

A close microscopic examination of samples of digesta found an almost complete absence of gut items that were sufficiently visually recognisable for them to be reliably assigned to taxa.

Samples of gut contents were each mixed by vortexing briefly before subsamples of each were taken to be extracted using each of the following DNA extraction kits by following the manufacturers' instructions; Gentra® Puregene® Tissue Kit (Qiagen®, Hilden, Germany), E.Z.N.A.® Mollusc DNA Kit (OMEGA Bio-tek, Inc., Georgia, USA), DNeasy® Tissue and Blood Kit (Qiagen®), and Mo Bio's Powerbiofilm® DNA Isolation Kit (now Qiagen® DNeasy PowerBiofilm Kit). The concentration of recovered DNA was measured for each using the NanoPhotometer® N60 (Implen, Munich, Germany). Comparisons revealed the Gentra® Puregene® Tissue Kit had superior DNA recovery and was used for subsequent DNA extraction of all *M. challengeri* gut samples.

Primer Selection and PCR Protocol

Two universal primer pairs were identified and selected from the literature: mlCOIintF (Leray et al. 2013) and jgHCO2198 (Geller et al. 2013) targeting a 313 base pair (bp) region of the mitochondrial COI gene, and Uni18SF and Uni18SR (Zhan et al. 2013) targeting a 425 bp region of the nuclear 18S rRNA V4 variable region. All primers included Illumina Nextera™ library transposase adapters (Table 1). The GC content of the primers are 50.8% for mlCOIintF, 47.5 % for jgHCO2198, 56.6 % for Uni18SF and 51.9 % for Uni18SR. Primers to block the

amplification of host DNA were not used, as previous related work on lobsters (O'Rorke et al. 2012; O'Rorke et al. 2014) had shown that they also block amplification of related crustaceans, which could be an important part of the diet of *M. challengeri*.

After preliminary trials of a variety of DNA polymerases and buffer mixes, two were tested extensively, Platinum *Taq* (Invitrogen™, Thermo Fisher Scientific Inc., Massachusetts, USA; Table S1) and the MyTaq™ Red Mix (Bioline, London, UK; Table S2). The PCR temperature profile was optimised to provide the best amplification for both the COI and 18S primers (Fig. S1). Each PCR cycle included a negative control (no DNA added) to check PCR reagents were not contaminated. The PCR products were run on 1.6% agarose gel and viewed in a Gel Doc™ XR+ (Bio-Rad Laboratories Inc., California, USA). The PCR products were visualised using Gel Red® (Biotium, California, USA).

The effect of both genomic DNA template concentration and BSA concentration were tested to try to reduce PCR inhibition. BSA (1%) was tested with volumes of 1 µl, 2 µl and 5 µl per 25 µl reaction. A range of different DNA dilutions (1:10, 1:50 and 1:100) were also tested.

Inconsistent PCR amplification of both genes was observed in the gut content samples, but not from muscle DNA extracts from the same individuals, likely due to the presence of PCR inhibitors. To test the effects of the presumed PCR inhibitors, gut content template DNA (1µl) was added to tail muscle template DNA (1 µl of 10 ng μl^{-1}) in the standard PCRs. Each individual was tested in three replicate PCRs.

DNA Metabarcoding and Analyses

Selection of individuals for metabarcoding

The PCR products for sequencing were selected from six individuals (70.2, 70.3, 70.9, Fro1, Fro2 and Fro3), as they provided consistent and strong amplifications for both COI and 18S genes. The PCR products encompassed the range of methodological factors that were to be addressed (Table 2). Individual 70.9 was used for comparison of the PCRs to determine which DNA polymerase and buffer mix was better (i.e., samples 1 and 2 versus 3 and 4). Preliminary data suggested the Platinum *Taq* reactions were inferior, so the remainder of the comparisons focussed on the Bioline reactions. The taxa that were identified from the digesta of the foregut and the hindgut were compared for the individuals 70.2 and 70.9 (i.e., samples 1 and 5 versus 2 and 6) to identify if the hindgut digesta could provide useful additional information about the diet, i.e., successive meals may be represented in the hindgut contents. The ethanol and frozen preservation methods were compared using three individuals each (i.e., samples 1, 2, 5, 6 and 7 versus 8 and 9), to assess the effects (if any) of differential DNA degradation of the digesta. All individuals were used for the comparison of the database identifications for the COI and 18S genes and thus a preliminary analysis of the diet was evaluated. Potential diet false positives were identified by evaluating the DNA negative.

DNA purification and pooling of selected samples

NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel, Düren, Germany) was used according to manufacturer's instructions to purify the PCR products and remove fragments of less than 164 bp in length for all selected individuals. The amplified DNA concentration of those products was determined using Qubit™ dsDNA HS Assay Kit (Invitrogen™, Thermo Fisher Scientific Inc.) following the manufacturer's protocol.

188 The purified COI and 18S PCR products from each sample were pooled and brought to equal
189 molarity where possible. Sequencing was done through New Zealand Genomics Ltd (Auckland,
190 New Zealand) at Massey University (Palmerston North, New Zealand) where indexing occurred
191 using the Nextera™ DNA library Prep Kit (Illumina, California, USA) before sequencing on an
192 Illumina MiSeq™ System (2 × 250 pair pair-end protocol).

193 ***Metabarcoding protocol***

194 The raw Illumina sequences were analysed through the New Zealand eScience Infrastructure
195 (NeSI; Auckland, New Zealand) high performance computing (HPC) facility. The resulting
196 Illumina metabarcoding sequenced reads were processed by firstly removing the primers (no
197 mismatch tolerated) using fastq-multx (version 1.3.1) and by pairing the reads with
198 SolexaQA++. Low quality 3' end sequences (Phred scores below 3) were truncated and reads
199 merged using the sequence analysis tool VSEARCH version 2.3.0 (Rognes et al. 2016), allowing
200 a maximum of five non-matching nucleotides in the overlap region. Merged reads were quality
201 filtered on the expected error value (<1) and dereplicated. Chimera checking and removal was
202 performed on the dereplicated sequences with the QIIME package (Caporaso et al. 2010) based
203 on the *de novo* and reference based methods of Usearch61 (Edgar 2010). Reference based
204 chimera detection of 18S was performed with the largest curated database available (SILVA;
205 Quast et al. 2013) and for COI, using the Midori database (Machida et al. 2017). The cleaned
206 reads (Table S3) were clustered into OTUs with the Swarm methodology (clustering threshold of
207 d2; Mahé et al. 2014).

208 ***Sequence assignment and analysis***

209 Representative OTU sequences (seed sequence of each OTU) were taxonomically assigned with
210 the Ribosomal Database Project (RDP) classifier (minimum confidence level of 80%; Wang et

al. 2007) using the curated databases Protist Ribosomal Reference (PR2; Guillou et al. 2013) and SILVA for 18S, and Midori database for COI. Representative sequences were also assigned to reference sequences from the National Center for Biotechnology Information (NCBI) Genbank database (Benson et al. 2013), using the megablast option of blastn (Morgulis et al. 2008) with an e-value threshold of 0.001. Only the best hit for each sequence was kept. NCBI Genbank is an all-inclusive database and thus was used for both the 18S gene sequences and COI gene sequences, however, sequences were only assigned to a taxon if the sequence could be assigned to a genus or species. NCBI Genbank is also not a curated database. The Midori database is specifically for metazoan mitochondrial DNA sequences. The PR2 database consists mainly of nuclear-encoded protistan sequences, and SILVA is an aligned database of small and large subunit rRNA genes. Taxonomically assigned OTUs were filtered for potential PCR or sequencing artefacts by selecting OTUs with five or more sequence reads ("hits"). The resulting filtered OTUs for COI were then further refined to keep only the OTUs that could be identified to either a genus or species level by Midori and the NCBI databases. The same was done for 18S with PR2, SILVA and NCBI. The resulting data for genus or species were used for assessing the methodological factors and preliminary diet which was analysed using Rstudio® (R Core Team 2017) and the collection of R packages in tidyverse (Wickham 2017).

To determine a list of identified taxa that were most likely to form part of the true diet of *M. challengeri*, taxa identified as contamination were removed from the final data. Any sequence reads identified as lobster (Astacidea) were presumed to be host contamination and removed. Sequence reads were also removed from further analysis if they matched to any taxa identified from the sequenced DNA negative sample, or to any taxa that could not be part of the *M.*

233 *challengeri* diet (e.g., terrestrial species). The final "hit counts" pertaining to digesta are referred
234 to as "diet hit counts".

235 Taxa were defined as "exclusive" if they were identified in only a single source, such as taxa
236 identified in the foregut but not in the hindgut, or solely identified in one PCR but not another.

Results

Factors Affecting PCR Amplification Success

Several important factors were identified that affected PCR amplification success prior to metabarcoding. These factors included PCR inhibition (and its reduction with BSA), optimal template DNA dilution, PCR reagents (particularly the DNA polymerase and buffer), and the success of amplification of different target genes (18S or COI). The success of the PCRs was determined by the presence and intensity of an appropriately-sized fragment on an agarose gel.

The sample DNA concentrations had a wide range (selected individuals ranged from 1.4 – 290.3 ng μl^{-1}) and, overall, had relatively poor purity ratios (as measured by A260:A280 and A260:A230 ratios; Table S4). PCR inhibition was minimised (visible increase in DNA amplification on agarose gel) when an optimal volume of 2 µl of 1% BSA was used in each 25 µl PCR, with some beneficial effect when using 1 µl or 5 µl of 1% BSA. PCR inhibition was found to occur inconsistently at a range of DNA concentrations, even with the addition of BSA. The inhibitory effect of the DNA extracted from gut material was demonstrated when amplifying both genes with DNA extracted from tail muscle tissue. PCR amplification from muscle DNA alone was usually strong, whereas PCR amplification was often dramatically reduced when DNA extracted from the gut (digesta DNA) was added. Amplification from digesta DNA was often very inconsistent between PCR replicates. Different dilutions of template DNA (1:10, 1:50 and 1:100) were also tested in an attempt to reduce PCR inhibition and provide optimal DNA amplification at the concentration best suited to the conditions of the individual sample, with the 1:10 dilution proving to result in the strongest DNA amplification overall.

Two different sets of PCR reagents (which differed primarily in their DNA polymerase and buffer) were compared. The Bioline reaction was generally more reliable in DNA amplification,

as not only was the intensity of the products generally greater, but the digesta DNA was more likely to produce PCR products for both the COI gene and 18S gene.

When comparing PCR amplification of the two target genes from digesta DNA, in general the COI gene did not amplify as well as the 18S. For several individuals, no PCR products were seen for the COI gene, while for the same DNA, the 18S gene amplified well. In addition, DNA from hindgut contents amplified more readily than DNA extracted from the foregut contents.

Factors Affecting Metabarcode Sequencing Results

PCR reagents (DNA polymerase and buffer)

A comparison between the Platinum *Taq* and the Bioline PCR reagents was made for individual 70.9, for which both sets of reagents were used (Table 2). Metabarcode sequences from both foregut and hindgut were pooled for each set of reagents (thus comparing samples 1 and 2 with 3 and 4). The total number of taxa identified and the total number of exclusive taxa identified (i.e., those found only in one category) was greater for both the COI and 18S genes when using the Bioline reaction (11 different taxa in total and six exclusive taxa), compared to the Platinum *Taq* reaction (six different taxa identified and one exclusive taxon; Table 3). The total diet hit count (number of sequence reads matching a potential dietary reference sequence in that database – i.e., excluding host matches) was also greater for the Bioline reaction.

Five of the taxa identified by Platinum *Taq* were a subset of those found in Bioline, which had a low summed hit count of 55 compared to Bioline's 468 hit count for the same taxa. Moreover, Bioline identified a higher number of exclusive taxa which averaged a hit count of ~32 hits per taxa whereas Platinum *Taq* only identified one exclusive taxon with a hit count of 13.

281 ***Hindgut versus foregut digesta***

282 The taxa identified from the hindgut and foregut of two individuals were compared (i.e., samples
283 1 and 5 versus 2 and 6). The hindgut digesta contained a higher number of both total taxa and
284 exclusive taxa compared to the foregut, when the sequences were matched to the COI gene
285 databases (Table 4). Conversely, the foregut contained a higher number of total taxa and
286 exclusive taxa with the 18S gene databases (Table 4).

287 ***Sample preservation method***

288 The effects of potential differential DNA degradation due to the method of preservation were
289 more difficult to assess directly, as the same individuals could not be compared across the two
290 conditions, and the comparison was somewhat confounded by being collected in different
291 locations. Instead, broad comparisons were made between the pooled results from the three
292 individuals preserved for each of the ethanol and frozen methods. Firstly, the potential
293 differential effects on preservation of diet DNA was assessed by calculating the diet hit count as
294 a percentage of the overall hit count (for samples 1, 2, 5, 6 and 7 versus 8 and 9; Table 5). A
295 noticeable difference was observed in the diet hit count percentage between frozen (0%) and
296 ethanol (6%) preserved individuals for the COI results (no frozen individuals provided any
297 dietary sequences; Table S5). No difference was seen between the ethanol and frozen individuals
298 for the 18S results, as both had a diet hit count percentage of 1% (Fro3 was the only frozen
299 individual to have any diet hit counts; Table S6).

300 **Database Comparison and Preliminary Assessment of *M. challengeri* Diet**

301 All individuals and their digesta amplified using the Bioline PCRs were used to provide a
302 preliminary assessment of the diet of *M. challengeri*. A total of 13,541 sequences were used to
303 analyse the diet: 22 taxa were identified from 10,611 COI sequences, and 25 taxa from 2,930

18S sequences. Among the 22 taxa identified from the COI databases were seven fish taxa (*Epinephelus epistictus*, *Helicolenus barathri, Hydrolagus novaezealandiae*, *Macruronus magellanicus*, *Notophycis marginata*, *Seriolella punctate* and *Thyrsites atun*), crab taxa (*Neosarmatium fourmanoiri*, *Microphrys branchialis* and *Homola* sp.), a marine worm (*Laonice cirrata*) and an anemone (*Diadumene leucolena*) (Fig. 1; Table S7 and Table S8). The 18S databases identified 25 taxa, which included a sea pen (*Umbellula* sp.), salps (*Pyrosoma* sp. and *Ihlea racovitzai*) and many parasites (i.e., apicomplexan *Eimeria variabilis* and dinoflagellate *Duboscquella*) (Fig. 2; Table S9 and Table S10). There were no common taxa identified between the reference databases when comparing 18S OTUs to COI OTUs.

Discussion

Methodological Issues in PCR Amplification Success

PCR inhibition

PCR inhibition is a substantial factor in this study. The spectrophotometer ratios of purity provided initial indicators of the presence of inhibitors in the extracted DNA (Watts 2014). An optimal volume of 2 µl 1% BSA successfully reduced PCR inhibition by increasing the amplification efficiency. An additional step effective in reducing PCR inhibition, was 1:10 dilution of the template genomic DNA, thus diluting the PCR inhibitors it presumably contained.

The presence of organic compounds in the foregut (i.e., bile salts in the gastric juices) is thought to be one of the main contributors to PCR inhibition which cannot always be alleviated by BSA (Lorenz 2012; Schrader et al. 2012; Harms-Tuohy, Schizas & Appeldoorn 2016). Bile salts often found in vertebrate gastric juices, are similar to an organic compound in crustaceans such as crabs, lobsters and crayfish, and suggests why BSA is not an alleviator in all DNA amplifications tested in this study (Borgström 1974). Different diets would also affect the secretion of the gastric juices, which may play a role in determining which PCR inhibitors are present and in what concentration in different parts of the gut (Rotllant et al. 2014). This demonstrates the problematic, high levels of PCR inhibition present in many genomic DNA extracts, and its influence on successfully amplifying gut content DNA. Increased quantitative monitoring of the PCR success, such as by utilising quantitative PCR, may be required in future.

One solution to PCR inhibition may be to pool replicate PCR products together. Another is the addition of the T4 gene 32 protein (gp32), which can be effective against inhibitors for which BSA is not, such as sodium dodecyl sulfate (SDS), sodium chloride and bile salts (Schrader et al. 2012).

PCR reagents (DNA polymerase and buffer)

Different DNA polymerases and buffers are available and are two of the key reagents for DNA amplification, each being susceptible to various components in the biological environment in which they amplify, thus affecting their performance (Rådström et al. 2004; Wolffs et al. 2004). A substantial difference could be seen visually in the DNA amplification when using the Bioline reaction compared to the Platinum *Taq* reaction. The Platinum *Taq* reaction appeared to be more susceptible to the effects of PCR inhibitors, and the DNA amplification was not as effective as in the Bioline reaction.

Amplification of the loci

The 18S gene region (~425 bp), in general, amplified more successfully compared to the COI gene region (~313 bp) targeted here. It is not common for shorter gene fragments to be more difficult to amplify, but it is relatively common to see some difficulties in amplification of the COI gene in some species, likely due to the greater sequence variation in the primer annealing sites (Chen, Jiang & Qiao 2012; Lv et al. 2014). However, when both genes amplified, their PCR product intensity appeared consistent between sample conditions, i.e., if the PCR product was more intense in the foregut for the 18S gene the same would be seen for the COI gene in the same individual.

Methodological Issues in Metabarcoding Success

PCR reagents (DNA polymerase and buffer)

Visualization of the PCR amplification can only suggest if amplification is occurring and not necessarily what the quality of amplification is. So although it was visually determined that the Bioline reaction generally resulted in stronger PCR amplifications, it was possible that it was amplifying only a subset of the possible templates compared to that of the Platinum *Taq* reaction.

Individual 70.9 amplified equally well for both the Platinum *Taq* and Bioline reactions, and also for both the hindgut and the foregut. The DNA sequencing revealed that the Bioline reaction amplified a higher number of taxa from the same individual's gut content DNA.

This was likely due to a higher tolerance of PCR inhibitors from the DNA polymerase and buffer in Bioline. This is suggested by the shared taxa, as Bioline (hit count 468) amplified approximately nine times more diet fragments than Platinum *Taq* (hit count 55). The exclusive taxa, presumed to have the lower DNA concentrations, identified by Bioline have ~32 hit counts per exclusive taxon, whereas, Platinum *Taq* only identified one exclusive taxon with a hit count of 13, thus indicating that the Bioline reaction appears to be more efficient at amplifying DNA from the digesta of *M. challengeri*.

Foregut versus hindgut

The examination of both the foregut and hindgut digesta provided a broader overview of the diet of *M. challengeri*. There was an increase in the number of taxa when the entire gut digesta was used, compared to using digesta from only the hindgut or the foregut. As ingested food takes a few hours to pass through the foregut and into the hindgut, the examination of the digesta from both may permit the assessment of successive meals over a broader time period before capture, or differences in gut transit times for different dietary items (Simon & Jeffs 2008; Lee, Hartstein & Jeffs 2015; Kamio et al. 2016). Also, as the chemical environment (i.e. pH) and degree of degradation of digesta differ between the gut sections, an assessment of both foregut and hindgut digesta appears to maximise the opportunity to detect the widest possible breadth of diet per individual.

It was initially expected that the hindgut contents may provide lower quality sequencing results, because it was more difficult to eliminate host tissue contamination during dissection, the more advanced state of degradation of the food, and the higher ratio of sediment to dietary items. However, a higher number of taxa of smaller marine organisms (i.e., dinoflagellates) were identified in the foregut than in the hindgut, and a higher number of taxa of larger marine organisms (i.e., mostly fish) were identified in the hindgut than in the foregut. Thus, to dismiss sampling from either the foregut or the hindgut would have narrowed the number of taxa that were able to be identified.

Host contamination

In most metabarcoding studies of diet in predators, host contamination of the PCR amplicons is an issue. However, the overall impact of host contamination can be reduced by introducing several molecular techniques that will help decrease any biased amplification of the host's DNA, and may broaden the taxonomic analysis (Polz & Cavanaugh 1998; Green & Minz 2005; Vestheim & Jarman 2008; Pompanon et al. 2012; Harms-Tuohy, Schizas & Appeldoorn 2016; Devloo-Delva et al. 2018). PCR "clamping", using additional DNA or peptide nucleic acid (PNA) oligonucleotide primers to bind to and inhibit PCR amplification of host DNA, is a favoured method for suppressing host DNA amplification in PCRs, and has been successful in diet studies (Egholm et al. 1993; Orum et al. 1993; Vestheim & Jarman 2008; Chow et al. 2011; O'Rorke et al. 2012). In this study, PCR clamping was not used, as previous research in lobsters had shown that it also restricted amplification from related crustaceans, many of which could be potential prey of *M. challengeri.* Other diet studies have also shown that the resulting sequences are sufficient to describe the diet without the use of PCR clamping (Piñol et al. 2014) and that there would not necessarily be any significant increase in diversity if they were used (Devloo-

Delva et al. 2018). The metabarcoding results showed that host DNA amplification did not completely outcompete the amplification of prey DNA, and there were sufficient sequence reads to identify a broad range of diet taxa.

Preservation methods

The quality of DNA extracted from the digesta is affected by the state of degradation due to the digestion rate in the gut (Rådström et al. 2004; Deagle & Tollit 2007; Troedsson et al. 2009). As the two different preservation methods used would have different rates in which the digesta was preserved, it was thought that one method may provide better preservation of digesta than the other. Determining which preservation method, frozen or ethanol, is better was relatively problematic to assess, due to the relatively high level of inter-individual variation and the confounded nature of the sample collection. Due to logistic constraints on the field sampling, different preservation methods could only be undertaken on different individuals taken from different locations.

The major difference between preservation methods was seen within the COI database results as the proportion of sequences matching diet taxa was 0% in the frozen individuals and 6% in the ethanol individuals. However, the diet hit count percentage from the 18S databases were both ~1% for ethanol and frozen individuals. Although not a particularly robust test due to sample size and lack of amplification in frozen individuals, it may still suggest that the frozen preservation method may increase the DNA degradation rate in the digesta.

Amplification of the loci and database review

Although there were difficulties in gaining equal success in the DNA amplification from both genes, together they provided a much broader overview of *M. challengeri* diet at the highest resolution possible.

The 18S gene and the COI gene identified a variety of taxa (Fig. 1 and Fig. 2). The COI gene is more commonly used than the 18S gene in studies of molecular taxonomy of large eukaryotes and is one of the pillars of DNA barcoding, making it a sought-after gene region for the discrimination of closely related Metazoan species (Hebert, Ratnasingham & deWaard 2003; Derycke et al. 2010; Tang et al. 2012). The 18S gene is more frequently used in molecular taxonomic studies for microbial organisms (i.e., micro-plankton and parasites; Wu, Xiong & Yu 2015).

Using a combination of databases to identify the OTUs, from a selected gene, to genus or species level proved invaluable; where the curated databases (Midori, PR2 and SILVA) could not provide a high resolution identification (i.e., only to order or family), NCBI was usually able to. Also, targeting two genes has allowed for a more inclusive capture of the different taxonomic groups because of the more complete coverage from distinct reference databases. For further analyses of the diet of *M. challengeri* both genes should be used, as this lobster, being relatively small in size, appears to take advantage of the availability of both large and small prey, both dead and alive.

Preliminary Assessment of *M. challengeri* Diet

This study undertook the first assessment of *M. challengeri* diet using DNA metabarcoding methods. Although this study investigated only a small number of individuals in this preliminary

444 analysis, it has already provided a fascinating snapshot of their diet, confirming some previous
445 expectations, and providing a broader understanding of their feeding ecology.

446 The components of the diet were identified at a much finer taxonomic level than any other
447 previous studies undertaken microscopically on any *Metanephrops* species. The diets previously
448 determined microscopically are at a poor taxonomic resolution, and included crustaceans, fish,
449 annelids and bivalves (Choi et al. 2008; Sahlmann, Chan & Chan 2011; Wahle et al. 2012).
450 *Metanephrops challengeri* are thought to be benthic foragers and scavengers, relying heavily on
451 chemosensory detection of potential food items (Major & Jeffs 2017). They are likely to be
452 scavengers of fish remains whether it is from trawl debris, sunken carcases or faeces, as well as
453 foraging for smaller dietary items such as sea pens. Using metabarcoding methods on digesta
454 also allowed identification of parasites that are either from the diet source or from *M. challengeri*
455 individuals sampled. Below, we examine some of the most interesting and prominent taxa
456 identified in the diet of *M. challengeri*, and suggest likely species identifications for those taxa
457 not classified to species or those whose counterparts are more likely to be the closest match.

458 *Metanephrops challengeri* have been reported to reside at a depth ranging from 140 – 640 m
459 (Holthuis 1991). Diet taxa were closely matched to several common fish species that reside
460 within this depth range on the Chatham Rise, including *H. novaezealandiae* (ghost shark), *N.*
461 *marginata* (dwarf codling), *S. punctata* (silver warehou) and *T. atun* (snoek) (Francis 1998;
462 Ministry for Primary Industries 2006b; Ministry for Primary Industries 2006a; Luna 2008;
463 Ministry for Primary Industries 2008; Priede 2017). The OTU identified as *M. magellanicus* is
464 likely instead to be *Macruronus novaezelandiae* (New Zealand hoki), as *M. magellanicus* is
465 located off the southern coast of Chile and Argentina and *M. novazealandiae* is located in the
466 Chatham Rise at depths from 209 – 904 m (D'Amato 2006; Connell, Dunn & Forman 2010;

Kobayashi, Mizuguchi & Matsuoka 2014). The OTU identified as *H. barathri* (sea perch) is likely to be *Helicolenus percoides,* a sea perch (Scorpaenidae) found at depths between 250 – 700 m in the Chatham Rise (Anderson 1998; Horn, Forman & Dunn 2012). *Epinephelus epistictus* (grouper) is another OTU that is probably instead its counterpart found in New Zealand waters at the appropriate depth, *Epinephelus octofasciatus*.

The only sea pen to be identified in the COI databases, *Funiculina quadrangularis*, is a common tall deep-sea sea pen that grows on muddy substrates and has been found between 20 – 2000 m in New Zealand waters (Hughes 1998). *Umbellula* sp. is a sea pen taxa identified by the 18S databases and has a world-wide distribution in depths 200 – 6260 m (Williams 1995; Williams 2014).

There are several OTUs identified as belonging to the Brachyuran crabs. Many families of this infraorder occur in New Zealand waters in the area of the Chatham Rise (Wilkens 2015), but it is not possible at this stage to identify the OTUs to their likely species from this region. OTUs identified to species in this infraorder include *M. branchialis* (Majidae, spider crabs – 12 species found on the Chatham Rise, according to the Ocean Biogeographic Information System [OBIS; OBIS 2017]), and *N. fourmanoiri* (Grapsidae), although none are known to be found in New Zealand. It is likely that these OTUs belong to genetically similar species that exist at the depth in which *M. challengeri* are found on the Chatham Rise. The *Homola* sp. may be *Homola orientalis*, as this crab is known to be distributed in New Zealand waters, living at depths of 500 m (Eldredge 1980).

The marine polychaete OTU was identified as *L. cirrata*, but this species has not been found in New Zealand waters, and it is likely to be DNA signal from a closely related benthic worm

489 species. Numerous plankton were identified, such as the deep-water salp, *I. racovitza* (tunicate),
490 which can grow to over 20 mm in diameter and are an appropriate size for *M. challengeri* to
491 handle with their feeding appendages (Pakhomov et al. 2011). This species is known to more
492 commonly occur in the high Antarctic cold-water zone, but has been found to occur in waters
493 south of Australia (Casareto & Nemoto 1986; Pakhomov et al. 2011; Ono & Moteki 2013).
494 *Ebria tripartita* (Rhizaria) has been reported world-wide, has oval cells 21-35 µm in length and
495 13-25 µm in width, and is likely to be consumed unintentionally due to their size (Tong et al.
496 1998). It is known to be a grazer of phytoplankton (nanoplanktonic diatoms and dinoflagellates)
497 and has been reported in the Hauraki Gulf, New Zealand (Gordon et al. 2012).

498 It is noticeable that no molluscs were detected, which may be due to more rapid degradation of
499 their soft tissues (Bell, Tuck & Dobby 2013). It will be interesting to examine the ratios of OTU
500 hit counts across a larger number of individuals from several locations.

501 At present it is impossible to entirely differentiate between OTUs that were directly consumed by
502 *M. challengeri*, and those that were consumed by their prey, i.e., secondary predation. However,
503 it is expected that, due to rapid decomposition, secondary predation is likely to account for only a
504 very small proportion of the taxa detected in *M. challengeri* digesta. It also cannot be dismissed
505 that *M. challengeri* may be cannibalistic, but for the purpose of this study sequences identified as
506 lobsters were removed from the diet data and considered as host contamination.

507 The parasites that have been identified are likely to be residing either within the gut of *M.*
508 *challengeri* or within the individuals that were consumed. Parasites known to infect fish were
509 identified, such as *Eimeria percae* which is known to parasitize the European perch, *Perca*
510 *fluviatilis* (Molnár et al. 2012). Syndiniales Dino-Group I and II are known to parasitize a variety

511 of marine organisms, including crustaceans (Guillou et al. 2010). Previous studies of gut contents
512 of lobster larvae using DNA methods have also identified a range of microscopic parasites
513 suggesting they may be a common component of the gut (O'Rorke et al. 2012; O'Rorke et al.
514 2015).

515 There was a small proportion of OTU sequences that were not matched to known sequences in
516 the databases and demonstrate the limitations of the reference databases. This is largely due to
517 the small amount of material from New Zealand deep-sea marine life that has been sequenced
518 and uploaded to the various databases. A possible short term solution for the analysis of the diet
519 would be to collect samples of bycatch species from the trawls, identify them, extract their DNA
520 and sequence the genes of interest. This would create a database against which the diet sequences
521 could be compared, potentially revealing more specific matches or even assigning identities to
522 unassigned sequences. Further insight into the diet of *M. challengeri* will come with a greater
523 number of individuals analysed. Data on the availability of prey will be needed in order to
524 determine diet preference, as relative concentration in the diet may simply be related to what is
525 more prevalent in their foraging area.

Conclusion

Many methodological issues have been addressed in this study, which has provided solutions for alleviating them, such as template DNA dilution and BSA addition for minimising PCR inhibition. The study has also identified considerable variation in PCR inhibition among individual samples, which indicates that it will be necessary to optimise PCR amplification for each sample. The study has also determined the optimum PCR reagents (particularly the DNA polymerase and buffer) and the benefit of using a selection of different databases for assigning OTUs to taxa for different genes.

The preliminary insight from this study into the varied diet of *M. challengeri* provides a foundation for both the production of a nutritional feed for aquaculture and an attractive bait for fishery pots. However, further examination of the diet is clearly needed. Given the variability among individuals, the minimum sample number that is analysed per collection site should be of around 10 individuals that have a moderate amount of digesta for DNA extraction. This will help in quantifying if *M. challengeri* have a diet preference and if it is related to sex, size and/or location. A future study should also proceed in collecting by-catch when *M. challengeri* are trawled, which will assist in determining prey availability, and provide a more complete sequence reference DNA database for comparison. It is clear that there is considerably more insight provided from identification using metabarcoding than from traditional microscopic identification. A further benefit is the store of diet DNA sequences that can be retained for future analysis (against updated reference databases), whereas microscopy results cannot be further analysed. Overall, this study shows there is great promise in analysing *M. challengeri* diet using metabarcoding methods.

549 **Acknowledgements**

550 We would like to thank Dr Ian Tuck (NIWA) and Mr Dean Stotter (NIWA) for providing
551 recorded information and access to samples. We thank the captain and crew of the R.V. Kaharoa
552 in facilitating the collection of the samples for this study. Lastly, thank you to Andrew Stanley
553 (Sanford) and Sophar Rach (Sanford) for assistance with the commercially harvested samples.

References

Anderson O, Bagley NW, Hurst RJ, Francis MP, Clark MR, McMillan PJ. 1998. *Atlas of New Zealand fish and squid distributions from research bottom trawls.* Wellington, NZ: NIWA.

Aylagas E, Borja Á, Irigoien X, and Rodríguez-Ezpeleta N. 2016. Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. *Frontiers in Marine Science* 3. https://doi.org/10.3389/fmars.2016.00096

Balss H. 1914. Ostasiatische Decapoden II: Die Natantia und Reptantia. *Abhandlungen der Bayrischen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Abteilung, Supp* 2:1-101.

Bell M, Tuck I, and Dobby H. 2013. Nephrops Species. In: Phillips B, ed. *Lobsters: Biology, Management, Aquaculture & Fisheries*. 2nd ed. Oxford, UK: John Wiley & Sons Ltd., 357-413.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, and Sayers EW. 2013. GenBank. *Nucleic Acids Research* 41:D36-42. https://doi.org/10.1093/nar/gks1195

Berry O, Bulman C, Bunce M, Coghlan M, Murray DC, and Ward RD. 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

Borgström B. 1974. Bile Salts—Their physiological functions in the gastrointestinal tract. *Acta Medica Scandinavica* 196:1-10. https://doi.org/10.1111/j.0954-6820.1974.tb00958.x

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, and Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7:335-336. https://doi.org/10.1038/nmeth.f.303

Casareto BE, and Nemoto T. 1986. Salps of the Southern Ocean (Australian sector) during the 1983-84 summer, with special reference to the species *Salpa thompsoni*, Foxton 1961. *Memoirs of National Institute of Polar Research Special issue* 40:221-239.

Chen R, Jiang L-Y, and Qiao G-X. 2012. The effectiveness of three regions in mitochondrial genome for aphid DNA barcoding: A case in Lachininae. *PLoS ONE* 7:e46190. https://doi.org/10.1371/journal.pone.0046190

Choi JH, Kim JN, Kim MH, Chang DS, Yoo JT, and Kim JK. 2008. Population biology and feeding habits of the nephropid lobster *Metanephrops thomsoni* (Bate, 1888) in the East China Sea. *Journal of Environmental Biology* 29:453-456.

Chow S, Suzuki S, Matsunaga T, Lavery S, Jeffs A, and Takeyama H. 2011. Investigation on natural diets of larval marine animals using peptide nucleic acid-directed polymerase chain reaction clamping. *Marine Biotechnology* 13:305-313. https://doi.org/10.1007/s10126-010-9301-3

Connell AM, Dunn MR, and Forman J. 2010. Diet and dietary variation of New Zealand hoki *Macruronus novaezelandiae*. *New Zealand Journal of Marine and Freshwater Research* 44:289-308. https://doi.org/10.1080/00288330.2010.515232

Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, and Arnaud-Haond S. 2015. Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS ONE* 10:e0117562. https://doi.org/10.1371/journal.pone.0117562

Crisol-Martínez E, Moreno-Moyano LT, Wormington KR, Brown PH, and Stanley D. 2016. Using next-generation sequencing to contrast the diet and explore pest-reduction services of sympatric bird species in macadamia orchards in Australia. *PLoS ONE* 11:e0150159. https://doi.org/10.1371/journal.pone.0150159

600 Cristo M, and Cartes JE. 1998. A comparative study of the feeding ecology of *Nephrops norvegicus* (L.),
601 (Decapoda: Nephropidae) in the bathyal Mediterranean and the adjacent Atlantic. *Scientia*
602 *Marina* 62:81-90.
603 D'Amato ME. 2006. Demographic expansion and subtle differentiation in the longtailed hake
604 *Macruronus magellanicus*: evidence from microsatellite data. *Marine Biotechnology* 8.
605 https://doi.org/10.1007/s10126-005-5075-4
606 Deagle BE, and Tollit DJ. 2007. Quantitative analysis of prey DNA in pinniped faeces: potential to
607 estimate diet composition? *Conservation Genetics* 8:743-747. https://doi.org/10.1007/s10592-
608 006-9197-7
609 Derycke S, Vanaverbeke J, Rigaux A, Backeljau T, and Moens T. 2010. Exploring the use of cytochrome
610 oxidase c subunit 1 (COI) for DNA barcoding of free-living marine nematodes. *PLoS ONE*
611 5:e13716. https://doi.org/10.1371/journal.pone.0013716
612 Devloo-Delva F, Huerlimann R, Chua G, Matley J, Heupel M, Simpfendorfer C, and Maes G. 2018. How
613 does marker choice affect your diet analysis: comparing genetic markers and digestion levels for
614 diet metabarcoding of tropical-reef piscivores. *Marine and Freshwater Research*.
615 https://doi.org/10.1071/MF17209
616 Dunn MR, Szabo A, McVeagh MS, and Smith PJ. 2010. The diet of deepwater sharks and the benefits of
617 using DNA identification of prey. *Deep-Sea Research Part I: Oceanographic Research Papers*
618 57:923-930. https://doi.org/10.1016/j.dsr.2010.02.006
619 Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460-
620 2461. https://doi.org/10.1093/bioinformatics/btq461
621 Egholm M, Buchardt O, Christensen L, Behrens C, Freier SM, Driver DA, Berg RH, Kim SK, Norden B, and
622 Nielsen PE. 1993. PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick
623 hydrogen-bonding rules. *Nature* 365. https://doi.org/10.1038/365566a0
624 Eldredge L. 1980. Two species of Homola (Dromiacea, Homolidae) from Guam. *Micronesica* 16:271-277.
625 FAO. 2016. Fisheries and aquaculture software. FishStatJ - software for fishery statistical time series.
626 *Available at http://www.fao.org/fishery/statistics/software/fishstatj/en* (accessed October
627 2017).
628 Farell EM, and Alexandre G. 2012. Bovine serum albumin further enhances the effects of organic
629 solvents on increased yield of polymerase chain reaction of GC-rich templates. *BMC Research*
630 *Notes* 5:257-257. https://doi.org/10.1186/1756-0500-5-257
631 Francis MP. 1998. New Zealand shark fisheries: Development, size and management. *Marine and*
632 *Freshwater Research* 49:579-591.
633 Geller J, Meyer C, Parker M, and Hawk H. 2013. Redesign of PCR primers for mitochondrial cytochrome c
634 oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular*
635 *Ecology Resources* 13:851-861. https://doi.org/10.1111/1755-0998.12138
636 Gordon DP, Diggles BK, Meisterfeld R, Hollis C, and Buchanan PK. 2012. Phylum Cercozoa: cercomonads,
637 filose testate amoebae, Phaeodaria, plasmodiophoras, Gromia, haplosporidians, and kin. In:
638 Gordon D, ed. *New Zealand inventory of biodiversity*. Christchurch, New Zealand: Canterbury
639 University Press, Christchurch, New Zealand, 233-241.
640 Green SJ, and Minz D. 2005. Suicide polymerase endonuclease restriction, a novel technique for
641 enhancing PCR amplification of minor DNA templates. *Applied and Environmental Microbiology*
642 71:4721-4727. https://doi.org/10.1128/AEM.71.8.4721-4727.2005
643 Guillou L, Alves-de-Souza C, Siano R, and González HE. 2010. The ecological significance of small,
644 eukaryotic parasites in marine systems. *Microbiology Today* May 2010:93-95.
645 Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, Decelle J, del
646 Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann M, Kooistra WHCF, Lara E, Le Bescot N,
647 Logares R, Mahé F, Massana R, Montresor M, Morard R, Not F, Pawlowski J, Probert I, Sauvadet

A-L, Siano R, Stoeck T, Vaulot D, Zimmermann P, and Christen R. 2013. The Protist Ribosomal Reference database (PR(2)): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. *Nucleic Acids Research* 41:D597-D604. https://doi.org/10.1093/nar/gks1160

Harms-Tuohy CA, Schizas NV, and Appeldoorn RS. 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 PDN, Ratnasingham S, and deWaard JR. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society B: Biological Sciences* 270:S96-S99. https://doi.org/10.1098/rsbl.2003.0025

Holthuis L. 1991. FAO species catalogue. Marine lobsters of the world: An annotated and illustrated catalogue of species of interest to fisheries known to date. Rome, Italy: FAO.

Horn PL, Forman JS, and Dunn MR. 2012. Dietary partitioning by two sympatric fish species, red cod (*Pseudophycis bachus*) and sea perch (*Helicolenus percoides*), on Chatham Rise, New Zealand. *Marine Biology Research* 8:624-634. https://doi.org/10.1080/17451000.2011.653543

Hughes DJ. 1998. Sea pens and burrowing megafauna (vol.III): an overview of dynamics and sensitivity characteristics for conservation management of Marine SACs. *Available at http://www.ukmarinesac.org.uk/search/sea-pens.htm* (accessed January 2018).

Jeffs A. 2010. Status and challenges of advancing lobster aquaculture globally. *The Marine Biological Association of India* 52:320-326.

Kamio M, Wakabayashi K, Nagai H, and Tanaka Y. 2016. Phyllosomas of smooth fan lobsters (*Ibacus novemdentatus*) encase jellyfish cnidae in peritrophic membranes in their feces. *Plankton and Benthos Research* 11:100-104. https://doi.org/10.3800/pbr.11.100

Kobayashi G, Mizuguchi T, and Matsuoka A. 2014. Isolation and autoxidation profile of fish myoglobin from hoki (*Macruronus magellanicus*). *Fukushima Journal of Medical Science* 60:31-34. https://doi.org/10.5387/fms.2014-5

Kreader CA. 1996. Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein. *Applied and Environmental Microbiology* 62:1102-1106. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC167874/

Kress WJ, García-Robledo C, Uriarte M, and Erickson DL. 2015. DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution* 30:25-35. http://dx.doi.org/10.1016/j.tree.2014.10.008

Lee S, Hartstein ND, and Jeffs A. 2015. Characteristics of faecal and dissolved nitrogen production from tropical spiny lobster, *Panulirus ornatus*. *Aquaculture International* 23:1411-1425. https://doi.org/10.1007/s10499-015-9893-8

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, and Machida RJ. 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* 10:34-34. https://doi.org/10.1186/1742-9994-10-34

Lorenz TC. 2012. Polymerase Chain Reaction: Basic protocol plus troubleshooting and optimization strategies. *Journal of Visualized Experiments*:3998. https://doi.org/10.3791/3998

Luna SM. 2008. *Notophycis marginata* (Günther, 1878) Dwarf codling. *Available at http://www.fishbase.org/summary/2008* (accessed September 2017).

Lv J, Wu S, Zhang Y, Chen Y, Feng C, Yuan X, Jia G, Deng J, Wang C, Wang Q, Mei L, and Lin X. 2014. Assessment of four DNA fragments (COI, 16S rDNA, ITS2, 12S rDNA) for species identification of the Ixodida (Acari: Ixodida). *Parasites & Vectors* 7:93-93. https://doi.org/10.1186/1756-3305-7-93

695 Machida RJ, Leray M, Ho S-L, and Knowlton N. 2017. Metazoan mitochondrial gene sequence reference
696 datasets for taxonomic assignment of environmental samples. *Scientific Data* 4:170027.
697 https://doi.org/10.1038/sdata.2017.27
698 Mahé F, Rognes T, Quince C, de Vargas C, and Dunthorn M. 2014. Swarm: robust and fast clustering
699 method for amplicon-based studies. *PeerJ* 2:e593. https://doi.org/10.7717/peerj.593
700 Major R, and Jeffs A. 2017. Orientation and food search behaviour of a deep sea lobster in turbulent
701 versus laminar odour plumes. *Helgoland Marine Research* 71:9.
702 https://doi.org/10.1186/s10152-017-0489-8
703 Mente E. 2010. Survival, food consumption and growth of Norway lobster (*Nephrops norvegicus*) kept in
704 laboratory conditions. *Integrative Zoology* 5:256-263. https://doi.org/10.1111/j.1749-
705 4877.2010.00211.x
706 Ministry for Primary Industries. 2006a. Barracouta (BAR). *Available at https://fs.fish.govt.nz* (accessed
707 October 2017).
708 Ministry for Primary Industries. 2006b. Fisheries Assessment: *Silver Warehou* (SWA). *Available at*
709 *https://fs.fish.govt.nz* (accessed October 2017).
710 Ministry for Primary Industries. 2008. Fisheries Assessment: Dark Ghost Shark (GSH). *Available at*
711 *https://fs.fish.govt.nz* (accessed October 2017).
712 Molnár K, Ostoros G, Dunams-Morel D, and Rosenthal BM. 2012. *Eimeria* that infect fish are diverse and
713 are related to, but distinct from, those that infect terrestrial vertebrates. *Infection, Genetics and*
714 *Evolution* 12:1810-1815. https://doi.org/10.1016/j.meegid.2012.06.017
715 Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, and Schaffer AA. 2008. Database indexing
716 for production MegaBLAST searches. *Bioinformatics* 24:1757-1764.
717 https://doi.org/10.1093/bioinformatics/btn322
718 O'Rorke R, Jeffs A, Wang M, Waite A, Beckley LE, and D. Lavery S. 2015. Spinning in different directions:
719 Western rock lobster larval condition varies with eddy polarity, but does their diet? *Journal of*
720 *Plankton Research* 37:542–553. https://doi.org/10.1093/plankt/fbv026
721 O'Rorke R, Lavery S, Chow S, Takeyama H, Tsai P, Beckley LE, Thompson PA, Waite AM, and Jeffs AG.
722 2012. Determining the diet of larvae of western rock lobster (*Panulirus cygnus*) using high-
723 throughput DNA sequencing techniques. *PLoS ONE* 7:e42757.
724 https://doi.org/10.1371/journal.pone.0042757
725 O'Rorke R, Lavery S, Wang M, Nodder SD, and Jeffs AG. 2014. Determining the diet of larvae of the red
726 rock lobster (*Jasus edwardsii*) using high-throughput DNA sequencing techniques. *Marine*
727 *Biology* 161:551-563. https://doi.org/10.1007/s00227-013-2357-7
728 OBIS. 2017. Ocean Biogeographic Information System. *Available at www.iobis.org* (accessed October
729 2017).
730 Ono A, and Moteki M. 2013. Spatial distributions and population dynamics of two salp species, *Ihlea*
731 *racovitzai* and *Salpa thompsoni*, in the waters north of Lützow-Holm Bay (East Antarctica) during
732 austral summers of 2005 and 2006. *Polar Biology* 36:807-817. https://doi.org/10.1007/s00300-
733 013-1305-9
734 Orum H, Nielsen PE, Egholm M, Berg RH, Buchardt O, and Stanley C. 1993. Single base pair mutation
735 analysis by PNA directed PCR clamping. *Nucleic Acids Research* 21:5332-5336.
736 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC310567/
737 Pakhomov EA, Dubischar CD, Hunt BPV, Strass V, Cisewski B, Siegel V, von Harbou L, Gurney L, Kitchener
738 J, and Bathmann U. 2011. Biology and life cycles of pelagic tunicates in the Lazarev Sea,
739 Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 58:1677-1689.
740 https://doi.org/10.1016/j.dsr2.2010.11.014

741 Piñol J, Andrés VS, Clare EL, Mir G, and Symondson WOC. 2014. A pragmatic approach to the analysis of
742 diets of generalist predators: the use of next-generation sequencing with no blocking probes.
743 *Molecular Ecology Resources* 14:18-26. https://doi.org/10.1111/1755-0998.12156
744 Polz MF, and Cavanaugh CM. 1998. Bias in template-to-product ratios in multitemplate PCR. *Applied and*
745 *Environmental Microbiology* 64:3724-3730.
746 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC106531/
747 Pompanon F, Deagle BE, Symondson WOC, Brown DS, Jarman SN, and Taberlet P. 2012. Who is eating
748 what: Diet assessment using next generation sequencing. *Molecular Ecology* 21:1931-1950.
749 https://doi.org/10.1111/j.1365-294X.2011.05403.x
750 Priede IG. 2017. *Deep-sea Fishes: Biology, Diversity, Ecology and Fisheries*. Cornwall, UK: Cambridge
751 University Press.
752 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, and Glöckner FO. 2013. The SILVA
753 ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic*
754 *Acids Research* 41:D590-D596. https://doi.org/10.1093/nar/gks1219
755 R Core Team. 2017. R: A language and environment for statistical computing. 1.0.143 ed. Vienna,
756 Austria: R Foundation for Statistical Computing.
757 Rådström P, Knutsson R, Wolffs P, Lövenklev M, and Löfström C. 2004. Pre-PCR processing. *Molecular*
758 *Biotechnology* 26:133-146. https://doi.org/10.1385/mb:26:2:133
759 Rognes T, Flouri T, Nichols B, Quince C, and Mahé F. 2016. VSEARCH: a versatile open source tool for
760 metagenomics. *PeerJ* 4:e2584. https://doi.org/10.7717/peerj.2584
761 Rossen L, Nørskov P, Holmstrøm K, and Rasmussen OF. 1992. Inhibition of PCR by components of food
762 samples, microbial diagnostic assays and DNA-extraction solutions. *International Journal of Food*
763 *Microbiology* 17:37-45. https://doi.org/10.1016/0168-1605(92)90017-W
764 Rotllant G, Mente E, Gisbert E, and Karapanagiotidis IT. 2014. Effects of different diets on the digestive
765 physiology of adult Norway lobster *Nephrops norvegicus*. *Journal of Shellfish Research* 33:1-9.
766 https://doi.org/10.2983/035.033.0101
767 Sahlmann C, Chan TY, and Chan BKK. 2011. Feeding modes of deep-sea lobsters (Crustacea: Decapoda:
768 Nephropidae and Palinuridae) in Northwest Pacific waters: Functional morphology of
769 mouthparts, feeding behaviour and gut content analysis. *Zoologischer Anzeiger* 250:55-66.
770 https://doi.org/10.1016/j.jcz.2010.11.003
771 Schrader C, Schielke A, Ellerbroek L, and Johne R. 2012. PCR inhibitors – occurrence, properties and
772 removal. *Journal of Applied Microbiology* 113:1014-1026. https://doi.org/10.1111/j.1365-
773 2672.2012.05384.x
774 Seafood New Zealand. 2017. New Zealand Seafood Exports: Seafood exports by species by country.
775 *Available at https://www.seafoodnewzealand.org.nz/publications/export-information/*
776 (accessed August 2017).
777 Simon CJ, and Jeffs A. 2008. Feeding and gut evacuation of cultured juvenile spiny lobsters, *Jasus*
778 *edwardsii*. *Aquaculture* 280:211-219. https://doi.org/10.1016/j.aquaculture.2008.05.019
779 Sousa LL, Xavier R, Costa V, Humphries NE, Trueman C, Rosa R, Sims DW, and Queiroz N. 2016. DNA
780 barcoding identifies a cosmopolitan diet in the ocean sunfish. *Scientific Reports* 6:28762.
781 http://dx.doi.org/10.1038/srep28762
782 Srivathsan A, Ang A, Vogler AP, and Meier R. 2016. Fecal metagenomics for the simultaneous
783 assessment of diet, parasites, and population genetics of an understudied primate. *Frontiers in*
784 *Zoology* 13. https://doi.org/10.1186/s12983-016-0150-4
785 Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, and Fontaneto D. 2012. The widely used
786 small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of
787 the meiofauna. *Proceedings of the National Academy of Sciences of the United States of America*
788 109:16208-16212. https://doi.org/10.1073/pnas.1209160109

789 Tong S, Nygaard K, Bernard C, Vørs N, and Patterson D. 1998. Heterotrophic flagellates from the water
790 column in Port Jackson, Sydney, Australia. *European Journal of Protistology* 34:162-194.
791 https://doi.org/10.1016/S0932-4739(98)80027-8
792 Troedsson C, Simonelli P, Nägele V, Nejstgaard JC, and Frischer ME. 2009. Quantification of copepod gut
793 content by differential length amplification quantitative PCR (dla-qPCR). *Marine Biology*
794 156:253-259. https://doi.org/10.1007/s00227-008-1079-8
795 Tshudy D, Chan TY, and Sorhannus U. 2007. Morphology based cladistic analysis of *Metanephrops*: The
796 most diverse extant genus of clawed lobster (Nephropidae). *Journal of Crustacean Biology*
797 27:463-476. https://doi.org/10.1651/S-2777.1
798 Vestheim H, and Jarman SN. 2008. Blocking primers to enhance PCR amplification of rare sequences in
799 mixed samples – a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology* 5:12.
800 https://doi.org/10.1186/1742-9994-5-12
801 Wahle RA, Tshudy D, Cobb JS, Factor J, and Jaini M. 2012. Infraorder Astacidea Latreille, 1802 p.p.: the
802 marine clawed lobsters. In: Schram FR, and von Vaupel Klein JC, eds. *Treatise on Zoology -*
803 *Anatomy, Taxonomy, Biology The Crustacea, Volume 9 Part B*. Boston: Brill, 3-108.
804 Wang Q, Garrity GM, Tiedje JM, and Cole JR. 2007. Naive Bayesian classifier for rapid assignment of
805 rRNA sequences into the new bacterial taxonomy. *Applied Environmental Microbiology* 73:5261-
806 5267. https://doi.org/10.1128/aem.00062-07
807 Watts GS. 2014. Interpreting nanodrop (spectrophotometric) results. *Available at*
808 *http://www.u.arizona.edu/~gwatts/azcc/InterpretingSpec.pdf* (accessed May 2018).
809 Wickham H. 2017. Tidyverse: Easily Install and Load 'Tidyverse' Packages. 1.1.1 ed. p R package.
810 Wilkens S, Ahyong, S. 2015. Coastal Crabs: A guide to the crabs of New Zealand. *Available at*
811 *https://www.niwa.co.nz/coasts-and-oceans/marine-identification-guides-and-fact-*
812 *sheets/Coastal%20Crabs* (accessed October 2017).
813 Williams G, Tracey, D., Mackay, E. 2014. Pennatulacea (sea pens) Descriptions for the New Zeaand
814 Region. A field guide of commonly sampled New Zealand sea pens including illustrations
815 highlighting technical terms and sea pen morphology. New Zealand Aquatic Environment and
816 Biodiversity Report No 132. Wellington. p 22.
817 Williams GC. 1995. Living genera of sea pens (Coelenterata: Octocorallia: Pennatulacea): illustrated key
818 and synopses. *Zoological Journal of the Linnean Society* 113:93-140.
819 https://doi.org/10.1111/j.1096-3642.1995.tb00929.x
820 Wolffs P, Grage H, Hagberg O, and Rådström P. 2004. Impact of DNA polymerases and their buffer
821 systems on quantitative real-time PCR. *Journal of Clinical Microbiology* 42:408-411.
822 https://doi.org/10.1128/JCM.42.1.408-411.2004
823 Wu S, Xiong J, and Yu Y. 2015. Taxonomic resolutions based on 18S rRNA genes: A case study of subclass
824 Copepoda. *PLoS ONE* 10:e0131498. https://doi.org/10.1371/journal.pone.0131498
825 Young JW, Hunt BPV, Cook TR, Llopiz JK, Hazen EL, Pethybridge HR, Ceccarelli D, Lorrain A, Olson RJ,
826 Allain V, Menkes C, Patterson T, Nicol S, Lehodey P, Kloser RJ, Arrizabalaga H, and Anela Choy C.
827 2015. The trophodynamics of marine top predators: Current knowledge, recent advances and
828 challenges. *Deep Sea Research Part II: Topical Studies in Oceanography* 113:170-187.
829 http://doi.org/10.1016/j.dsr2.2014.05.015
830 Zhan A, Hulák M, Sylvester F, Huang X, Adebayo AA, Abbott CL, Adamowicz SJ, Heath DD, Cristescu ME,
831 and MacIsaac HJ. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in
832 aquatic communities. *Methods in Ecology and Evolution* 4:558-565.
833 https://doi.org/10.1111/2041-210X.12037

834

Table 1(on next page)

Universal primer pairs used to target the selected regions of the COI and 18S genes.

The COI primers, mlCOIintF and jgHCO2198, target a 313 base pair (bp) region. The 18S rRNA primers, Uni18SF and Uni18SR, target a 425 bp V4 variable region. Illumina Nextera™ library adapters (NexAd) have been added to the primers and are underlined.

Primer Name	Target	Sequence (adapters underlined)
mlCOIintF_NexAd (Forward)	COI 313 bp region	5' TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GGW ACW GGW TGA ACW GTW TAY CCY CC 3'
jgHCO2198_NexAd (Reverse)	COI 313 bp region	5' GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTA IAC YTC IGG RTG ICC RAA RAA YCA 3'
Uni18SF_NexAd (Forward)	18S V4 425 bp region	5' TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AGG GCA AKY CTG GTG CCA GC 3'
Uni18SR_NexAd (Reverse)	18S V4 425 bp region	5' GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGR CGG TAT CTR ATC GYC TT 3'

1

Table 2(on next page)

DNA metabarcoding samples, comprised of six individuals and one DNA negative control (Fro1 and Fro2 are two individuals).

Sample	Individuals	Digesta Source	PCR Reagent	Preservation
1	70.9	Hindgut	Bioline	Ethanol
2	70.9	Foregut	Bioline	Ethanol
3	70.9	Hindgut	Platinum *Taq*	Ethanol
4	70.9	Foregut	Platinum *Taq*	Ethanol
5	70.2	Hindgut	Bioline	Ethanol
6	70.2	Foregut	Bioline	Ethanol
7	70.3	Foregut	Bioline	Ethanol
8	Fro1 & Fro 2	Foregut	Bioline	Frozen
9	Fro3	Foregut and Hindgut	Bioline	Frozen
10	DNA Negative	NA	Bioline	NA

1

Table 3(on next page)

A comparison of the number of assigned diet taxa using the Bioline and the Platinum *Taq* reactions.

The taxa identified in each of the reactions pertain to diet taxa, therefore taxa identified as lobster (Astacidea), matching the DNA negative sample or identified as terrestrial were removed. The 'Diet Hit Count' refers to the number of sequence reads matching a potential dietary reference sequences (identified as diet taxa) in the databases. The 'Exclusive Taxa' are those taxa found only in one category i.e., in the Bioline reaction but not in the Platinum *Taq* reaction.

	Taxa Identified in the Bioline Reaction (Diet Hit Count)	Taxa Identified in the Platinum *Taq* Reaction (Diet Hit Count)
Midori/NCBI (COI)	9 (648)	5 (62)
NCBI/PR2/SILVA (18S)	2 (10)	1 (6)
Total Taxa	11 (658)	6 (68)
Total Exclusive Taxa	6 (190)	1 (13)

1

Table 4(on next page)

A comparison of the total and exclusive diet taxa identified in the foregut and hindgut.

The taxa identified in each of the reactions pertain to diet taxa, therefore taxa identified as lobster (Astacidea), matching the DNA negative sample or identified as terrestrial were removed. The 'Diet Hit Count' refers to the number of sequence reads matching a potential dietary reference sequences (identified as diet taxa) in the databases. The 'Exclusive Taxa' are those taxa found only in one category i.e., in the foregut but not in the hindgut.

Database	Total Taxa Identified (Diet Hit Count) in Foregut	Total Taxa Identified (Diet Hit Count) in Hindgut	Exclusive Taxa Identified in Foregut	Exclusive Taxa Identified in Hindgut
Midori/NCBI (COI)	3 (448)	12 (274)	2	11
NCBI/PR2/ SILVA (18S)	14 (736)	2 (14)	13	1

1

Table 5(on next page)

A comparison of the digesta amplification from ethanol and frozen preserved individuals.

The ‘Genus/Species Hit Count’ refers to the number of sequence reads that matched a high degree of taxonomic resolution (genus and/or species level). The ‘Diet Hit Count’ refers to the number of sequence reads matching potential dietary reference sequences in the databases. The ‘Diet Hit Count (%)’ is the percentage of the ‘Diet Hit Count’ out of the total ‘Genus/Species Hit Count’.

Database and Preservation	Genus/Species Hit Count	Diet Hit Count	Diet Hit Count (%)
Midori /NCBI (COI)			
Ethanol	171652	10611	6
Frozen	36691	0	0
NCBI/PR2/SILVA (18S)			
Ethanol	165566	2028	1
Frozen	94218	902	1

1

Figure 1

Taxa identified from the diet of *M. challengeri* using the Midori and NCBI databases for the COI sequences.

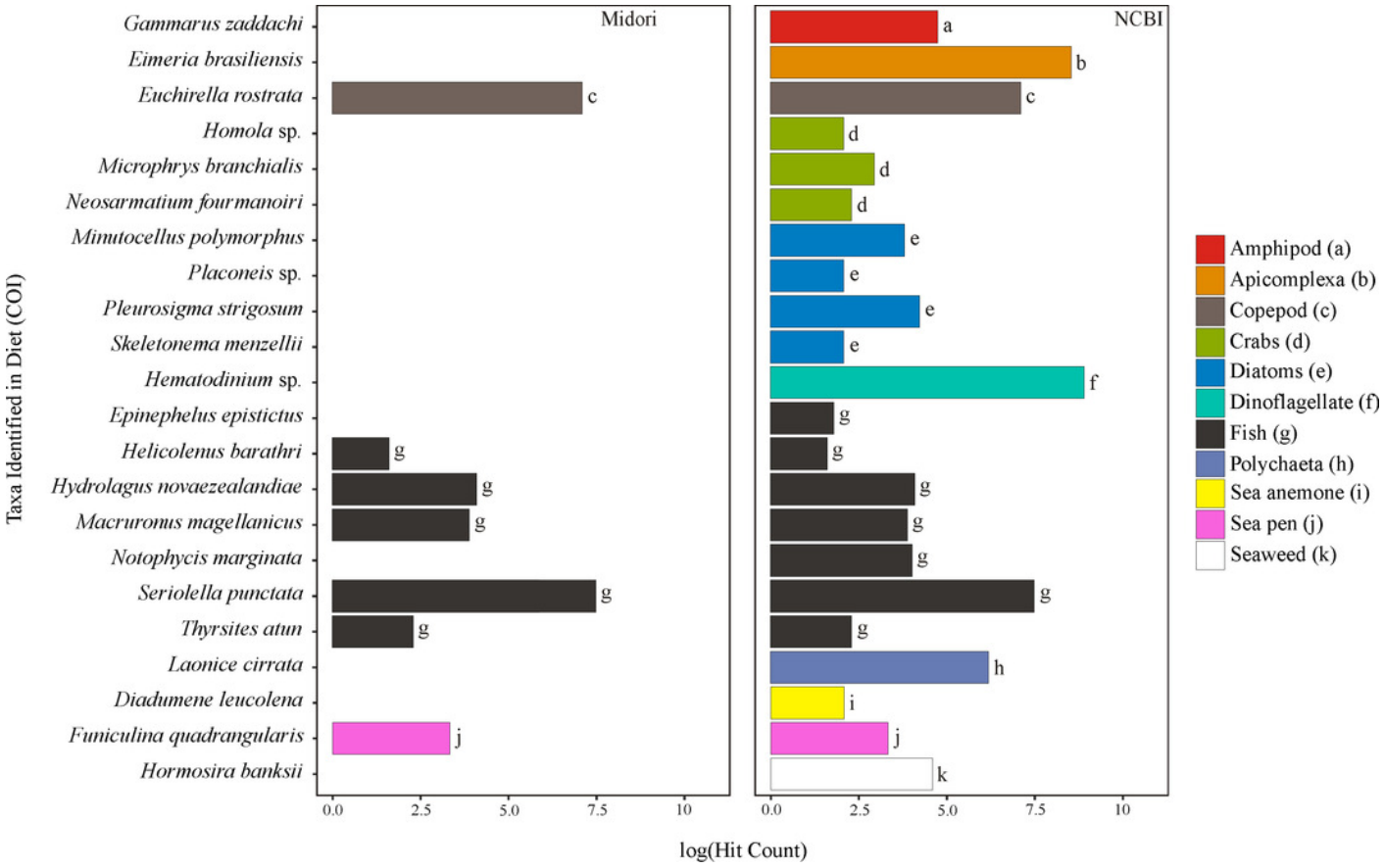

Figure 2

Taxa identified from the diet of *M. challengeri* using the PR2, SILVA and NCBI databases for the 18S sequences.