

Nanopore short-read sequencing: A quick, cost-effective and accurate method for DNA metabarcoding

Aimee L. van der Reis¹  | Lynnath E. Beckley²  | M. Pilar Olivar³  | Andrew G. Jeffs¹ 

¹Institute of Marine Science and School of Biological Sciences, University of Auckland, Auckland, New Zealand

²Environmental and Conservation Sciences, Murdoch University, Perth, Western Australia, Australia

³Institut de Ciències del Mar, CSIC, Passeig Marítim de la Barceloneta, Barcelona, Spain

Correspondence

Aimee L. van der Reis, Institute of Marine Science and School of Biological Sciences, University of Auckland, Auckland, New Zealand.

Email: avan398@aucklanduni.ac.nz

Funding information

Australian Marine National Facility, Grant/Award Number: IN2019_V03

Abstract

Dietary and predator–prey studies are more frequently relying on DNA metabarcoding methods, typically achieving results that have a better taxonomic resolution (e.g., species-level) than previous methods. With the continuous advancement in sequencing technology, what was previously accessible only as a large, fixed structure in a laboratory, which had a limited number of users, has now advanced to a small and readily usable device. In this study, we used the gut (content and lining) from juvenile lanternfish (*Hygophum*) specimens to compare the short-read sequencing capability of the portable Nanopore MinION with the Illumina MiSeq. Primers common in dietary DNA metabarcoding work (COI “Leray primers” and 18S rRNA V4 “Zhan primers”) were used, with an additional comparison of cost-effective COI “Lobo primers” (targeting the same COI fragment) for the proficiency in species detection of a broad range of taxa. Our results indicate high congruency between sequencing machines for, not only taxonomic assignments, but also relative read abundance of the main dietary items. We also identified that Nanopore sequencing is more cost-effective. The Lobo primers are comparable to that of Leray, but substantially reduce the primer set price without compromising detection of taxa. Using both COI and 18S broadened the taxonomic scope, providing greater prey detection. Overall, this preliminary study was successful in creating a foundation for future dietary work involving larvae and transformation stage fishes whereby the content of the gut need not be separated from the gut lining to detect prey. The *Hygophum* diet detected here aligns with previous research that suggests the main dietary items to be calanoid copepods, but using molecular methods, soft prey was more readily identified compared to studies using visual methods of identification of dietary items. Overall, this study found that Nanopore sequencing is suitable for short-read DNA metabarcoding and can provide rapid access to sequencing results.

KEYWORDS

18S rRNA, COI, diet analysis, DNA metabarcoding, *Hygophum*, Illumina, lanternfish, Nanopore

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Environmental DNA* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

DNA metabarcoding is a molecular method that targets a conserved gene region (e.g., COI and 18S rRNA V4) that also has adequate DNA sequence variation to allow discrimination between closely related taxa. Typically, this method is used in large-scale species identification studies where the source material is composed of many species. This technique has been increasingly used in dietary studies where the primary samples are gut content, regurgitated and/or fecal matter (e.g., Berry et al., 2015; Carroll et al., 2019; De Barba et al., 2014; Kartzinel et al., 2015; Oehm et al., 2017; van der Reis et al., 2018). Moreover, it is a promising method to investigate predator–prey interactions and thereby reveal potential trophic interactions in complex food webs, as often these interactions cannot be directly observed in the natural environment because of limited accessibility, complex logistics and/or cost, for example, the deep sea (Compson et al., 2019; van der Reis et al., 2020; Zamora-Terol et al., 2020).

Until recently, the most frequent way of analyzing gut contents of predators has been manual and visual sorting using taxonomic methods that rely on morphology to differentiate the identity of the species that had been consumed. However, the species consumed cannot always be identified from the gut content due to insufficient remaining morphological structure as a result of mastication and/or rapid digestion (e.g., soft bodied organisms such as gelatinous animals; Arai et al., 2003; O'Rorke, Lavery, Chow, et al., 2012), which can often lead to the importance of certain dietary species being under-estimated. Furthermore, the smaller the predator (e.g., larval and transformation fish stages), the more difficult it is to identify its increasingly small prey. This is one of the major advantages of DNA metabarcoding, which can deliver taxonomic identification irrespective of the morphological state (or size) of the gut contents. Also, this molecular method of identification does not require taxonomic expertise and is relatively rapid, cost-effective, and accurate, but is not without some limitations (e.g., reviewed in: O'Rorke, Lavery, & Jeffs, 2012; Ruppert et al., 2019; van der Loos & Nijland, 2021).

Rapid initial DNA metabarcoding work (e.g., DNA extraction and amplification), such as whilst working in the field, is challenging because it requires a translocation of fundamental laboratory equipment. However, any subsequent work that follows this, such as DNA sequencing, has been slowed or impractical in the field situation because the sequencing devices were too expensive to risk translocation, and a trained technician was needed for their operation. However, the advent of small and highly portable sequencing devices (e.g., Oxford Nanopore Technologies—ONT—MinION devices; Jain et al., 2016) has led to effective and efficient real-time sequencing, which makes sequencing possible on-site with results available within 24–48h (Runtuwene et al., 2019). These devices are generally used for whole genome sequencing and community metagenomics because in these samples the DNA is intact and of relatively high quality, and thus capitalizes on the ability of the device to reliably sequence very long reads. DNA metabarcoding does feature in the Nanopore suite but is targeted to the full-length 16S rRNA (i.e., EPI2ME 16S workflow). In contrast, these devices were not designed to be a major sequencing

platform for short-read environmental or dietary DNA where the DNA is more likely to be in a degraded state (Deiner et al., 2017; Pompanon et al., 2012). However, recent studies suggest that, despite the lower recorded sequencing accuracy (Santos et al., 2020), these portable devices are highly consistent compared to other sequencing devices that have higher accuracy for shorter read lengths (e.g., Illumina MiSeq; Chang, Ip, Ng, & Huang, 2020), when comparing molecular operational taxonomic units (MOTUs) (Chang, Ip, Bauman, & Huang, 2020). Furthermore, Nanopore basecalling is continuously developing and improving in accuracy (e.g., Zeng et al., 2020).

In this preliminary DNA metabarcoding study we explore the proficiency of using the ONT MinION Mk1B for diet analysis, targeting species-level resolution using short-read sequencing (e.g., COI ca. 313bp; Leray et al., 2013), and compare it to the results from standard Illumina MiSeq. In doing so, a rudimentary, yet adequate, bioinformatics pipeline is assembled from existing bioinformatic tools for the MinION sequence filtering. We also assess the capability of two different primer pair sets, which target the same popular “mini” COI region differing only in the reverse primer, to identify dietary components and, moreover, use an additional primer pair targeting the 18S rRNA V4 region to broaden taxa detection. Here, the diet of juvenile (ca. 2cm) lanternfish, *Hygophum* sp., sampled in the eastern Indian Ocean were investigated. Additional objectives of this work were to determine if there was a difference in the variety of identified taxa when comparing DNA from gut content versus the gut lining they were extracted from, and to compare the cost-effectiveness of both sequencing technologies. The first additional objective was aimed at ensuring future work using molecular methods to identify prey that includes the gut lining of smaller mesopelagic fish samples will not be dominated by host DNA.

2 | METHODS AND MATERIALS

2.1 | Sample collection

Lanternfish samples were collected during the 2nd International Indian Ocean Expedition (<https://iioe-2.incois.gov.in/>), on board a 94m oceanographic research vessel (*RV Investigator*, CSRIO, Hobart, Tasmania), which took place between 13 May 2019 and 14 June 2019. The expedition's sampling stations were along the 110° E meridian off western Australia. The samples were collected either by means of a neuston net surface tow (1mm mesh size) or captured during an EZ multinet tow (0.5mm mesh size). All lanternfish that were caught were microscopically identified to the lowest possible taxonomic level on board (pertinent species identification literature: Hulley & Paxton, 2016a, 2016b; Paxton et al., 1995; Paxton & Williams, 2019). Fish specimens set aside for dietary analysis were stored in 70% ethanol and subsequently placed into the –20°C freezer on board. For this preliminary study of comparative techniques a single lanternfish species, visually identified as a species of *Hygophum*, was chosen and selected samples collected from stations 3, 12 and 13 were used (Table S1).

2.2 | Sample dissection, DNA extraction, and amplification

The total length of each lanternfish specimen was recorded before microscopic dissection took place in a UV sterilized laminar flow cabinet. Using sterilized fine tweezers, for each sample, the foregut and the hindgut, if possible, were carefully removed. Subsequently, any gut content that could be separated from the gut lining was removed and the gut lining and gut content were each stored separately in 80% ethanol at -20°C .

Shortly after dissection, the DNA was extracted from the gut content and gut lining separately using the NucleoSpin XS Tissue DNA extraction kit following the manufacturer's instructions (Macherey-Nagel), in a UV sterilized laminar flow cabinet. The polymerase chain reactions (PCRs) were done in triplicate using MyTaq Red Mix (BioLigne; Meridian Bioscience) master mix; 12.5 μl MyTaq Red Mix, 0.5 μl of each primer (10 μM), 8.5 μl UltraPure DNase/RNase-Free Distilled Water (Invitrogen; Thermo Fisher Scientific), 1 μl DNA and 2 μl BSA (1%). Negative controls were included in the DNA extractions (extraction blank - no tissue added) and subsequently in the PCRs, to check for possible contamination. Possible contamination was also monitored by including a PCR blank (no DNA added) in every PCR run.

For the recovery of dietary DNA three different universal primer pairs were used, two targeting the mitochondrial DNA COI region and another targeting the nuclear 18S rRNA V4 region (Table 1). The COI primer sets targeted the same area of ~ 313 base pairs (bp) in COI and only differed by the reverse primer. The reverse general primer jgHCO2198 (Geller et al., 2013) contains the degenerate base inosine (I), which has affinity to nucleotides, in the order of stability, I-C > I-A > I-T \approx I-G and thus somewhat increases specificity (Case-Green & Southern, 1994; Martin et al., 1985). This reverse primer is often used in conjunction with M1COIintF (Leray et al., 2013) in DNA metabarcoding studies since the versatility of this primer set was revealed. This primer set has commonly become known as the "Leray primers," and thus will be referred to as such in this study. The second

COI reverse primer, LoboR1 (Lobo et al., 2013), replaces the first occurrence of I with an A and the remaining Is with W (W = A/T). Like jgHCO2198, LoboR1 was designed to cover a wide variety of marine taxa replacing the "Folmer primers" (Folmer et al., 1994) with more universal ones, and the combination of M1COIintF and LoboR1 will be referred to as the "Lobo primers." The 18S rRNA V4 primer pair targets ~ 425 bp (referred to as "Zhan primers"; Zhan et al., 2013). The PCR products were run on a 1.6% agarose gel, and visualized using Gel Red (Biotium), in a Gel Doc XR+ (Bio-Rad).

2.3 | PCR product clean-up

The PCR triplicates were pooled together by sample before proceeding with DNA clean-up using Agencourt AMPure XP (Beckman Coulter). The clean-up followed the Illumina protocol for PCR clean-up (Illumina, 2013) for all gene regions. The concentration of the purified PCR products was determined using Qubit dsDNA HS Assay Kit (Invitrogen) following the manufacturer's instructions.

2.4 | Nanopore preparation, sequencing, bioinformatics, and filtering

The Nanopore native barcoding amplicons protocol was followed as per the manufacturer's instructions with minor changes (Table S2). The reaction volume was halved for all PCR amplicons and reagent volumes for "end-prep" and "native barcode ligation" steps, to make the sequencing more cost-effective. However, for the "adapter ligation and clean-up" and "priming and loading the SpotON flow cell" steps the volumes were kept as per the manufacturer's instructions. Two barcodes were used per sample so that each barcode represented a different COI primer set, which otherwise would not be able to be called out separately when demultiplexing. A 2 h sequencing trial run on a MinION Mk1b, using R9.4.1 flow cell technology, with sample 47F was done to ensure that halving reagent volumes

TABLE 1 Primer pairs, gene targeted, expected amplicon size and PCR protocols followed. The differences in the COI reverse primers have been indicated by underlining relevant bases. For Illumina sequencing, Nextera adapters were added to the 5' end of the primers (Illumina, 2013).

Gene region	Primer pair	PCR protocol
COI (313bp)	F: M1COIintF (Leray et al., 2013) GGWACWGGWTGAACWGTWTAYCCYCC R: jgHCO2198 (Geller et al., 2013) TA <u>I</u> ACYTC <u>I</u> GGRTG <u>I</u> CCRAARAAYCA R: LoboR1 (Lobo et al., 2013) TAA <u>A</u> ACYTC <u>W</u> GGRTG <u>W</u> CCRAARAAYCA	Lobo et al. (2013) (modified) 94°C 60s 35 \times [94°C 30s; 54°C 90s; 72°C 60s] 72°C 5 min
18S V4 (425 bp)	F: Uni18SF (Zhan et al., 2013) AGGGCAAKYCTGGTGCCAGC R: Uni18SR (Zhan et al., 2013) GRCGGTATCTRATCGYCTT	Clarke et al. (2017) 98°C 30s 30 \times [98°C 5 s; 53°C 20s; 72°C 20s] 72°C 5 min
Nextera adapters	F: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG R: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	

and amplicon volumes worked adequately. The trial was successful and thus the remaining samples were sequenced over a 24 h period. The flow cell was washed at the end of each run following the manufacturer's instructions.

The raw data (fast5 file format) bases were called using ONT sequencing software Guppy v4.2.2. The resulting fastq files, per barcode, were concatenated into a single fastq file. Cutadapt v3.3 (Martin, 2011) was used to search for primer matches against the sequences (allowing for a 10% error, but ensuring a full overlap between primer and sequence). Once matched, the primers were trimmed in two steps. The first step was done using the data in the original read direction and the second was using the remaining untrimmed sequences and reverse complementing them (SeqKit v0.15.0; Shen et al., 2016), then searching for the primers again. These trimmed reads were then concatenated by their respective barcodes. It was noted what the total number of reads retained were and if the reads were trimmed based on the forward or the reverse primer. Furthermore, all trimmed reads then underwent quality filtering whereby the 3' end of the read was truncated when a base was identified with a quality score < 10. The reads were then filtered for length (COI \geq 263bp and \leq 363bp; 18S V4 \geq 375bp and \leq 475bp) and converted to fasta format.

VSEARCH (Rognes et al., 2016) in Qiime2 (v2021.2; Bolyen et al., 2019) was used to dereplicate sequences, cluster the sequences de novo and perform chimera detection on the consensus sequences. Research has indicated that VSEARCH is appropriate for clustering Nanopore data when using shorter sequences that are similar to the read length used in this study (Calus et al., 2018). Clustering at 95% was done to account for the suggested 5% error rate (Santos et al., 2020). The consensus sequences were then run through scikit-learn (within Qiime2; Pedregosa et al., 2011), which assigned taxonomy with the greatest possible resolution to every consensus read at a minimum confidence threshold of 80% using SILVA v138.1 (18S rRNA; Quast et al., 2013), MetaZooGene v2.2 (MZG; COI; Bucklin et al., 2021) and MIDORI v239 (COI; Machida et al., 2017). However, it was found that MIDORI was not satisfactory for COI assignments and OTUs with high numbers of reads were not being assigned when using MZG (when sequences were "blasted" using GenBank they were assigned to zooplankton), and poor resolution was achieved when using SILVA for 18S. It was also found that when doing taxonomic assignments using these curated databases (i.e., MIDORI, MZG and SILVA), the reads extracted in reverse orientation were rarely assigned. Thus, GenBank's MegaBLAST option in the BLASTn suite (Morgulis et al., 2008) was used for assignments using the BLAST v2021-05 database (Benson et al., 2013), minimum E-value threshold of 0.001 and percentage identity of \geq 80%.

Once OTUs were assigned, the taxonomic identity was run through the World Register of Marine Species (WoRMS; Horton et al., 2022) and only the assignments taxonomically classified within the marine database were retained (i.e., removing any non-marine assignments/terrestrial contamination). OTUs were further filtered out if they did not meet the minimum thresholds

of 90% query coverage alignment and 90% percentage identity. Percentage identity was a parameter previously set to 80% for BLAST assignment, but subsequent analyses provided confidence that sufficient data would be retained if increased. OTUs were only retained if they were classed as not being chimeric and to combat sequencing error only those OTUs that had formed clusters with >10 reads were retained (per sample). To avoid any host contamination, all OTUs assigned to ray-finned fishes (class Actinopteri) were removed. Any influence of possible (cross-)contamination or PCR artifacts was reduced by proportional subtraction of the negative control from the respective assignments.

2.5 | Illumina preparation, sequencing, bioinformatics, and filtering

The PCR products were brought to equal molarity, $2 \text{ ng } \mu\text{l}^{-1}$ where possible. Sequencing was done through Auckland Genomics (New Zealand) where indexing, using the Nextera DNA Library Prep Kit and the second round of PCR clean-up occurred before sequencing on an Illumina MiSeq System (2×300 paired-end; single lane). It must be noted that these PCR products were pooled into a larger sampling run, but as our research aim was to compare the accuracy of the assignments with the greater relative read abundance and not the presence of rarer assignments (which may also be influenced by PCR stochasticity), a greater sequencing depth would only have potentially benefitted identifying rarer taxa.

Demultiplexed data were received from Auckland Genomics and Cutadapt was used to trim the forward and reverse primers from the sequences. Primers were removed if an exact sequence match could be found anchored at the beginning of the sequence, and no indels or errors were allowed in the primer sequence. All untrimmed sequences were discarded. Qiime 2 (version 2021.2; Bolyen et al., 2019) was used to visualize the initial sequence quality. DADA2 (within Qiime 2; Callahan et al., 2016) was used for sequence filtering based on quality scores, denoising, merging and chimera formation to ensure only high quality paired-end sequences were retained. Similar attempts to use the curated databases for sequences were made, but using the RDP classifier (Wang et al., 2007). However, this resulted in the same issues experienced previously with OTUs, that is, amplicon sequence variants (ASVs; Callahan et al., 2017) with high read numbers not being assigned a taxonomic identity or poor taxonomic resolution was achieved. Therefore, the resulting ASVs were run through the BLAST database using the parameters as described for the Nanopore MinION data and further filtering followed the same steps as outlined previously. The only step that differed was the omission of filtering for a minimum read depth because Illumina sequencing is extremely accurate (99.9%; Santos et al., 2020) and when investigating those assignments that were singletons, identified in the content of samples 46A (18S) and 46B (18S and COI),

their occurrence was confirmed by being present in the gut lining in higher numbers.

2.6 | Data analyses and sequencing price comparison

All data analyses was done using R v4.1.0 (R Core Team, 2021) and plots were made using ggplot2 v3.3.3 (Wickham, 2016). To determine the cost-effectiveness of the two different sequencing technologies, the per plate costs for Nanopore and Illumina MiSeq sequencing were investigated and contrasted.

3 | RESULTS

3.1 | Nanopore quality filtering

In total 17,021,907 reads were sequenced from the seven fish specimens from a 24 h run (Table S3). Primer removal, Phred quality score (Q score) and length filtering resulted in 2,998,265 reads retained for Leray, 2,865,143 for Lobo and 2,906,050 for Zhan primers. Forward and reverse primer removal from the original and reverse complement reads had an approximate 1:1 ratio for all primer pairs except Leray, which had an approximate 2:1 ratio (forward:reverse) (Figure S1).

Clustering at 95% sequence similarity produced 2,019,624, 1,936,704 and 2,263,944 OTUs for Leray, Lobo and Zhan, respectively. BLAST assigned taxonomic identities to 98% OTUs for Leray (1,963,030 OTUs; 2,928,921 reads), 97% for Lobo (1,873,809; 2,787,351) and ~100% for Zhan primers (2,262,665; 2,904,771). OTUs were retained if they were identified as non-chimeric and met the minimum thresholds of OTU clustering, query coverage alignment and percentage identity. After filtering, a total of 12% (3629 OTUs; 361,963 reads), 12% (3317; 331,425) and 10% (4232; 299,057) of reads remained for Leray, Lobo and Zhan primers, respectively. OTUs that could not be taxonomically classified through the WoRMs database or were identified as fish were removed (Leray: 3312 OTUs and 350,439 reads; Lobo: 2778 and 310,457; Zhan: 1270 and 131,686) (Tables S4 and S5). While Leray and Lobo remained above 10% for reads retained, the removal of fish assignments decreased Zhan reads by ~5%. The relative read abundances were similar within fish samples for both the reads classified as forward and reverse (Figures S2 and S3). Regarding negative controls, there were no reads detected in the 18S (Zhan) control and only one species (*Labidocera detruncata*) was detected in the COI (Lobo and Leray) controls. Proportional subtraction highlighted some taxa, which had "low" read numbers that warranted further investigation to determine if doing only proportional subtraction was adequate. Taking samples that had data for both gut content and lining, a substantial difference was seen in read numbers for *L. detruncata* (e.g., 47F content

had 61,143 reads versus lining with 165 reads; Table S6). Thus, the low read numbers were not false positives and proportional subtraction was found to be adequate to control the possible contamination.

3.2 | Illumina quality filtering

The raw number of Illumina reads for the fish samples were 226,520, 312,018 and 147,242 for Leray, Lobo and Zhan primers, respectively (Table S7). After filtering via DADA2 (including denoising, merging and chimera removal) there were 52% (118,573 reads; 376 ASVs), 50% (154,598; 414) and 45% (66,575; 125) reads remaining for Leray, Lobo and Zhan primers, respectively. BLAST assigned identities to 96% of Leray (360 ASVs; 117,552 reads), 93% of Lobo (383; 152,748) and 100% of Zhan (125; 66,575) ASVs.

After the removal of ASVs that (1) did not meet the minimum thresholds of query coverage alignment and percentage identity, (2) were assigned to host DNA (all fish ASVs removed), or (3) could not be taxonomically classified by the WoRMs database, 32% (173 ASVs; 73,541 reads), 30% (197; 92,531) and 21% (58; 30,750) of reads remained for Leray, Lobo and Zhan primers, respectively (Tables S4 and S5). Again, it was possible to mitigate possible contamination using proportional subtraction from the results obtained from the negative controls. As above, the lower read numbers after subtraction were typically retained as the comparison of the gut content and lining data indicated the detection of the taxa were true positives (Table S6).

3.3 | Comparison of Illumina and Nanopore

For Leray, Lobo and Zhan primers there were 69, 69 and 70 unique BLAST assignments (i.e., typically species-level assignments; Tables S4 and S5) post filtering for the Nanopore OTUs, respectively. In comparison, for Leray, Lobo and Zhan primers, among the Illumina ASVs there were 67, 72 and 40 unique assignments, respectively.

Overall, the assignments made, as well as their respective frequencies were mirrored more closely between Nanopore and Illumina in COI (66 assignments in common, 10 unique to Nanopore and eight unique to Illumina) than 18S (33 assignments common, 37 unique to Nanopore and seven unique to Illumina). The percentage identity was typically lower for the Nanopore assignments, with ~2% difference (standard error of ± 0.8). At genus-level there were 41 genera in common for COI, five genera were unique to Nanopore and one was unique to Illumina (Figure 1). For 18S there were 27 genera in common between the sequencing platforms and Nanopore had 24 unique genera and Illumina had four (Figure 2). The relative read abundance was similar between Nanopore and Illumina for the major dietary genera (Figures 1 and 2).

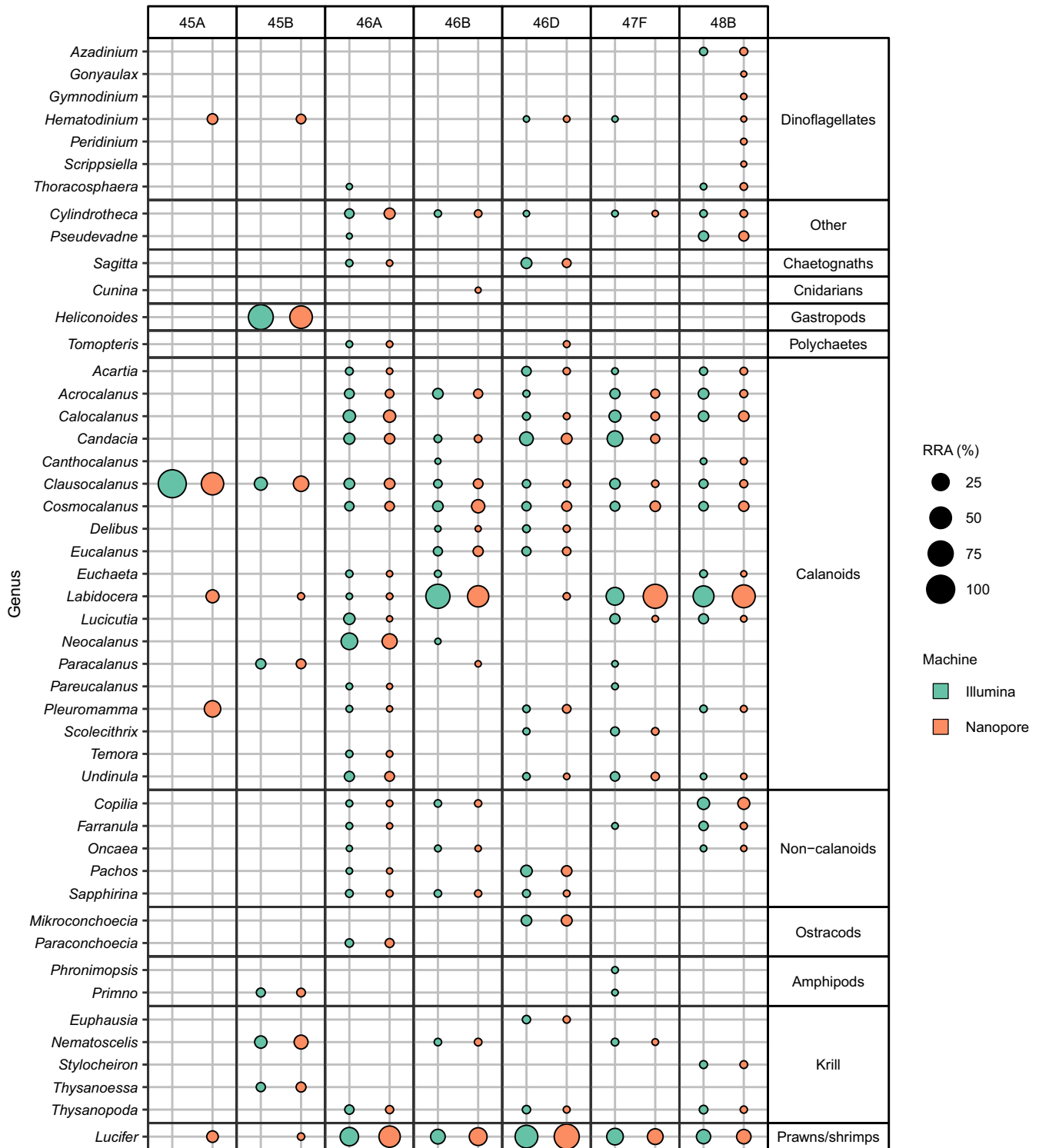


FIGURE 1 Illumina and Nanopore COI BLAST assignments for genera identified from the gut of seven *Hygophum* fish specimens, sampled from the eastern Indian Ocean. Fish samples are individually grouped in columns. Genera are grouped by higher taxonomic classification using common names, genera listed on the left y-axis and common names on the right y-axis. The relative read abundance (RRA) percentage is shown in relation to the area of the circle (i.e., the smaller the circle size, the lower the RRA). Copepods have been split as calanoids and non-calanoids.

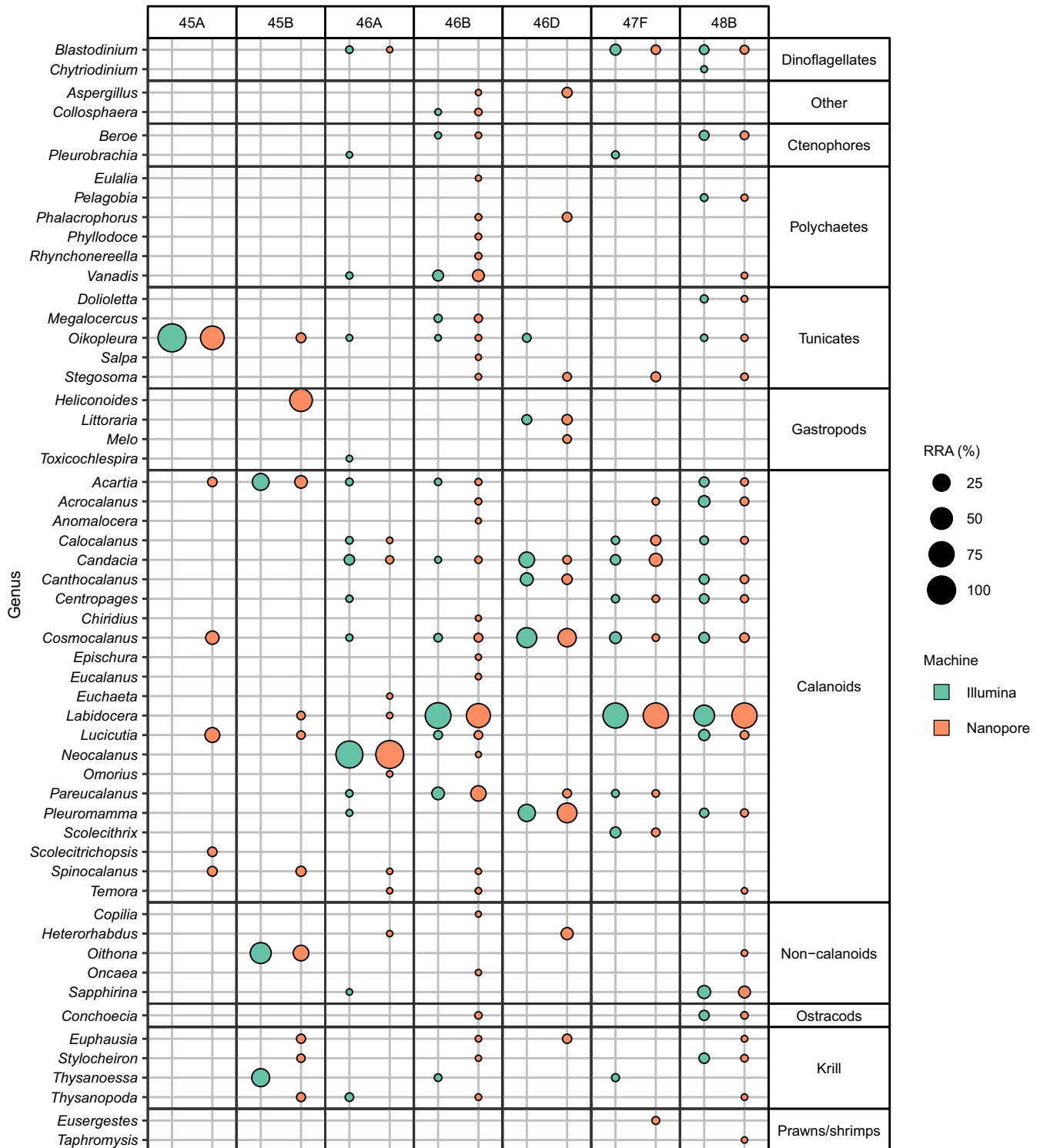


FIGURE 2 Illumina and Nanopore 18S rRNA V4 BLAST assignments for genera identified from the gut of seven *Hygophum* fish specimens, sampled from the eastern Indian Ocean. Fish samples are individually grouped in columns. Genera are grouped by higher taxonomic classification using common names, genera listed on the left y-axis and common names on the right y-axis. The relative read abundance (RRA) percentage is shown in relation to the area of the circle (i.e., the smaller the circle size, the lower the RRA). Copepods have been split as calanoids and non-calanoids.

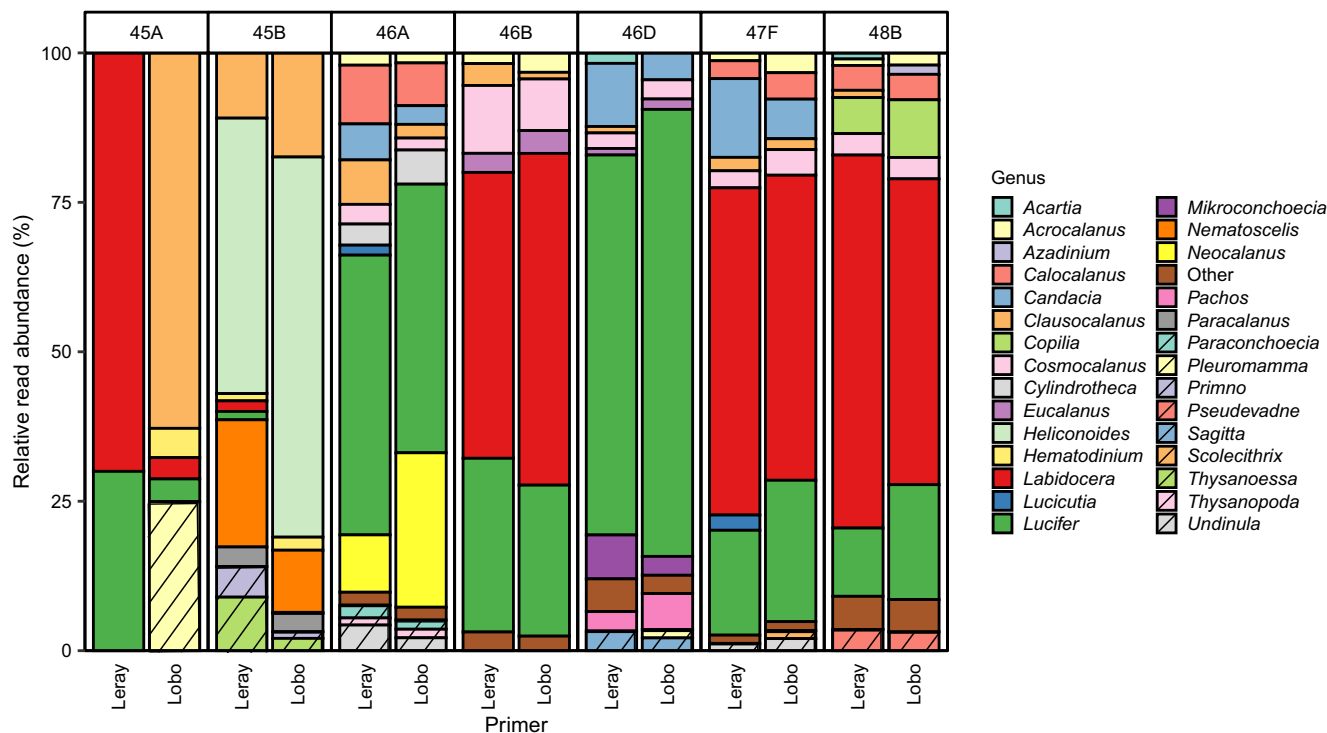


FIGURE 3 Comparison of COI primer pair sets targeting the same gene region, which differ only by their reverse primer. BLAST assignments are shown at genus-level. The relative read abundance (%) is shown on the y-axis and the different primer set results are shown on the x-axis. The results are grouped in columns by their respective samples, shown on top. Any genera with a relative read abundance <2% are represented by “other”.

TABLE 2 The number of BLAST assignments at genus-level that were identified from the lanternfish samples and categorized into either being identified from the gut content and gut lining, unique to the content or unique to the lining. The total number of genera identified from the gut per sample is provided.

Fish ID	Content and lining	Content only	Lining only	Total
46A	22	3	11	36
46B	10	3	30	43
47F	15	1	9	25
48B	27	7	7	41

3.4 | COI primer pair comparison: Leray versus Lobo

Based on the agarose gel results the Lobo primers amplified more readily compared to the Leray primers. There were 62 assignments in common between Leray and Lobo for the Nanopore OTUs, and seven assignments unique to both Leray and Lobo. For Illumina ASVs there were 65 assignments in common between Leray and Lobo, and two and seven assignments unique to Leray and Lobo, respectively. Irrespective of the sequencing device used, there were 45 genera identified in common between primer pairs (Figure 3). Only Lobo detected unique genera (*Scrippsiella* and *Gonyaulax* - dinoflagellates), but the genera were very low with respect to relative read abundance and were only detected in one individual.

3.5 | Lining versus content

Four fish specimens had enough gut content that it could be removed from the gut lining. In general, there were a lower number of organisms that could be assigned from the content compared to the lining. Organisms that were identified in both the content and the lining were generally higher in number than organisms found to be unique to the content or lining alone (Table 2). Fish 46B was an exception where there were 30 unique genera identified in the lining alone and three unique to the content, with only 10 in common.

3.6 | Items detected from the gut

Nanopore results detected 22 genera in common between COI and 18S, with 24 and 29 genera being unique, respectively. For Illumina there were 16 genera in common between COI and 18S, with 26 and 15 genera being unique, respectively. Irrespective of sequencing device used, there were only 23 genera in common between the two gene regions (more than half from the order Calanoida), with a further 24 genera unique to COI and 32 unique to 18S. Some taxa were only identified by a single gene, for example, chaetognath genera were only identified by COI and tunicates only by 18S (Table S8). There were also instances where genera could be identified by both genes, but were only identified by one gene within a sample (Table S8). For example, the calanoid *Lucicutia* could be identified by

both COI and 18S and was identified in six samples, but not always by both genes. In three of the samples *Lucicutia* was identified by 18S, in another two by COI and in one by both COI and 18S.

The most prevalent genera identified from the gut were calanoid copepods. In total 26 genera were identified, but only three were identified in all fish (*Acartia*, *Clausocalanus* and *Labidocera*; Table S8). Other common genera identified included *Lucifer* (prawn), *Oikopleura* (tunicate), *Thysanopoda* (krill) and *Cylindrotheca* (diatom). Detected to lesser extents were amphipods, ostracods, gastropods and radiolarians (e.g., *Collospira*). Organisms that were detected, but were unlikely to be dietary items, included parasites such as *Hematodinium* (dinoflagellate).

3.7 | Sequencing price comparison

Sequencing per plate costs were identified from reputable sources (Table 3). Nanopore sequencing (~\$2100; all prices stated in USD) was cheaper than a standard Illumina MiSeq run (~\$2600), when based on a sequencing run of 96 indexed samples. However, Nanopore reagents would need to be bought in bulk and, thus, the cost was higher if reagents were only purchased for once-off sequencing (~\$4400). A nano Illumina MiSeq run was the cheapest option (~\$900), but would be constrained by sequencing depth (1.6–2 million reads, which is <10% of reads delivered in a standard run; Illumina, 2017) making it suitable for preliminary work where low read depth is sufficient or sample size is small.

4 | DISCUSSION

The Nanopore MinION has recently been of interest to researchers due to its user-friendly operation, its ability to sequence long-read

lengths, rapid real-time sequencing capability and portability. A primary concern is its reported sequencing accuracy, which could subsequently impact taxa determination from short reads. In this study we demonstrated that Nanopore sequencing is a suitable, cost-effective, and rapid alternative for short-read DNA metabarcoding, with the high relative read abundance results aligning closely to those of Illumina.

4.1 | Comparison of Illumina and Nanopore

There is a notable difference in the clustering capabilities between the two sequencing technologies, which was expected given the higher error rate of Nanopore (at the time of this study). Illumina had greater clustering capabilities with on average >4× reads clustered per ASV compared to Nanopore OTUs (post-WoRMS filtering). However, the congruency between the results indicates that the error rate and clustering does not greatly impact species determination when using conservative filtering methods for Nanopore data. The basecalling technology for Nanopore is continuously improving (e.g., Zeng et al., 2020), and thus as the error rate improves, increasing the read accuracy, so too will the clustering capabilities. However, despite the error rate, research has shown that partitioning long reads to shorter lengths, such as the read length used in this study, improves clustering capabilities when dealing with errors (Calus et al., 2018). Furthermore, filtering for taxonomic identity and coverage helped to mitigate the sequencing error present in the Nanopore data to ensure correct taxonomic assignment that was comparable to Illumina data. It must be noted that other Nanopore pipelines have been constructed for bulk DNA metabarcoding (e.g., Baloglu et al., 2021; Maestri et al., 2019), but reference-free identification and distinguishing between closely related species was

TABLE 3 Cost calculation (USD) comparison of sequencing for DNA metabarcoding using Nanopore or Illumina technology.

Company	Item	Price (USD)	Library of 96 (USD)	Quantity
Nanopore	Flow Cell	900	900	When stored correctly can be reused
	Ligation Sequencing Kit	599	100	6 libraries
	Flow Cell Wash Kit	99	17	6 washes
	Native Barcoding Expansion*	1200	100	12 libraries
	NEBNext Ultra II End Repair/dA-Tailing Module**	795	795	96 rxns
	NEB Blunt/TA Ligase Master Mix**	410	158	250 rxns
	NEBNext® Quick Ligation Module**	328	17	20 libraries
	Total	4331	2087	
Illumina	MiSeq (standard) v3 2×300†	3252	–	
	MiSeq (standard) v2 2×250†	2603	–	
	MiSeq (nano) v2 2×250†	901	–	

Note: Items were priced in March 2022. Not included in the pricing are costs that would be fixed regardless of the sequencing method chosen, for example, DNA extraction kits, PCR reagents and DNA concentration quantification. Items marked with *, **, † and ‡ had their prices sourced from reputable and publicly available listings, Oxford Nanopore Technologies (<https://www.nanoporetech.com>), New England BioLabs (<https://www.neb.com>), Cornell University (<https://www.biotech.cornell.edu/brc/genomics/services/price-list#miseq>) and Illumina (<https://www.illumina.com>), respectively. Please note that the total cost for an Illumina MiSeq run is including the Nextera index kit (24 indexes, 96 samples; \$270)‡.

essential to this dietary DNA metabarcoding study. Thus, for comparison purposes, a pipeline needed to be developed that followed a similar framework to that implemented for the Illumina data.

For both COI and 18S the majority of high relative read abundance Illumina BLAST assignments were also assigned to Nanopore with similar abundance. A noticeable difference though was for *Thysanoessa* (krill; 18S), which was filtered out when using the >10 reads per OTU threshold for the Nanopore data. The assignments unique to either Nanopore or Illumina generally had low relative read abundance. It is probable that PCR stochasticity accounts for some of the unique assignments. It is also possible that sequencing accuracy or differences in bioinformatic sequence filtering played a role and may account for the somewhat larger number of unique Nanopore assignments. However, in many cases the unique Nanopore assignments were identified in more than one individual. Furthermore, if Nanopore sequencing inaccuracy was prevalent it would be expected that “random” assignments would occur in a more inconsistent manner across groups and not within a single sample for a specific group (e.g., diversity seen in dinoflagellates for COI and polychaetes for 18S). It would also not be expected that both of the COI primers used would detect the unique assignments if there was sequencing error.

The number of unique 18S assignments for Illumina was ~2× lower than for Nanopore, and although precautions were taken to ensure final pooling was equal in molarity for the Illumina run, it was clear that this was not reflected in the raw read ratios (Leray and Lobo had >1.5× raw reads than Zhan). Having a lower read depth would have affected the detection of those 18S sequences with lower abundance (i.e., low DNA template), making it more difficult to directly compare with the Nanopore results. The assignments in common for both sequencing devices, for COI and 18S, tended to have high relative read abundance percentages. This likely reflects items that would have passed through the gut more recently. The congruency between sequencing devices highlights the accuracy of the assignments and increases the confidence in the Nanopore results, despite the ~2% lower percentage identity than the Illumina results and the poorer clustering results (OTUs:reads vs ASVs:reads). Nanopore's sequencing depth capabilities also promotes its use for dietary analyses, as the high coverage increases the probability of diversity discovery/detection, although new taxa detection has been shown to plateau after a certain depth (Alberdi et al., 2018).

Undertaking a short-read DNA metabarcoding study (plate of 96 samples) proved to be more cost-effective on the Nanopore MinION (~\$2100; all prices stated in USD) compared to doing a standard Illumina MiSeq sequencing run (starting at ~\$2600) (Table 3). However, the upfront costs of reagents for Nanopore need to be taken into consideration if the laboratory does not frequently run DNA metabarcoding experiments. A way to “extend” Nanopore reagents, as done in this study, is to halve reaction volumes (e.g., NEBNext Ultra II End Repair/da-Tailing Module and NEB Blunt/TA Ligase Master Mix) until the final library is pooled—reducing the price by \$477. Besides the cost-saving and high coverage, a mixture of long and

short amplicons can be sequenced in a single run. A current disadvantage is that barcodes are only available for 96 samples, compared to Illumina, which offers 384 samples to be barcoded. This can be mitigated by adding 5′ primer tags (≤10 nucleotides long; Binladen et al., 2007; Carøe & Bohmann, 2020; Coissac, 2012), but multiple primers would need to be purchased. The greatest advantages for the use of Nanopore is the size, portability and the rapid return of sequencing results in real-time or, alternatively, as soon as the run has been completed (~24–48 h). It must be noted that an Illumina MiSeq nano run may be a suitable and potentially a more feasible option for DNA metabarcoding studies, but this is dependent on amplicon size, and the desired quality (i.e., Illumina's v3 chemistry being better than v2 chemistry) and read depth (v2 nano: ~1.6–2 million reads; v2 standard: 24–30 million reads; v3 standard: 44–50 million reads) (Illumina, 2017). An advantage of the Illumina runs is that the sequencing is done by trained individuals, where a minimum number of reads is often guaranteed, promoting the safe handling of precious samples.

4.2 | Leray versus Lobo primer pairs

The decision to trial both Leray and Lobo primers was that Leray has previously proved to be more difficult to amplify (van der Reis et al., 2018) and agarose gel results had indicated amplification was better when the Lobo primers were used. In addition, the Lobo primers are specifically designed for marine organisms (Lobo et al., 2013), and the reverse primer has been incorporated in COI studies when wanting to amplify particularly challenging marine organisms (Ip et al., 2019). Furthermore, the replacement of the inosine base (Table 1) with other degenerate primers makes it about four times cheaper to purchase (pricing from Integrated DNA Technologies, March 2022). The results in this study indicated little difference between the results of the two COI primer sets (Figure 3) and utilizing the Lobo reverse primer is an effective and cheaper alternative to the Leray reverse primer.

4.3 | General diet

The ocean's mesopelagic zone (200–1000m) is host to an abundance of mesopelagic fishes and, specifically, lanternfish (family Myctophidae) species are estimated to be the highest in abundance (biomass and quantity) and diversity, and thus play a crucial ecological role in oceanic food webs (Gjøesaeter & Kawaguchi, 1980; Olivar et al., 2019). Their daily diel vertical migration through the water column creates an important carbon transport link between the more productive surface layers at night (feeding opportunities) and the mesopelagic depths during the day (predator refuge), which also ensures a well-connected food web (Bernal et al., 2015; Hudson et al., 2014; Mehner & Kasprzak, 2011; Raes et al., 2022; Saunders et al., 2019). Studies have identified that these fish mainly consume zooplankton (e.g., Bernal et al., 2020; Clarke, 1980; Clarke

et al., 2020; Cohen & Beckley, 2021; McClain-Counts et al., 2017). They are preyed on by many marine predators (e.g., squid—Merten et al., 2017; predatory fishes—Connell et al., 2010; dolphins—Giménez et al., 2018; and seabirds—Cherel et al., 2002; Orben et al., 2018; Watanuki & Thiebot, 2018).

While this preliminary study focused on a single species of *Hygophum* using specimens captured in the south-east Indian Ocean, one of its overarching goals was to ensure that the methodology created a foundation for the study to be expanded to other genera/species. Two main aspects were tested, namely, to ensure dietary work on fish of early life stages would be successful if the entire gut was used (i.e., no separation of the content from the lining), and to confirm the suitability of the primers that were selected.

The results indicated that the removal of the gut content from the lining, and subsequently the extraction, only slightly improved the detection of organisms from the gut content. This supports the use of entire intestinal tracts for determining the diet from transformation and/or larval lanternfish (i.e., smaller fish than the ca. 2 cm samples used in this study), as the host DNA does not preclude dietary items from being detected. Despite the growing literature to support the selected primers, a recently published DNA metabarcoding diet study investigating lanternfish (4.5–17.6 cm) from the southern Kerguelen Axis region reported host DNA usually being assigned to 99% of reads for COI (Leray primers) and that it was poorly suited for diet analysis in lanternfish (Clarke et al., 2020). It was suggested 18S rRNA V9 (~100 bp) had better dietary recovery due to its short size, thus its ability to recover degraded dietary DNA that is likely less than the 313 bp COI fragment targeted (Clarke et al., 2020). However, preliminary *in silico* primer pair tests against the SILVA database suggested taxonomic coverage would be less for 18S rRNA V9 (e.g., >2× Copepoda entries for V4 than V9), so therefore, we decided to use the Zhan primers (425 bp) targeting V4 to broaden taxonomic coverage and/or potentially obtain better the resolution (O'Rorke et al., 2022). In this study, ~10% of total COI and ~5% 18S reads were assigned to dietary items for Nanopore (notably a greater percentage of reads were assigned to dietary items, but were filtered out due to not meeting the read number OTU clustering threshold) and, ~30% and ~20% for COI and 18S for Illumina, respectively. This indicated that host DNA did not prevent dietary detection and that these amplicon sizes are suitable for use in future work without the use of methods to block host DNA amplification (O'Rorke, Lavery, Chow, et al., 2012; O'Rorke, Lavery, & Jeffs, 2012).

This preliminary work also provided an indication of *Hygophum* spp. general diet. Calanoids were identified to be the main dietary item, which is in line with microscopic analysis done on a *Hygophum* species (Kinzer et al., 1993) in the Arabian Sea, as well as other *Hygophum* species collected in the western Mediterranean Sea (Bernal et al., 2015). Other minor dietary groupings previously identified, and in this study, included Euphausiacea (krill), Decapoda (e.g., *Lucifer*), Amphipoda and Ostracoda (Bernal et al., 2015). The majority of the genera identified in this study have also been

visually identified in the diets of three other myctophid species (*Ceratoscopelus warmingii*, *Diaphus effulgens* and *Symbolophorus evermanni*) collected across the southern tropical gyre in the Indian Ocean, indicating an overlap in prey items, and possible competition among myctophid species and size classes (Bernal et al., 2020). Soft zooplankton prey, such as tunicates, particularly Appendicularia (*Oikopleura* spp.) appear as relevant prey items from our study, but previously only sporadic or null occurrences in the microscope visual analyses of gut contents have been reported (Bernal et al., 2015; Kinzer et al., 1993; Figure 2). Due to the conservative nature of filtering to ensure all host DNA was removed and no cross-contamination occurred during DNA extraction (other mesopelagic species gut contents were extracted and considered for dietary work), all fish assignments identified were removed. However, fish have been detected as a minor dietary component of *Hygophum* in previous studies (using microscope visual gut content identification methods; Bernal et al., 2015; Kinzer et al., 1993).

Unlike microscopic analysis, DNA metabarcoding provides a wealth of information without the need for taxonomic expertise and can identify species present when partially/fully digested (e.g., gelatinous items and content in the hindgut). It also has the ability to identify the residue from organisms that have passed through the gut content previously by amplifying any tissue/DNA that may remain, increasing the timeline in which organisms can be detected. The use of 18S and COI was complementary, providing a more informative investigation into the diet.

5 | CONCLUSION

This study, to our best knowledge, is the first to compare short-read DNA metabarcoding dietary results between Nanopore MinION and Illumina MiSeq. The Nanopore BLAST assignment results were comparable to that of Illumina MiSeq. Using COI and 18S rRNA V4 gene regions proved to be beneficial in widening the taxonomic breadth detected, revealing a greater dietary composition. Savings can be made on COI universal primers by pairing the Lobo reverse primer with the Leray forward, with no decrease in taxa discovery. Nanopore is typically more cost-effective than Illumina MiSeq, with substantial cost reduction possible by halving specific volumes in the protocol. The assessment of the diet of juvenile *Hygophum* sp. specimens relying on extraction from the gut (inclusive of lining and content) was successful. Our results highlight that Nanopore, despite its lower reported sequencing accuracy, is suitable for short-read DNA metabarcoding, which can also aid in more rapid availability of molecular results.

AUTHOR CONTRIBUTIONS

ALVDR, AGJ, LEB, and MPO contributed to the conception of the study and collection of specimens. MPO and LEB carried out the taxonomic identification of the specimens. ALVDR carried out laboratory work, performed data analyses, prepared figures and tables, and wrote the manuscript. AGJ provided significant guidance

regarding the development of the project and advised on the analyses. LEB and AGJ oversaw the project's logistics and administration. All authors contributed to the development of the manuscript and approved the final manuscript.

ACKNOWLEDGMENTS

Research vessel time for this investigation was funded by Australian Marine National Facility (IN2019_V03). We acknowledge the assistance of the captain and crew of *RV Investigator* in enabling this research and especially thank the CSIRO electronics and technical team for their help with the deployment of the EZ multinet. Plankton was collected under Murdoch University ethics permit number R2885/16 and collections at stations from 29 to 23°S were undertaken under Australian Government permit AU-COM2019-446, Australian Fisheries Management Authority scientific permit 1004152, and those in the Abrolhos Marine Park under permit PA2018-00065-1. The views expressed in this publication do not necessarily represent the views of the Director of National Parks or the Australian Government.

Daniel Cohen and Danielle Hodgkinson are thanked for their enthusiastic onboard assistance. MPO acknowledges the support of her institution through the "Severo Ochoa Centre of Excellence" accreditation (CEX2019-000928-S). We are especially grateful to Dr John Paxton for his advice and for providing us with access to his unpublished report on myctophids in Australian waters.

We thank Dr Annabel Whibley for her time, Nanopore advice, and expertise. We also wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI; www.nesi.org.nz) high performance computing facilities and consulting support services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's (MBIE) Research Infrastructure program.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All the data used in this study is publicly available. The raw sequence data is available on NCBI's Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/bioproject/898436>) with associated metadata. The raw data provided for Illumina is fastq. The raw data provided for Nanopore is fast5 (a single file in archived format; i.e., pre-guppy) and fastq (post-guppy). All R code and associated data is available on Figshare as an R project (<https://doi.org/10.17608/k6.auckl.and.21482826>).

ORCID

Aimee L. van der Reis  <https://orcid.org/0000-0003-3054-284X>

Lynnath E. Beckley  <https://orcid.org/0000-0001-6872-7896>

M. Pilar Olivar  <https://orcid.org/0000-0002-8887-9181>

Andrew G. Jeffs  <https://orcid.org/0000-0002-8504-1949>

REFERENCES

- Alberdi, A., Aizpuru, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, 9(1), 134–147. <https://doi.org/10.1111/2041-210X.12849>
- Arai, M. N., Welch, D. W., Dunsmuir, A. L., Jacobs, M. C., & Ladouceur, A. R. (2003). Digestion of pelagic Ctenophora and cnidaria by fish. *Canadian Journal of Fisheries and Aquatic Sciences*, 60(7), 825–829. <https://doi.org/10.1139/f03-071>
- Baloglu, B., Chen, Z., Elbrecht, V., Braukmann, T., MacDonald, S., & Steinke, D. (2021). A workflow for accurate metabarcoding using Nanopore MinION sequencing. *Methods in Ecology and Evolution*, 12(5), 794–804. <https://doi.org/10.1111/2041-210X.13561>
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. *Nucleic Acids Research*, 41(D1), D36–D42. <https://doi.org/10.1093/nar/gks1195>
- Bernal, A., Olivar, M. P., Maynou, F., & Fernández de Puelles, M. L. (2015). Diet and feeding strategies of mesopelagic fishes in the western Mediterranean. *Progress in Oceanography*, 135, 1–17. <https://doi.org/10.1016/j.pocean.2015.03.005>
- Bernal, A., Toresen, R., & Riera, R. (2020). Mesopelagic fish composition and diets of three myctophid species with potential incidence of microplastics, across the southern tropical gyre. *Deep Sea Research Part II: Topical Studies in Oceanography*, 179, 104706. <https://doi.org/10.1016/j.dsr2.2019.104706>
- Berry, O., Bulman, C., Bunce, M., Coghlan, M., Murray, D. C., & Ward, R. D. (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>
- Binladen, J., Gilbert, M. T. P., Bollback, J. P., Panitz, F., Bendixen, C., Nielsen, R., & Willerslev, E. (2007). The use of coded PCR primers enables high-throughput sequencing of multiple homolog amplification products by 454 parallel sequencing. *PLoS One*, 2(2), e197. <https://doi.org/10.1371/journal.pone.0000197>
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., & Chase, J. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37, 852–857. <https://doi.org/10.1038/s41587-019-0209-9>
- Bucklin, A., Peijnenburg, K. T. C. A., Kosobokova, K. N., O'Brien, T. D., Blanco-Bercial, L., Cornils, A., Falkenhaus, T., Hopcroft, R. R., Hosa, A., Laakmann, S., Li, C., Martell, L., Questel, J. M., Wall-Palmer, D., Wang, M., Wiebe, P. H., & Weydmann-Zwolicka, A. (2021). Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology*, 168(6), 78. <https://doi.org/10.1007/s00227-021-03887-y>
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, 11, 2639–2643. <https://doi.org/10.1038/ismej.2017.119>
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7), 581–583. <https://doi.org/10.1038/nmeth.3869>
- Calus, S. T., Ijaz, U. Z., & Pinto, A. J. (2018). NanoAmpli-seq: A workflow for amplicon sequencing for mixed microbial communities on the Nanopore sequencing platform. *GigaScience*, 7(12), giy140. <https://doi.org/10.1093/gigascience/giy140>
- Carøe, C., & Bohmann, K. (2020). Tagsteady: A metabarcoding library preparation protocol to avoid false assignment of sequences to samples. *Molecular Ecology Resources*, 20(6), 1620–1631. <https://doi.org/10.1111/1755-0998.13227>

- Carroll, E. L., Gallego, R., Sewell, M. A., Zeldis, J., Ranjard, L., Ross, H. A., Tooman, L. K., O'Rourke, R., Newcomb, R. D., & Constantine, R. (2019). Multi-locus DNA metabarcoding of zooplankton communities and scat reveal trophic interactions of a generalist predator. *Scientific Reports*, 9(1), 281. <https://doi.org/10.1038/s41598-018-36478-x>
- Case-Green, S. C., & Southern, E. M. (1994). Studies on the base pairing properties of deoxyinosine by solid phase hybridisation to oligonucleotides. *Nucleic Acids Research*, 22(2), 131–136. <https://doi.org/10.1093/nar/22.2.131>
- Chang, J. J. M., Ip, Y. C. A., Bauman, A. G., & Huang, D. (2020). MinION-in-ARMS: Nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from autonomous reef monitoring structures. *Frontiers in Marine Science*, 7(448), 1–12. <https://doi.org/10.3389/fmars.2020.00448>
- Chang, J. J. M., Ip, Y. C. A., Ng, C. S. L., & Huang, D. (2020). Takeaways from Mobile DNA barcoding with BentoLab and MinION. *Genes*, 11(10), 1121. <https://doi.org/10.3390/genes11101121>
- Cherel, Y., Pütz, K., & Hobson, K. A. (2002). Summer diet of king penguins (*Aptenodytes patagonicus*) at the Falkland Islands, southern Atlantic Ocean. *Polar Biology*, 25(12), 898–906. <https://doi.org/10.1007/s00300-002-0419-2>
- Clarke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution*, 7(3), 873–883. <https://doi.org/10.1002/ece3.2667>
- Clarke, L. J., Trebilco, R., Walters, A., Polanowski, A. M., & Deagle, B. E. (2020). DNA-based diet analysis of mesopelagic fish from the southern Kerguelen Axis. *Deep Sea Research Part II: Topical Studies in Oceanography*, 174, 104494. <https://doi.org/10.1016/j.dsr2.2018.09.001>
- Clarke, T. A. (1980). Diets of fourteen species of vertically migrating mesopelagic fishes in Hawaiian waters. *Fishery Bulletin*, 78, 619–640.
- Cohen, D. L., & Beckley, L. E. (2021). Diet and prey selectivity of the mesopelagic lanternfish *Myctophum asperum* from the Perth canyon, Western Australia. *Ichthyological Research*, 68(2), 294–302. <https://doi.org/10.1007/s10228-020-00782-2>
- Coissac, E. (2012). OligoTag: A program for designing sets of tags for next-generation sequencing of multiplexed samples. In F. Pompanon & A. Bonin (Eds.), *Data production and analysis in population genomics: Methods and protocols* (pp. 13–31). Humana Press. https://doi.org/10.1007/978-1-61779-870-2_2
- Compson, Z. G., Monk, W. A., Hayden, B., Bush, A., O'Malley, Z., Hajibabaei, M., Porter, T. M., Wright, M. T. G., Baker, C. J. O., Al Manir, M. S., Curry, R. A., & Baird, D. J. (2019). Network-based biomonitoring: Exploring freshwater food webs with stable isotope analysis and DNA metabarcoding. *Frontiers in Ecology and Evolution*, 7(395), 1–16. <https://doi.org/10.3389/fevo.2019.00395>
- Connell, A. M., Dunn, M. R., & Forman, J. (2010). Diet and dietary variation of New Zealand hoki *Macruronus novaezelandiae*. *New Zealand Journal of Marine and Freshwater Research*, 44(4), 289–308. <https://doi.org/10.1080/00288330.2010.515232>
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, 14(2), 306–323. <https://doi.org/10.1111/1755-0998.12188>
- Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., Creer, S., Bista, I., Lodge, D. M., Vere, N., Pfrender, M. E., & Bernatchez, L. (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>
- 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(5), 294–299. <http://europemc.org/abstract/MED/7881515>
- Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources*, 13(5), 851–861. <https://doi.org/10.1111/1755-0998.12138>
- Giménez, J., Marçalo, A., García-Polo, M., García-Barón, I., Castillo, J. J., Fernández-Maldonado, C., Saavedra, C., Santos, M. B., & de Stephanis, R. (2018). Feeding ecology of Mediterranean common dolphins: The importance of mesopelagic fish in the diet of an endangered subpopulation. *Marine Mammal Science*, 34(1), 136–154. <https://doi.org/10.1111/mms.12442>
- Gjøsaeter, J., & Kawaguchi, K. (1980). A review of the world resources of mesopelagic fish. *FAO Fisheries Technical Papers (FAO)*, 13, 1–151.
- Horton, T., Kroh, A., Ah Yong, S., Bailly, N., Bieler, R., Boyko, C. B., Brandão, S. N., Gofas, S., Hooper, J. N. A., Hernandez, F., Mees, J., Molodtsova, T. N., Paulay, G., Bouirig, K., Decock, W., Dekeyzer, S., Vandepitte, L., Vanhoorne, B., Adlard, R., ... Zullini, A. (2022). *World Register of Marine Species (WoRMS)*. <https://www.marinespecies.org>
- Hudson, J. M., Steinberg, D. K., Sutton, T. T., Graves, J. E., & Latour, R. J. (2014). Myctophid feeding ecology and carbon transport along the northern mid-Atlantic ridge. *Deep Sea Research Part I: Oceanographic Research Papers*, 93, 104–116. <https://doi.org/10.1016/j.dsr.2014.07.002>
- Hulley, P. A., & Paxton, J. R. (2016a). Myctophidae. In A. N. D. A. K. Carpenter (Ed.), *Bony fishes, part 1 (Elopiformes-Scorpaeniformes), the living marine resources of the eastern Central Atlantic* (Vol. 3, pp. 1860–1928). FAO.
- Hulley, P. A., & Paxton, J. R. (2016b). Neoscolopelidae. In A. N. D. A. K. Carpenter (Ed.), *Bony fishes, part 1 (Elopiformes-Scorpaeniformes), the living marine resources of the eastern Central Atlantic* (Vol. 1, pp. 1855–1859). FAO.
- Illumina. (2013). *16S metagenomic sequencing library preparation: Preparing 16S ribosomal RNA gene amplicons for the Illumina MiSeq system* (Report no. 15044223 Rev. B). https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf
- Illumina. (2017). *MiSeq™ System. 1–5, Article M-GL-00006 v1.0*. <https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/miseq-system-data-sheet-m-gl-00006/miseq-data-sheet-m-gl-00006.pdf>
- Ip, Y. C. A., Tay, Y. C., Gan, S. X., Ang, H. P., Tun, K., Chou, L. M., Huang, D., & Meier, R. (2019). From marine park to future genomic observatory? Enhancing marine biodiversity assessments using a bio-code approach. *Biodiversity Data Journal*, 7, e46833. <https://doi.org/10.3389/BDJ.7.e46833>
- Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: Delivery of Nanopore sequencing to the genomics community. *Genome Biology*, 17(1), 239. <https://doi.org/10.1186/s13059-016-1103-0>
- Kartzinel, T. R., Chen, P. A., Coverdale, T. C., Erickson, D. L., Kress, W. J., Kuzmina, M. L., Rubenstein, D. I., Wang, W., & Pringle, R. M. (2015). DNA metabarcoding illuminates dietary niche partitioning by African large herbivores. *Proceedings of the National Academy of Sciences*, 112(26), 8019–8024. <https://doi.org/10.1073/pnas.1503283112>
- Kinzer, J., Böttger-Schnack, R., & Schulz, K. (1993). Aspects of horizontal distribution and diet of myctophid fish in the Arabian Sea with reference to the deep water oxygen deficiency. *Deep Sea Research Part II: Topical Studies in Oceanography*, 40(3), 783–800. [https://doi.org/10.1016/0967-0645\(93\)90058-U](https://doi.org/10.1016/0967-0645(93)90058-U)
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., Boehm, J. T., & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing

- coral reef fish gut contents. *Frontiers in Zoology*, 10, 34. <https://doi.org/10.1186/1742-9994-10-34>
- Lobo, J., Costa, P. M., Teixeira, M. A. L., Ferreira, M. S. G., Costa, M. H., & Costa, F. O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecology*, 13(1), 34. <https://doi.org/10.1186/1472-6785-13-34>
- Machida, R. J., Leray, M., Ho, S.-L., & Knowlton, N. (2017). Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data*, 4, 170027. <https://doi.org/10.1038/sdata.2017.27>
- Maestri, S., Cosentino, E., Paterno, M., Freitag, H., Garces, J. M., Marcolungo, L., Alfano, M., Njunjić, I., Schilthuizen, M., Slik, F., Menegon, M., Rossato, M., & Delledonne, M. (2019). A rapid and accurate MinION-based workflow for tracking species biodiversity in the field. *Genes*, 10(6), 468.
- Martin, F. H., Castro, M. M., Aboul-ela, F., & Tinoco, I., Jr. (1985). Base pairing involving deoxyinosine: Implications for probe design. *Nucleic Acids Research*, 13(24), 8927–8938. <https://doi.org/10.1093/nar/13.24.8927>
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*, 17(1), 3. <https://doi.org/10.14806/fej.17.1.200>
- McClain-Counts, J. P., Demopoulos, A. W. J., & Ross, S. W. (2017). Trophic structure of mesopelagic fishes in the Gulf of Mexico revealed by gut content and stable isotope analyses. *Marine Ecology*, 38(4), e12449. <https://doi.org/10.1111/maec.12449>
- Mehner, T., & Kasprzak, P. (2011). Partial diel vertical migrations in pelagic fish. *Journal of Animal Ecology*, 80(4), 761–770. <https://doi.org/10.1111/j.1365-2656.2011.01823.x>
- Merten, V., Christiansen, B., Javidpour, J., Piatkowski, U., Puebla, O., Gasca, R., & Hoving, H.-J. T. (2017). Diet and stable isotope analyses reveal the feeding ecology of the orangeback squid *Sthenoteuthis pteropus* (Steenstrup 1855) (Mollusca, Ommastrephidae) in the eastern tropical Atlantic. *PLoS One*, 12(12), e0189691. <https://doi.org/10.1371/journal.pone.0189691>
- Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., & Schäffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics*, 24(16), 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322>
- Oehm, J., Thalinger, B., Eisenkölbl, S., & Traugott, M. (2017). Diet analysis in piscivorous birds: What can the addition of molecular tools offer? *Ecology and Evolution*, 7(6), 1984–1995. <https://doi.org/10.1002/ece3.2790>
- Olivar, M. P., Bode, A., López-Pérez, C., Hulley, P. A., & Hernández-León, S. (2019). Trophic position of lanternfishes (Pisces: Myctophidae) of the tropical and equatorial Atlantic estimated using stable isotopes. *ICES Journal of Marine Science*, 76(3), 649–661. <https://doi.org/10.1093/icesjms/fsx243>
- Orben, R. A., Kokubun, N., Fleishman, A. B., Will, A. P., Yamamoto, T., Shaffer, S. A., Paredes, R., Takahashi, A., & Kitaysky, A. S. (2018). Persistent annual migration patterns of a specialist seabird. *Marine Ecology Progress Series*, 593, 231–245. <https://doi.org/10.3354/meps12459>
- O'Rourke, R., Lavery, S., Chow, S., Takeyama, H., Tsai, P., Beckley, L. E., Thompson, P. A., Waite, A. M., & Jeffs, A. G. (2012). Determining the diet of larvae of western rock lobster (*Panulirus cygnus*) using high-throughput DNA sequencing techniques. *PLoS One*, 7(8), e42757. <https://doi.org/10.1371/journal.pone.0042757>
- O'Rourke, R., Lavery, S., & Jeffs, A. (2012). PCR enrichment techniques to identify the diet of predators. *Molecular Ecology Resources*, 12(1), 5–17. <https://doi.org/10.1111/j.1755-0998.2011.03091.x>
- O'Rourke, R., van der Reis, A., von Ammon, U., Beckley, L. E., Pochon, X., Zaiko, A., & Jeffs, A. (2022). eDNA metabarcoding shows latitudinal eukaryote micro- and mesoplankton diversity stabilizes across oligotrophic region of a >3000 km longitudinal transect in the Indian Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 205, 105178. <https://doi.org/10.1016/j.dsr2.2022.105178>
- Paxton, J. R., Lavenberg, R. J., & Sommer, C. (1995). Myctophidae, linternillars. In F. Fischer, W. Krupp, C. Schneider, K. E. Sommer, V. Carpenter, & W. Niem (Eds.), *Guía FAO para la identificación de especies para los fines de la pesca en el Pacífico centrooriental* (pp. 1315–1321). FAO.
- Paxton, J. R., & Williams, A. (2019). *Key to the genera of Australian lanternfishes, family Myctophidae*. Australian Museum.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Pompanon, F., Deagle, B. E., Symondson, W. O. C., 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>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), D590–D596. <https://doi.org/10.1093/nar/gks1219>
- R Core Team. (2021). *R: A language and environment for statistical computing (version 4.1.0; R studio version 1.4.1106)*. R Foundation for Statistical Computing.
- Raes, E. J., Hörstmann, C., Landry, M. R., Beckley, L. E., Marin, M., Thompson, P., Antoine, D., Focardi, A., O'Brien, J., Ostrowski, M., & Waite, A. M. (2022). Dynamic change in an ocean desert: Microbial diversity and trophic transfer along the 110 °E meridional in the Indian Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 201, 105097. <https://doi.org/10.1016/j.dsr2.2022.105097>
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. <https://doi.org/10.7717/peerj.2584>
- Runtuwene, L. R., Tuda, J. S. B., Mongan, A. E., & Suzuki, Y. (2019). On-site MinION sequencing. In Y. Suzuki (Ed.), *Single molecule and single cell sequencing* (pp. 143–150). Springer Singapore. https://doi.org/10.1007/978-981-13-6037-4_10
- Ruppert, K. M., Kline, R. J., & Rahman, M. S. (2019). Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecology and Conservation*, 17, e00547. <https://doi.org/10.1016/j.gecco.2019.e00547>
- Santos, A., van Aarle, R., Barrientos, L., & Martínez-Urtaza, J. (2020). Computational methods for 16S metabarcoding studies using Nanopore sequencing data. *Computational and Structural Biotechnology Journal*, 18, 296–305. <https://doi.org/10.1016/j.csbj.2020.01.005>
- Saunders, R. A., Hill, S. L., Tarling, G. A., & Murphy, E. J. (2019). Myctophid fish (family Myctophidae) are central consumers in the food web of the Scotia Sea (Southern Ocean). *Frontiers in Marine Science*, 6(530), 1–22. <https://doi.org/10.3389/fmars.2019.00530>
- Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One*, 11(10), e0163962. <https://doi.org/10.1371/journal.pone.0163962>
- van der Loos, L. M., & Nijland, R. (2021). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. *Molecular Ecology*, 30(13), 3270–3288. <https://doi.org/10.1111/mec.15592>
- van der Reis, A. L., Jeffs, A. G., & Lavery, S. D. (2020). From feeding habits to food webs: Exploring the diet of an opportunistic benthic generalist. *Marine Ecology Progress Series*, 655, 107–121. <https://doi.org/10.3354/meps13511>

- van der Reis, A. L., Laroche, O., Jeffs, A. G., & Lavery, S. D. (2018). Preliminary analysis of New Zealand scampi (*Metanephrops chalcongeri*) diet using metabarcoding. *PeerJ*, 6, e5641. <https://doi.org/10.7717/peerj.5641>
- Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied Environmental Microbiology*, 73(16), 5261–5267. <https://doi.org/10.1128/aem.00062-07>
- Watanuki, Y., & Thiebot, J.-B. (2018). Factors affecting the importance of myctophids in the diet of the world's seabirds. *Marine Biology*, 165(4), 79–14. <https://doi.org/10.1007/s00227-018-3334-y>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Zamora-Terol, S., Novotny, A., & Winder, M. (2020). Reconstructing marine plankton food web interactions using DNA metabarcoding. *Molecular Ecology*, 29(17), 3380–3395. <https://doi.org/10.1111/mec.15555>
- Zeng, J., Cai, H., Peng, H., Wang, H., Zhang, Y., & Akutsu, T. (2020). Causalcall: Nanopore basecalling using a temporal convolutional network. *Frontiers in Genetics*, 10(1332), 1–11. <https://doi.org/10.3389/fgene.2019.01332>
- Zhan, A., Hulák, M., Sylvester, F., Huang, X., Adebayo, A. A., Abbott, C. L., Adamowicz, S. J., Heath, D. D., Cristescu, M. E., & MacIsaac, H. J. (2013). High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods in Ecology and Evolution*, 4(6), 558–565. <https://doi.org/10.1111/2041-210X.12037>

SUPPORTING INFORMATION

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

How to cite this article: van der Reis, A. L., Beckley, L. E., Olivar, M. P., & Jeffs, A. G. (2023). Nanopore short-read sequencing: A quick, cost-effective and accurate method for DNA metabarcoding. *Environmental DNA*, 5, 282–296. <https://doi.org/10.1002/edn3.374>