Title

Capturing open ocean biodiversity: comparing environmental DNA metabarcoding to the continuous plankton recorder

Running Title

eDNA for capturing open ocean biodiversity

Authors

Leonie Suter[1], Andrea Maree Polanowski[1], Laurence John Clarke[1,2], John Andrew Kitchener[1], Bruce Emerson Deagle[1,3]

Affiliations

1. Australian Antarctic Division, Department of Agriculture, Water and the Environment, 203 Channel Highway, Kingston TAS 7050, Australia
2. Institute of Marine and Antarctic Studies, University of Tasmania, 20 Castray Esplanade, Battery Point TAS 7004, Australia

3. Commonwealth Scientific and Industrial Research Organisation, Castray Esplanade, Battery Point TAS 7004, Australia

Corresponding Author

Leonie Suter, Email: leonie.suter@awe.gov.au, Tel: +61 3 6232 3630

Abstract

Environmental DNA (eDNA) metabarcoding is emerging as a novel, objective tool for monitoring marine metazoan biodiversity. Zooplankton biodiversity in the vast open ocean is currently monitored through continuous plankton recorder (CPR) surveys, using ship-based bulk plankton sampling and morphological identification. We assessed whether eDNA metabarcoding (2 L filtered seawater) could capture similar Southern Ocean zooplankton biodiversity as conventional CPR bulk sampling (~1500 L filtered seawater per CPR sample). We directly compared eDNA metabarcoding with (i) conventional morphological CPR sampling and (ii) bulk DNA metabarcoding of CPR collected plankton (two transects for each comparison, 40 and 44 paired samples respectively). A metazoan-targeted cytochrome c oxidase I (COI) marker was used to characterize species-level diversity. In the 2 L seawater eDNA samples this marker amplified large amounts of non-metazoan picoplanktonic algae, but eDNA metabarcoding still detected up to 1.6 times more zooplankton species than morphologically analysed bulk CPR samples. COI metabarcoding of bulk DNA samples mostly avoided non-metazoan amplifications and recovered more zooplankton species than eDNA metabarcoding. However, eDNA metabarcoding detected roughly two thirds of metazoan species and identified similar taxa contributing to community differentiation across the subtropical front separating transects. We observed a diurnal pattern in eDNA data for copepods which perform diel vertical migrations, indicating a surprisingly short temporal eDNA signal. Compared to COI, a eukaryote-targeted 18S ribosomal RNA marker detected a higher proportion, but lower diversity, of metazoans in eDNA. With refinement and standardization of methodology, eDNA metabarcoding could become an efficient tool for monitoring open ocean biodiversity.

Keywords

Environmental DNA, biodiversity, metabarcoding, continuous plankton recorder, open ocean, biomonitoring

53

1 Introduction

Biodiversity surveying and long-term monitoring of the open ocean is crucial to capture the effect of rapidly changing climate on marine ecosystems (Constable et al., 2014; Hoegh-Guldberg & Bruno, 2010). Plankton communities are particularly useful as an indicator of change in this environment, as they reflect a diversity of biological responses to climatic variation at a range of timescales (Beaugrand, 2005; McQuatters-Gollop et al., 2019). These communities also form the base of the oceanic food web which supports global marine ecosystems and fisheries (Frederiksen, Edwards, Richardson, Halliday & Wanless, 2006). However, carrying out biological monitoring in the open ocean has always been challenging, particularly for metazoan (i.e. animal) species, as they range in size enormously and a variety of sampling techniques are required. Currently, most data on metazoans comes from either using indirect methods such as hydro-acoustics (Egerton et al., 2018; Kloser, Ryan, Sakov, Williams & Koslow, 2002), datasets from fisheries catch (Cheung, Watson & Pauly, 2013; Gordon, Beaumont, MacDiarmid, Robertson & Ahyong, 2010), using nets to sample the water column during dedicated marine science voyages (Hosie, Schultz, Kitchener, Cochran & Richards, 2000; Koubbi et al., 2010; O'Hara, Williams, Althaus, Ross & Bax, 2020), or using underway sampling methods such as the continuous plankton recorder (CPR) (Hosie, Fukuchi & Kawaguchi, 2003; Reid, Colebrook, Matthews, Aiken & Team, 2003). Recently, DNA metabarcoding has emerged as an important tool to characterize marine biodiversity and has provided new insights into life in the world's oceans (Stat et al., 2019). DNA-based approaches promise to circumvent some challenges associated with traditional survey methods particularly the time-consuming microscopic identification of plankton (Deagle, Clarke, Kitchener, Polanowski & Davidson, 2018; Leduc et al., 2019) which is currently the foundation of many marine biology investigations. The other primary advantage of a DNA-based approach is the potential for a standardised taxonomy with less reliance on direct input from taxonomic experts (Ji et al., 2013; Ruppert, Kline & Rahman, 2019). This is especially relevant as the expertise required to identify species from such a diverse array of phyla is in decline (Hutchings, 2019; Kholia & Fraser-Jenkins, 2011).

There are two broad ways that DNA metabarcoding has been implemented in biodiversity surveys of metazoans. In the first, whole organisms are collected using standard sampling methods (e.g. plankton net sampling) and DNA is extracted from a bulk sample of organisms (e.g. Hirai, Kuriyama, Ichikawa, Hidaka & Tsuda, 2015). The bulk sampling approach is generally applied to smaller planktonic organisms, but can also be applied to large species by pooling tissue samples (Ratcliffe et al., 2020) or extracting DNA from sample preservative (Hajibabaei, Spall, Shokralla & van Konynenburg, 2012). The second approach is to focus on trace amounts of environmental DNA (eDNA) generally collected by filtering small volumes of water through fine filters (< 1 µm pore size; Rees, Maddison, Middleditch, Patmore & Gough, 2014). The sources of eDNA in the ocean include: excreted faeces, mucous, and gametes; shed cells and scales; degrading tissue (e.g. moults and carcasses) and whole live organisms (Taberlet, Bonin, Zinger & Coissac, 2018). In reality these two approaches, 'bulk DNA metabarcoding' (Taberlet, Coissac, Pompanon, Brochmann & Willerslev, 2012) and 'eDNA metabarcoding', are part of a continuum of DNA sampling strategies – from coarse filters with large volumes of water to very fine filters and small water volumes, but for practical purposes we consider them separately here.

One of the main benefits of using eDNA metabarcoding rather than bulk DNA metabarcoding (henceforth called 'bulk metabarcoding') in biomonitoring applications is the relative ease of sample collection and ability to standardise sample processing. However, the eDNA metabarcoding methodology still needs to be optimised and evaluated against other approaches before it can be considered as a viable alternative (Goldberg et al., 2016; Macher et al., 2018). In marine environments, it remains unclear how effective eDNA metabarcoding is in detecting species and community compositions, to what extent this depends on sampled water volume, and for how long eDNA signals remain detectable in the open ocean (Sepulveda et al., 2019; Thomsen et al., 2012). These initial validation steps are particularly important for applications to long-term monitoring because of the effort required to collect these large datasets and the critical need to maintain a methodology that ensures comparability of data (Pawlowski et al., 2018). To date much of the eDNA metabarcoding

validation work for metazoan communities has focussed on freshwater lakes and streams (Blackman et al., 2019; Deiner, Walser, Mächler & Altermatt, 2015). In the marine environment, studies are most common in biologically rich estuaries (Brown, Chain, Zhan, MacIsaac & Cristescu, 2016) and the coastal neritic zone (< 200 m) above the drop-off of the continental shelf (Bucklin et al., 2019). Less research has been done in the largest marine biome, the open ocean zone beyond the continental shelf.

To validate open ocean eDNA metabarcoding, we compared eDNA sampling to bulk plankton sampling of the continuous plankton recorder (CPR). The CPR survey is a large, long-term biomonitoring effort that samples open ocean plankton communities (Hosie et al., 2003; Reid et al., 2003). The survey uses a device that is towed behind a ship to capture plankton from surface water at a continuous rate over large areas of the ocean and is often deployed from merchant ships. In a previous study we compared metazoan DNA metabarcoding of bulk CPR samples with conventional morphological analysis of CPR samples (Deagle et al., 2018). The DNA-based approach identified more taxa, but the CPR unit is not an ideal genetic sampler, as there is potential for cross-contamination and the protocol we used to preserve samples was quite time consuming – a critical consideration when sampling from ships of opportunity. Here we examine whether eDNA has a similar capacity to detect species and community composition as the CPR survey by comparing metabarcoding of eDNA samples to metabarcoding of CPR samples as well as to traditional morphologic analysis of CPR samples.

One of the most important decisions in the design of eDNA metabarcoding studies is the choice of metabarcoding marker, as this both determines which taxa can be detected as well as the taxonomic resolution (Bucklin, Lindeque, Rodriguez-Ezpeleta, Albaina & Lehtiniemi, 2016; Leray & Knowlton, 2016; Tang et al., 2012). Different metabarcoding markers are used in studies of the full diversity of marine eukaryotic organisms (e.g. De Vargas et al., 2015; López-Escardó et al., 2018) than in focussed surveys into the distribution of large taxa (e.g. sharks or fish; Boussarie et al., 2018; Thomsen et al., 2016). In the open ocean, bacterial communities have been extensively characterized

using 16S ribosomal RNA (rRNA) markers (e.g. Raes et al., 2018), surveys of eukaryotes have been conducted using 18S rRNA markers (De Vargas et al., 2015), and bulk metabarcoding of plankton and eDNA barcoding analysis targeting specific taxa using species-specific assays have also been done (Sigsgaard et al., 2017; Yamahara et al., 2019). Here, we chose a cytochrome c oxidase I (COI) assay, designed to resolve a large variety of metazoan taxa to species level (Leray et al., 2013). The high taxonomic resolution, combined with an extensive reference database, provides the best direct comparison of eDNA metabarcoding to conventional morphological CPR. We also characterised overall eukaryotic diversity in our eDNA samples using an 18S rRNA marker to gain a broader perspective of taxa present in the eDNA.

We collected eDNA and CPR samples in parallel along four open ocean transects between Tasmania, Australia and the subantarctic Macquarie Island, spanning the subtropical front. For eDNA sampling, two litres of surface seawater were collected from the ship's underway seawater line at roughly 5 nautical mile intervals. For the CPR sampling, plankton material continuously collected over each 5 nautical mile interval (filtering approximately 1 500 L of seawater (Hunt & Hosie, 2006)) was considered as a single sample. Each CPR sample therefore had a corresponding eDNA sample that was collected in parallel. The CPR sampling was divided into either morphological or genetic transects, as the sample storage requirements for genetic and morphological analyses are different and mutually exclusive (see methods). Therefore, two transects were analysed using conventional morphological methods, while the remaining two transects were analysed by bulk metabarcoding. eDNA samples from all four transects were characterized using high throughput amplicon sequencing of a COI and an 18S marker. The bulk metabarcoding CPR samples were characterised with the same COI marker as used for the eDNA metabarcoding. The COI marker was also used to examine two additional large-volume eDNA samples (> 2000 L of seawater) collected by continuously filtering water over the full length of two transects.

We specifically address the following questions: i) Can eDNA metabarcoding capture a similar number of metazoan species as bulk metabarcoding or morphological analysis of CPR samples within transects? ii) Can eDNA metabarcoding detect similar metazoan community composition as bulk metabarcoding or morphological analysis of CPR samples across transects? iii) How long does the eDNA of a targeted taxon remain detectable in the open ocean? iv) How does metabarcoding of large volume eDNA samples compare to small volume eDNA metabarcoding?

2 Material and Methods

2.1 CPR sampling for morphological and genetic analysis

Samples were collected on board the *RSV Aurora Australis* in the Southern Ocean south of Tasmania, Australia (Figure 1A) in March 2018 (9 – 24 March). A CPR unit (Type II, Mark V) fitted with 270 µm nylon mesh was towed 100 m behind the ship at about 10 m depth. A total of four CPR tows were conducted (Figure 1A); two conventional tows ("Morph-A" and "Morph-B") and two genetic tows ("Gen-A" and "Gen-B"). Transects Morph-A (46° 52' 30'' S to 48° 18' 43'' S) and Morph-B (47° 52' 05'' S to 46° 51' 24'' S) followed standard CPR methods (Hosie et al., 2003) and the CPR silk spools were fixed in 10% buffered formalin for morphological identification of metazoan species. For the two genetic transects, Gen-A (43° 26' 31'' S to 45° 30' 41'' S) and Gen-B (53° 05' 43'' S to 51° 34' 51'' S), a sheet of plastic was added between the mesh layers on the spool to minimize cross-contamination from eDNA potentially seeping through the spool layers (see Deagle et al., 2018) and no formalin was added, as formalin is not a suitable preservative for downstream genetic analyses. For genetic analysis the mesh was unrolled once on board, cut into 5.5 cm segments (each representing approximately 5 nautical miles of a tow and approximately 1 500 L of filtered seawater) and frozen at -80°C on board. Gen-A was sampled north of the subtropical front (Sikes et al., 2009) in relatively warm waters (17.4 – 19.1 °C), while the other three transects were sampled south of the subtropical front in much cooler waters of the Southern Ocean (Morph-A: 12.0 – 14.3 °C ; Morph-B: 10.9 -13.0 °C; Gen-B: 7.4 – 8.6 °C).

2.2 eDNA sampling

eDNA was collected in parallel to the CPR plankton samples during each of the four transects. Two litre of seawater were collected approximately every 30 minutes directly from the ship's uncontaminated seawater line (4 ± 2 m depth). Each eDNA sample provides a snapshot sample which corresponds to one of the CPR segments we collected in parallel (i.e. plankton collected on mesh by

CPR unit over 5 nautical miles from 1 500 L of seawater). The eDNA water samples were vacuum filtered through 47 mm, 0.45 µm polyethersulfone filters (Pall Life Sciences, New York, NY, USA) using a Sentino microbiology pump (Pall Life Sciences). The membrane discs were cut in half and frozen immediately at −80 °C. One half of the disk was used for genetic analysis, while the other half was kept in storage at -80 °C as a backup and for potential long-term biomonitoring studies in the future. This 'eDNA biobanking' of samples is crucial to allow for future changes in eDNA processing methods (Jarman, Berry & Bunce, 2018). During sampling, a total of eight blank samples were taken by filtering 500 mL of the freshwater used in the laboratory. The 2 L Nalgene bottles and filtration equipment were rinsed with freshwater between filtrations and sterilised by soaking in 10% bleach after every fourth sample.

2.3 Large-volume eDNA sampling

We investigated the characteristics of eDNA in larger sample volumes and coarser filters (large-volume eDNA), by continuously filtering seawater from the underway seawater line through a piece of 270 µm silk mesh over two entire transects. One sample was taken at the same time as transect Gen-B (filtering approximately 2160 L of seawater over 9 hours), overlapping with eDNA and CPR samples of that transect (51° 29' 52'' S to 53° 15' 09'' S). The other sample was taken at the end of the voyage (filtering approximately 2880 L of seawater over 12 hours), with a near-perfect overlap (42° 56' 52.6236'' S to 45° 9' 38.7036'' S) to samples of transect Gen-A, which was taken 13 days earlier. The mesh from large-volume eDNA sampling were placed in falcon tubes and frozen at -80 °C on board.

2.4 Morphological analysis

Samples from the morphological transects were processed following standard methods of the Southern Ocean CPR survey (see Hosie et al., 2003; McLeod, Hosie, Kitchener, Takahashi & Hunt, 2010). Briefly, silks were unrolled and cut into segments representing 5 nautical miles. All

zooplankton on each segment were identified and counted under a dissecting microscope. Identification was to the lowest taxonomic level possible, based on the Register of Antarctic Marine Species (de Broyer & Danis, 2014). Euphausiids and copepods were also identified to developmental stage where possible. Taxa within some broad lineages were not differentiated (e.g. hydrozoans and larval echinoderms) and damaged specimens were grouped at coarser taxonomic levels where necessary (e.g. indeterminate euphausiid).

2.5 DNA extraction

2.5.5 Bulk DNA

To extract bulk DNA from CPR segments, 50 mL of 80% ethanol was added to the storage tube containing the CPR silk (n = 40). The tubes were shaken vigorously several times and then left overnight. Using a wide bore pipette tip, 1.2 mL of settled sediment was transferred into a tube, centrifuged at 3000 rpm for 1 min and the ethanol was poured off to provide material for extraction. This sample processing method has previously been shown to be optimal to capture metazoan diversity in CPR samples (Deagle et al., 2018). DNA extraction was carried out using the QIAGEN DNeasy Blood & Tissue kit (QIAGEN) following the manufacturer's instructions (including the Proteinase K digestion step). Extracts were eluted in 150 µL buffer EB and stored at -20 °C

2.5.2 Large-volume eDNA

The large-volume eDNA samples (n = 2) followed the same protocol as the CPR segments, with the exception that three technical replicates, consisting of 1.2 mL of settled sediment, were taken from each sample.

2.5.3 eDNA

DNA was extracted from one half of each filter membrane (n = 84) using the NucleoSpin Tissue system (Macherey-Nagel, Düren, Germany) at the Australian Genome Research Facility (AGRF, Adelaide, Australia; http://www.agrf.org.au).

2.6 PCR and sequencing

2.6.1 COI

A mitochondrial cytochrome c oxidase subunit I (COI) marker (approximately 310 bp) was amplified using highly degenerate metazoan primers (forward primer: mlCOIintF, reverse primer: jgHCO2198; Leray et al., 2013). PCR amplifications were performed in two rounds, the first to amplify the target locus and add sample-specific 6 bp multiplex-identifier (MID) tags (forward and reverse primer) and Illumina sequencing primers, the second to add sequencing adapters and additional 10 bp MIDs (Clarke, Beard, Swadling & Deagle, 2017). All pre-PCR preparations were carried out in a separate clean room designed for eDNA work, which was physically separated from the main laboratory. For first-round PCR, the 10 µL reaction mixes contained 1 µL DNA (diluted to 5 ng μL^{-1}), 0.5 µM of each primer, 2 µg BSA, 5 µL AmpliTaq Gold™ 360 Master Mix (Life Technologies) and nuclease free water (QIAGEN). Thermal cycling conditions were 95°C for 10 min; followed by 35 cycles for bulk DNA and large-volume eDNA sampels and 40 cycles for eDNA samples of 95°C for 30 s, 46°C for 30 s and 72°C for 45 s; and then a final extension at 72°C for 5 min. For the second round, samples were diluted 1:10 and 1 µL of this used as template in a 10 cycle PCR. Conditions were as in the first round except for 55°C annealing temperature. PCR products were pooled, purified using Agencourt AMPure XP beads (Beckman Coulter) and sequenced in both directions using a Miseq V2 Reagent Kit (500 cycles).

All eDNA and bulk DNA samples were sequenced in a first sequencing run. As five eDNA samples had failed during this first run, they were sequenced again in a second sequencing run, together with 12 repeated eDNA samples and the two large-volume eDNA samples. Sequencing depth in the second

sequencing run was higher than in the first one (average number of sequences of repeated samples: first round sequencing: 30 995; second round sequencing: 106 337). Negative controls included 8 water blanks (500 mL of tap water filtered on board between eDNA sampling, see above), 2 extraction blanks (DNA extraction without a sample) and 3 PCR blanks (no DNA added to the PCR amplification).

2.6.2 18S

The V4 region of the 18S rRNA gene (Piredda et al., 2017) was amplified following the protocol for the BPA-Marine Microbiome Projects (Brown et al., 2018) at the Ramaciotti Centre of Genomics (UNSW, Australia). In brief, the 25 µL PCR mastermix contained 1 x KAPA HiFi Hot Start Readymix, 0.5 µM forward primer (V4_18SNext.For with indexed adapter sequences for sequencing on MiSeq platform), 0.5 µM reverse primer (V4_18SNext.Rev with indexed adapter sequences) and 1 µL of undiluted template DNA. Thermal cycling conditions were 98°C for 30 s; followed by 10 cycles of 98°C for 10 s, 44°C for 30 s and 72°C for 15 s; 20 cycles of 98°C for 10 s, 62°C for 30 s and 72°C for 15 s and then a final extension at 72°C for 7 min. PCR products were cleaned and normalized using SequalPrep Normalization plates according to manufacturer's instructions. Normalized amplicons were pooled, the pool was diluted to 4 nM and sequenced in both directions using a Miseq V2 Reagent Kit (500 cycles).

2.7 Data processing

2.7.1 COI

The 10 bp MIDs were used initially to assign sequences to sample-specific fastq files on the MiSeq. The command "fastq_mergepairs" in USEARCH v10.0.240 (Edgar, 2010) was used to merge paired reads. We identified and kept reads flanked by exact matches of the first round 6 bp MID tags and COI primer sequences using the R package "ShortRead" (Morgan et al., 2009). These processed sequences were then pooled into one fasta file and unique sequences were identified (dereplication) using the

"fastx_uniques" command in USEARCH. From the resulting dereplicated fasta file, sequencing errors and chimaerae were removed using the USEARCH command "unoise3" to create an OTU (operational taxonomic unit) table of unique sequences (no clustering of sequences was applied). To assign read numbers of individual samples to this OTU table, sample sequences were firstly given sample-specific labels using the USEARCH command "fastx_relabel" and these sequences were then pooled into one fasta file. The command "fastx_uniques_persample" was used to identify unique sequences per sample and from the resulting fasta file the sequences per sample were assigned to the OTU table using the USEARCH command "otutab". The raw sequencing data and processing pipeline described here are available on the data repository Dryad (Suter, Polanowski, Clarke, Kitchener & Deagle, 2020).

The OTU sequences were searched against the NCBI nucleotide database using the command blastn (Madden, 2013) (settings: -num_descriptions 50 -num_alignments 50 -num_threads 16 -perc_identity 80; environmental samples, metagenomes and unidentified organisms were excluded using the command -negative_seqidlist). We used MEGAN to assign taxonomy (Huson et al., 2016; LCA parameters "min score": 300; "top percent": 5; "min support": 1), and then manually curated all metazoan assignments, including double-checking metazoan sequences using the online BOLD identification engine for additional entries (http://www.barcodinglife.com/index.php/IDS_OpenIdEngine; Ratnasingham & Hebert, 2007), verifying that the identified species occurred in the sampled range and checking whether closely related species existing in the range were present or absent in the genetic reference databases. If two species could not be separated based on their COI sequence, the respective OTU sequences were called e.g. "*Pleuromamma gracilis/piseki*". If a closely related species was missing in the reference databases, the name was added in brackets, e.g. "*Euphausia similis/(spinifera)*". Additionally, taxonomic assignments were verified by aligning all OTU sequences of metazoan groups (e.g. all OTU assigned to the phylum Cnidaria) and calculating a phylogenetic tree using MEGA version 10.0.5 (Kumar, Stecher, Li, Knyaz & Tamura, 2018). If OTU sequences that were only resolved to a high taxonomic level (e.g. Hydrozoa) formed a closely related monophyletic group of at least two

sequences resembling a species, these sequences were renamed to e.g. “Hydrozoa sp. A” (see Figure S1).

Extraction blanks and PCR blanks from the first round of sequencing contained no metazoan OTUs, indicating that there was no cross-contamination of samples from laboratory work. The PCR blank from the second round of sequencing as well as all other samples from the second round of sequencing contained very low levels of cross-contamination from other samples. Consequently, read numbers < 10 per OTU per sample were set to 0 in the second round sequencing data. Water blanks taken between eDNA samples contained 1 – 4 metazoan species each, but detected species did not correspond to species detected in eDNA samples taken immediately before the water blank, and species composition of water blank samples did not correspond to eDNA samples of the same transect (see results).

OTUs that were annotated to likely contaminant species were removed from the data set. Species included *Homo sapiens* (human), *Canis lupus* (dog), *Demodex brevis* (human face mite) and *Bactrocera tryoni/aquilonis* (Queensland fruit fly).

Samples with low sequencing coverage (less than 7 500 total reads excluding contaminants) were removed from the data set. Only a dataset of matched sample pairs was retained (i.e. eDNA sample matched to a CPR sample collected at the same time). Negative controls were kept in the data set regardless of sequencing depth.

Five eDNA samples from transect Morph-B had failed in the first sequencing round and were re-sequenced at higher sequencing depth in a second round of sequencing (see above). These five samples were rarefied to the average sequencing depth of all other eDNA samples of transect Morph-B (35’850 reads) using the function “rrarefy” of the R package “vegan” (Oksanen et al., 2007) to make them comparable to the other samples. They were then included in all further analyses.

2.7.2 18S

18S sequences were processed following the Australian Marine Microbial Biodiversity Initiative protocol described in Brown et al. (2018). In brief, sequences were merged using FLASH (Magoč & Salzberg, 2011), then sequences were de-replicated and an OTU table was generated using USEARCH. The taxonomy of OTUs was identified using SILVA (Quast et al., 2012) and classified using MOTHUR's implementation of the Wang classifier (Schloss et al., 2009; Wang, Garrity, Tiedje & Cole, 2007).

2.8 Data analysis

Most data analyses were performed in R (R Core Team, 2018) and the R markdown code for all analyses as well as all processed data is available on Dryad (Suter et al., 2020). Most figures were visualized using the R package "ggplot2" (Wickham, 2016).

To get an initial overview of the genetic data, OTUs were classified as 'metazoa', 'non-metazoa' or 'not classified' and read numbers of each category were plotted per transect and sampling method. To establish whether eDNA metabarcoding could detect similar alpha and beta diversities as bulk metabarcoding, we did not rarefy sequence counts to standardise metazoan read numbers. The overall read numbers were similar between eDNA samples and bulk DNA samples, so the sequencing effort was equivalent between the sample types, even though eDNA samples had much lower metazoan read numbers (see results). Rarefying bulk DNA metazoan read numbers to a similar level as eDNA metazoan read numbers would have drastically reduced the number of species detected in bulk DNA samples, which would impede the purpose of verifying how much diversity can be detected in eDNA samples compared to traditional bulk sampling methods.

To compare alpha diversity, Hill numbers were calculated for each sample (Alberdi & Gilbert, 2019a). For q = 0, Hill numbers are equivalent to species richness, i.e. the number of species that were detected. For q = 1 (equivalent to Shannon diversity) and q = 2 (equivalent to Simpson diversity), more abundant species have an increasing weight, or species with low abundance are increasingly

neglected. When eDNA metabarcoding was compared to bulk metabarcoding, Hill numbers were calculated with q = 0, q = 1 and q = 2, whereas when eDNA metabarcoding was compared to morphological analysis of CPR samples, Hill numbers were only calculated with q = 0 (species richness), as sampling methods were too different to compare species abundance. Hill numbers and diversity profiles were calculated using the R package "hilldiv" (Alberdi & Gilbert, 2019b), and pairwise *t*-tests were calculated to test for statistical difference between groups.

To characterize species composition between transects (Gen-A vs Gen-B; Morph-A vs Morph-B) and sampling methods (eDNA vs bulk metabarcoding vs large-volume eDNA metabarcoding; eDNA metabarcoding vs morphological CPR analysis), principal coordinates analyses (PCoA) were calculated based on binary Jaccard distances (all transects) and Bray-Curtis distances (genetic transects only). PERMANOVAs were calculated to assess the separation of these groups, and a SIMPER analysis (Clarke, 1993) was performed to establish which species contribute most to the separation of the two genetic transects for eDNA metabarcoding and bulk metabarcoding, respectively. PCoAs, PERMANOVAs and SIMPER analyses were calculated using the R package "vegan" (Oksanen et al., 2007). Additionally, we determined whether species detection within metazoan phyla was shared between methods or limited to one sampling method only, separately for each transect.

To determine whether taxonomic levels or taxonomic group impacted the comparison of the sampling methods, proportion of sequence reads of metazoan phyla, calanoid copepod genera and chordate species were compared per transect and sampling method.

To investigate how long the eDNA from zooplankton lasted in the water column, we examined data from copepods known to make diel vertical migrations, i.e. they forage in surface waters at night and move to deeper layers during the day to avoid predation (Lo, Shih & Hwang, 2004). We assessed whether the abundance of species belonging to the copepod genus *Clausocalanus* corresponded to day or night sampling, both for morphological CPR samples and eDNA samples.

To establish whether in the non-metazoan reads amplified with the COI marker in eDNA samples represented the non-metazoan diversity accurately, we compared the COI reads to 18S reads at class level resolution in all eDNA samples.

To assess whether there was a bias in metazoan group detection of the COI marker, we compared metazoan COI reads to metazoan 18S reads at phylum level resolution in all eDNA samples.

To visualise taxonomic diversity in the samples, neighbour-joining trees were calculated in MEGA version 10.0.5 (Kumar et al., 2018), and trees were drawn using the online tool 'interactive tree of life (iTOL)' (Letunic & Bork, 2006), adding species presence, abundance (normalized against metazoan reads and square-root-transformed) and OTU numbers as additional datasets.

To determine whether sequencing depth affected the number of metazoan species detected in eDNA samples, rarefaction plots were calculated and drawn for original and repeated eDNA sample pairs using the R package "iNEXT" (Chao et al., 2014). Additionally, we compared proportion of metazoan reads as well as detected metazoan species between sequencing depth sample pairs.

3 Results

3.1 Data overview and taxa detected

For the COI marker, a total of 6.74 million Illumina paired-end reads were retained after filtering and quality control. These were allocated to 12 339 OTUs, of which 1 292 were assigned to metazoan taxa. While overall read numbers of eDNA, bulk DNA and large-volume eDNA samples were comparable (mean read number per transect, eDNA: 660 535; bulk DNA: 482 419; large volume eDNA: 611 408; Table S1), the proportion of metazoan reads in eDNA samples was much lower than in bulk DNA and large-volume eDNA samples (eDNA: 1.4 – 9.7 % of total reads; bulk DNA: 70.1 and 83.8 %; large-volume eDNA: 89.7 and 97.0 %; Figure 1B, Figure S2, Table S1).

Of the 1 292 metazoan COI OTUs, 1 055 could be taxonomically resolved to species level and were assigned to 198 species (115 in eDNA, 168 in bulk DNA, 87 in large-volume eDNA, Figure S3 A, Table S1, however, most species had very low abundance (Figure S3 B). The number of OTUs per species varied greatly (probably reflecting differences in intraspecific variation), ranging from 102 OTUs for the krill species *Nyctiphanes australis* to only 1 OTU for many other species (Figure S3 C). Morphological analysis of CPR samples detected 20 species (Table S1). At phylum level, eDNA metabarcoding detected 13 phyla (4 transects), bulk metabarcoding 12 phyla (2 transects), large-volume eDNA metabarcoding 7 phyla (2 transects) and morphological CPR 5 phyla (2 transects).

3.2 Alpha diversity

We compared metazoan alpha diversity detected with eDNA metabarcoding to bulk metabarcoding in two transects (Gen-A and Gen-B; Figure 1). When comparing species richness (Hill numbers calculated with q = 0), the northern transect Gen-A detected about four times more species than the southern transect Gen-B, both for eDNA metabarcoding and bulk metabarcoding. Within each transect, bulk metabarcoding detected about 1.5 times more species than eDNA metabarcoding

(Figure 2A, Table S1). With increasing q values, the difference between eDNA metabarcoding and bulk metabarcoding decreased for Gen-A, until with q = 2 the methods became comparable (Figure 2A, Table S2), suggesting that bulk metabarcoding detected more rare species, but similar numbers of abundant species as eDNA metabarcoding. In contrast, in transect Gen-B, bulk metabarcoding always detected higher Hill numbers than eDNA metabarcoding, irrespective of q values (Figure 2A, Table S2). Large-volume eDNA metabarcoding detected less species than both other methods over the transects (samples within methods and transects merged), and for increasing q values Hill numbers became comparable to both eDNA metabarcoding and bulk metabarcoding (Gen-A) or eDNA metabarcoding only (Gen-B; Figure 2C).

Species richness detected with eDNA metabarcoding was compared to conventional morphological analysis of CPR samples in transects Morph-A and Morph-B. In both transects, eDNA metabarcoding detected more species than CPR (Morph-A: 1.3 times more; Morph-B: 1.6 times more; Figure 2B, Table S2).

3.4 Beta diversity

3.4.1 eDNA metabarcoding vs bulk metabarcoding

To characterize species composition between genetic transects (Gen-A vs Gen-B, separated by 812 km) and sampling methods (eDNA vs bulk DNA vs large-volume eDNA), a PCoA was calculated based on binary Jaccard distances (Figure 3A). Transects separated strongly along the first axis, whereas eDNA sampling and bulk DNA sampling methods separated along the second axis, indicating that the geographic differences were larger than the differences introduced by sampling methodology. PERMANOVA also revealed strong separation between these groups (Table S3). Large-volume eDNA samples grouped closely with bulk DNA samples. In an additional PCoA based on Bray-Curtis distances taking species abundance into account, transects and sampling methods were not as strongly

445 separated as in the binary Jaccard PCoA, but there was still no overlap between the groups (Figure S4,
446 Table S3).

447 To determine statistically which species contribute most to the separation of the two transects
448 within sampling methods, a SIMPER analysis was performed. The 20 species with the strongest
449 contribution to separation of transects accounted for 90 % (eDNA samples) and 81 % (bulk DNA
450 samples) of total contribution. Among these 20 species, 10 were shared between the sampling
451 methods, indicating that both methods detected similar species contributing strongly to community
452 differentiation (Figure S5).

453 While more species were detected with bulk metabarcoding than with eDNA metabarcoding, the
454 overlap between the two methods was large, in particular in transect Gen-A (87 of 149 species (58%)
455 were detected with both methods in Gen-A, 17 of 41 species (41%) in Gen-B, Figure 3D). The species
456 detected with large-volume eDNA metabarcoding had a larger overlap to the species detected with
457 bulk metabarcoding than eDNA metabarcoding in transect Gen-A(Figure S6 A), whereas in Gen-B, the
458 overlap of species detection with the other two methods was identical, i.e. most species detected with
459 large-volume eDNA metabarcoding were also detected with both other methods, while the few
460 remaining detections were exclusive to large-volume eDNA metabarcoding (Figure S6 B).

461

462 ### 3.4.2 eDNA metabarcoding vs morphological CPR analysis

463 To compare eDNA metabarcoding to morphologically analysed CPR samples in transects Morph-A and
464 Morph-B, a PCoA based on binary Jaccard distances was calculated (Figure 3B). The sampling methods
465 separated along the first axis, while within sampling methods there was a complete overlap of Morph-
466 A and Morph-B samples, indicating that re-sampling the same transect 13 days later resulted in highly
467 repeatable data within sampling methods. PERMANOVA results also indicated a dominant effect of
468 sampling method (Table S3). The overlap of species detection between eDNA metabarcoding and

morphologically analysed CPR samples was low and limited to the phylum Arthropoda (5 of 27 species (19%) were detected with both methods in Morph-A, 5 of 40 species (13%) in Morph-B; Figure 3E).

3.4.3 eDNA only

When comparing eDNA samples of all four transects to each other using a PCoA based on binary Jaccard distances, three groups clearly separated following the geographic pattern of sampling: Gen-A samples and Gen-B samples from the north and south respectively formed separate groups (Figure 3C). The third group included all eDNA samples of the Morph-A and Morph-B transects (Figure 3C). Samples of these two transects were collected in the same open ocean area (half way between Gen-A and Gen-B, see Figure 1A), but 13 days apart. When negative water controls (tap water filtered between eDNA samples) were added to this analysis, seven out of eight water controls did not group with eDNA samples of the same transect (Figure S7), indicating negligible cross-contamination between eDNA samples.

3.5 Taxonomic biases

The comparability of sampling methods depended on the taxonomic level and taxonomic group of comparison. At phylum level, bulk metabarcoding detected more phyla overall than eDNA metabarcoding or large-volume eDNA metabarcoding, and larger proportions of Cnidaria reads than both other sampling methods, while eDNA metabarcoding and large-volume eDNA metabarcoding detected larger proportions of Arthropoda reads than bulk metabarcoding (Figure 4A). At lower taxonomic levels, sampling methods were more comparable in small, more uniformly distributed organisms. For example, the calanoid copepod genera detected with different sampling methods largely overlapped within transects (Figure 4B), whereas there was virtually no overlap in detection of the less uniformly distributed chordate species between sampling methods (Figure 4C).

3.5 Effect of sequencing depth

To determine whether sequencing depth affected the number of metazoan species detected with eDNA metabarcoding, a subset of samples was re-sequenced at a higher sequencing depth (see methods). Repeated samples (eDNA rep) did not detect more species (Figure S8 A, top (q = 0)). However, the repeated samples were sequenced in a separate sequencing run that suffered from low levels of cross contamination between samples, and consequently OTUs with less than 10 reads within a sample were set to 0 in repeated samples (see methods), potentially generating some false negatives. The proportion of metazoan vs non metazoan reads did not change with increased sequencing depth (both 'eDNA rep' (repeated) and 'eDNA rep rar' (repeated rarefied), Figure S8 B). Species composition between repeated samples (including rarefied data) was remarkably similar, further indicating that relatively low sequencing depth (~30 000 reads per eDNA sample) may be sufficient to capture the metazoan diversity present in a sample (Figure S8 C). To fully explore the effect of sequencing depth, technical duplicates of samples should be added to the same sequencing run at different DNA concentrations.

3.6 Persistence of eDNA signal

To investigate how long the eDNA from metazoan organisms may last in the water column, we examined data from copepods known to make diel vertical migrations. As expected, species of the genus *Clausocalanus* (*C. brevipes* and *C. laticeps*) were mainly detected during the night in the morphologically analysed CPR samples of transects Morph-A and Morph-B, (Figure 5), reflecting their diel vertical migration patterns. Two of the three *Clausocalanus* species in the eDNA samples of the same transects showed the same pattern (*C. brevipes* and *C. pergens*), while a third detected species (*C. ingens*) was mainly detected during the day, potentially suggesting a reversed diel vertical

517 migration pattern for this species (Ohman, Frost & Cohen, 1983) (Figure 5). Detected diel vertical
518 migration patterns did not correspond to overall or metazoan sequencing depth in the eDNA samples
519 (Figure S2 B).

520

521 ## 3.7 Comparison of COI and 18S datasets

522 In the eDNA samples, the proportion of metazoan reads detected with COI was much lower than in
523 the other genetic samples (bulk DNA, large-volume eDNA; Figure 1B, Figure S2). To establish what
524 component of the non-metazoan community was detected with COI in the eDNA samples , we
525 compared the COI reads to 18S reads at class-level resolution (Figure 6). Interestingly, the proportion
526 of metazoan reads was larger for 18S than COI, even though the COI primers were specifically designed
527 to capture metazoan diversity (Leray et al., 2013). Large proportions of the non-metazoan COI reads
528 could not be classified to class level, likely due to missing information on databases for non-metazoan
529 species (Bakker et al., 2019). Of the reads resolved to class-level with COI, the majority belonged to
530 Mamiellophyceae, a class of green algae. Most reads within this group were assigned to two globally
531 abundant picophytoplankton species (*Bathycoccus prasinos* and *Micromonas pusilla*), which both
532 have near perfect matches to both forward and reverse primers of the COI marker used (Table S4), In
533 contrast to COI, the 18S marker detected more diverse groups, with the most common ones belonging
534 to Dinophyceae (a class of dinoflagellates), Syndiniales (a group of parasitic dinoflagellates) and
535 Maxillopoda (a class of crustaceans).

536 To assess if particular taxonomic groups of metazoan were differentially represented in the COI
537 and 18S data, all eDNA samples were compared at phylum level. In the metazoan reads, 18S
538 metabarcoding generally detected a larger proportion of Arthropoda reads than COI metabarcoding,
539 as well as small proportions of Xenacoelomorpha reads, which were largely absent in COI
540 metabarcoding results (Figure S9). The non-Arthropoda phyla detected in larger proportions by COI

metabarcoding were primarily from Chordata, Cnidaria and Gastrotrichia. Within the phylum Arthropoda, 18S metabarcoding detected two classes, Maxillopoda and Malacostraca, at similar relative abundances as COI metabarcoding. COI metabarcoding again detected more diversity and further arthropod classes, such as Branchiopoda and Ostracoda. This indicates that COI metabarcoding may overall detect more metazoan diversity in eDNA than 18S metabarcoding.

4 Discussion

4.1 Overview

Our evaluation of metabarcoding approaches to characterize zooplankton biodiversity directly compared eDNA samples taken at a single time point from half of a 2 L water sample (filter split for biobanking) with bulk zooplankton samples filtered continuously over 30 minutes from approximately 1500 L of seawater. The number of species detected with small volume eDNA metabarcoding was roughly two thirds of the species detected through bulk metabarcoding, with a large overlap of detected species. This was the case both for the species-rich northern transect as well as the less diverse southern transect, suggesting that the methods have a consistent rate of species detection across these open ocean environments. Furthermore, eDNA metabarcoding identified similar species contributing strongly to community differentiation between the two transects as bulk metabarcoding, and the three sampled transect regions showed clearly different species compositions in the eDNA samples. When we consider that eDNA was sampled from a tiny fraction of the water, and that eDNA metazoan read numbers were 6 – 30 times smaller than metazoan read numbers from bulk samples (due to increase of non-target COI amplification from eDNA), small volume eDNA metabarcoding was remarkably effective at detecting metazoan species.

We found collecting multiple small-volume (2L) eDNA samples detected more species along a transect than a single large-volume (> 2000 L) eDNA sample filtered through a coarse filter. This is despite the fact that large-volume eDNA samples were sequenced at a sequencing depth similar to the combined reads from all the small volume eDNA samples from a transect and had a much larger proportion of metazoan reads than small volume eDNA samples. Collecting multiple large-volume eDNA samples in parallel might improve species detection rates, however, collecting multiple small volume eDNA samples along a transect gives both higher diversity and higher spatial resolution, so may be more suitable to measure biodiversity than large samples spanning the whole transect.

When comparing eDNA metabarcoding to traditional morphological surveys, eDNA metabarcoding detected 1.3 – 1.6 times more species than morphologically analysed CPR samples and the overlap of species detected between methods was small. The small overlap may partially be explained by missing data in the reference sequence databases. Additionally, while traditional morphological analyses have the strength of being able to identify larval stages of certain taxa (e.g. krill) to species level, there are many other juvenile forms that are only classified broadly (e.g. fish larvae, and eggs). In contrast, taxonomic identification through eDNA metabarcoding can identify organism regardless of developmental stage and does not even require intact specimens, as long as DNA can be recovered and a reference sequence exists. However, eDNA metabarcoding cannot provide information on developmental stages of detected organisms. While species detection differed between sampling methods, the two transects that compared conventional morphological versus eDNA sampling detected very similar species compositions within methods, showcasing that sampling the same transects within short time frames (13 days apart) can deliver repeatable results in open ocean environments.

4.2 Persistence of eDNA signal

One surprising finding was that eDNA data for two copepods (*Clausocalanus* species) showed a diel vertical migration pattern (i.e. the species were only detected at night) which mirrors expected movement of these crustaceans. This suggests that eDNA signals were primarily detected when the organism was present in the water column. This is contrary to findings of marine experimental studies, where eDNA signals were detected days after the organism was removed from the water column (Sassoubre, Yamahara, Gardner, Block & Boehm, 2016; Thomsen et al., 2012). However, recent eDNA studies from ocean environments have found that transfer of eDNA between sampling sites is negligible even on small geographic scales, either due to rapid degradation of eDNA in marine environments, or due to dilution effects resulting in a rapid drop of eDNA concentrations below

detection thresholds (Jeunen et al., 2019a; Jeunen et al., 2019b; Stat et al., 2019). Our observation might also indicate that the eDNA we are detecting from these small copepods is from the whole organism, or faecal pellets that may sink from the surface relatively soon after the organism disappears, rather than from free DNA molecules. Here, we used a relatively long metabarcoding marker to characterize eDNA, which may not be very effective at detecting degraded or extremely diluted eDNA targets. As a result, the quick disappearance of detectable eDNA signals after an organism has passed through provides a snapshot of the immediate composition of the sampled environment. If a more complete picture of the community is desired, multiple samples at different times of the day should be considered to account for diel vertical migration patterns of many species (Brierley, 2014). Furthermore, future studies could examine whether more sensitive, shorter and/or group-specific primers could detect eDNA signals for a longer period in open ocean environments (Rees et al., 2014; Thomsen & Willerslev, 2015).

4.3 Considerations for open ocean eDNA sampling and monitoring

For COI, the proportion of metazoan reads in eDNA samples was much smaller than in bulk DNA or large-volume eDNA samples. The non-specific amplification using COI markers with eDNA template is similar to findings of Collins et al. (2019), and our data from bulk samples show that sampling methods can strongly influence the performance of metabarcoding markers. Several factors may have contributed to the small proportion of metazoan reads in eDNA samples: firstly, the volume of open ocean seawater sampled was very small (2 L), therefore the amount of metazoan DNA per sample was likely minuscule. Secondly, we used 0.45 µm filter, which likely captured a large amount of non-target material that would have passed through the coarse mesh used with the bulk sampling. In eDNA samples the presence of picoplanktonic algae (Mamiellophyceae, 1 – 2 µm diameter) accounted for much of the non-target signal as two Mamiellophyceae species were amplified with the chosen COI marker particularly well, possibly as a result of homoplasy in the primer binding region between very

distantly related taxa (Deagle, Jarman, Coissac, Pompanon & Taberlet 2014). Thirdly, we used a comparatively long COI marker, reducing the potential to detect eDNA from degraded metazoan tissues and preferentially detecting intact cells..

While eDNA metabarcoding was generally remarkably efficient in detecting metazoan species in small water volumes, some taxonomic groups were more readily detected than others and this appears to be related to the size of the organism. The detection of small, relatively evenly distributed organisms such as copepods was comparable between eDNA metabarcoding and bulk metabarcoding. Larger species such as fish, with a less uniform distribution, were often missed in the eDNA samples. Based on the very limited number of fish species we detected with the broad metazoan COI metabarcoding marker, this marker is unlikely to be suitable for open ocean eDNA fish surveys. This seems to be in contrast with studies conducted in species rich, coastal environments such as reefs or bays, where small volume eDNA metabarcoding assays have been effective at detecting fish species (Stat et al., 2019; Yamamoto et al., 2017). However, these studies usually use group-specific primers. Using targeted group-specific assays with higher specificity and sensitivity could clarify whether fish species can be reliably detected in open ocean eDNA samples such as the ones we examined, but it is possible that the eDNA from fish in the open ocean surface water may be so dilute that reliable detection is not feasible for routine monitoring. While there is room to improve eDNA sampling and processing (see below), gaps in detecting certain taxa will likely remain. However, even with a somewhat incomplete picture, eDNA metabarcoding did detect similar species contributing strongly to community differentiation as bulk metabarcoding across open ocean environments. We suggest that a simple, repeatable and cost-effective eDNA sampling and processing approach could become a highly useful and practical monitoring tool to detect major shifts in community compositions across geographic scales and time.

Several steps could be investigated to improve eDNA detection rates for pelagic metazoan species. Firstly, slightly larger water volumes (e.g. ~20 L) could be sampled at each location to increase

metazoan DNA per sample, and several biological replica could be taken in parallel. Secondly, coarser filters (e.g. ~5 µm) could be trialled to filter the seawater, reducing capture of non-target organisms such as picoplankton and bacteria, and preventing filters from clogging when filtering larger water volumes (Robson et al., 2016). Thirdly, shorter metabarcoding markers that are more efficient in detecting low concentration, partially degraded DNA, or group specific primers could be used (Bylemans, Gleeson, Hardy & Furlan, 2018; Jarman, Deagle & Gales, 2004; Miya et al., 2015; Thomsen & Willerslev, 2015). Alternatively, multiple metabarcoding assays could be combined to get more comprehensive results (Berry et al., 2019; Djurhuus et al., 2020; Lacoursière-Roussel et al., 2018; Stat et al., 2017). However, as this substantially increases laboratory work, data analysis and costs, a multi-assay approach has to be carefully weighed up against a more simple, cost-effective protocol, possibly only involving one or two COI markers, particularly in studies processing large numbers of samples. Lastly, PCR parameters could be optimized. Here we used a comparatively low annealing temperature (46 °C) for COI amplifications, as these settings have previously been shown to increase the number of species detected for COI in bulk plankton samples (Clarke et al., 2017). However, in our eDNA samples with very low metazoan template DNA, these relatively unspecific annealing conditions may have contributed to large amounts of off-target amplifications, highlighting that optimal PCR protocols may vary depending on sample types.

The metazoan-targeted COI approach used in this study focussed on taxonomically described species, which allows comparison with existing Southern Ocean morphological datasets. Even with the large COI DNA barcoding database that is available, missing reference sequences limited our ability to identify many sequences we recovered. Other markers providing species-level identifications have considerably fewer reference sequences and would have provided only a fraction of the species matches. The 18S marker we also sequenced is not variable enough to provide species-level taxonomic information but is widely used for characterising marine eukaryotic microbes and a strong case can be made for using this marker in routine open ocean biomonitoring due to its broad coverage. Potential biomonitoring approaches could also rely on taxonomy-free or even function-based methods (Cordier

et al., 2020). While each of these approaches has strengths and weaknesses, the probably most important aspect for ecological monitoring is not which marker is used, but that the sampling methodology is consistent and can be done at appropriate temporal and spatial scales. One of the benefits of eDNA biomonitoring is that the DNA samples can be archived and interrogated using different strategies retrospectively if required (Jarman et al. 2018).

5 Conclusions

Oceanic ecosystems are projected to be strongly affected by climate change, likely leading to a shift in ecosystem compositions (Constable et al., 2014; Poloczanska et al., 2016). To provide a clear view of the impacts on biodiversity and species composition, it is of paramount importance to have a robust, comparable, yet relatively simple biomonitoring system in place that can easily be used by many research groups. Our eDNA assay using a single COI metabarcoding marker and 2 litre water samples detected more zooplankton biodiversity than conventional morphological methods and identified similar species driving community differentiation as bulk plankton metabarcoding. There are many adjustments to eDNA methodology that could be made to fine tune the information obtained from an eDNA monitoring program. If high taxonomic resolution metazoan data (such as that obtainable with the COI marker) is required, sampling with slightly coarser filters might help decrease the amount of non-target single-celled eukaryote signal. Larger water volumes or targeted assays using shorter and more sensitive markers, may increase species detection rates for larger or elusive taxa such as fish. The use of very broad eukaryotic-targeted makers such as the 18S provides a different picture of biodiversity and this may be preferable depending on the ultimate application of the data. With a refinement of eDNA sampling and processing methods and standardization between studies, eDNA sampling has the potential to create an unprecedented biodiversity monitoring capacity for the open ocean.

Acknowledgements

The research was funded by the Australian Antarctic Science Program (AAS Projects 4556 and 4107). We thank the crew aboard the *RSV Aurora Australia* for CPR deployments, James Marthick (Menzies Research Institute, University of Tasmania) for facilitating use of the MiSeq Genome Sequencer and Andrew Bissett (CSIRO Hobart) for 18S data processing. This work was supported by the Australian Government's Business Cooperative Research Centres Programme through the Antarctic Climate and Ecosystems Cooperative Research Centre.

705

706 References

707

708 Alberdi A, Gilbert MTP (2019a) A guide to the application of Hill numbers to DNA-based diversity
709 analyses. *Molecular Ecology Resources*.
710 Alberdi A, Gilbert MTP (2019b) hilldiv: an R package for the integral analysis of diversity based on Hill
711 numbers. *bioRxiv*, 545665.
712 Bakker J, Wangensteen OS, Baillie C, Buddo D, Chapman DD, Gallagher AJ, . . . Mariani S (2019)
713 Biodiversity assessment of tropical shelf eukaryotic communities via pelagic eDNA
714 metabarcoding. *Ecology and Evolution*.
715 Beaugrand G (2005) Monitoring pelagic ecosystems using plankton indicators. *ICES Journal of Marine*
716 *Science, 62*(3), 333-338.
717 Berry TE, Saunders BJ, Coghlan ML, Stat M, Jarman S, Richardson AJ, . . . Bunce M (2019) Marine
718 environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies
719 ecosystem responses to anomalous climatic events. *PLoS Genetics, 15*(2).
720 Blackman RC, Mächler E, Altermatt F, Arnold A, Beja P, Boets P, . . . Jones JI (2019) Advancing the use
721 of molecular methods for routine freshwater macroinvertebrate biomonitoring–the need for
722 calibration experiments. *Metabarcoding and Metagenomics, 3*(49-57.
723 Boussarie G, Bakker J, Wangensteen OS, Mariani S, Bonnin L, Juhel J-B, . . . Robbins WD (2018)
724 Environmental DNA illuminates the dark diversity of sharks. *Science advances, 4*(5),
725 eaap9661.
726 Brierley AS (2014) Diel vertical migration. *Current Biology, 24*(22), R1074-R1076.
727 Brown EA, Chain FJ, Zhan A, MacIsaac HJ, Cristescu ME (2016) Early detection of aquatic invaders
728 using metabarcoding reveals a high number of non-indigenous species in Canadian ports.
729 *Diversity and Distributions, 22*(10), 1045-1059.
730 Brown MV, Van De Kamp J, Ostrowski M, Seymour JR, Ingleton T, Messer LF, . . . Bibiloni-Isaksson J
731 (2018) Systematic, continental scale temporal monitoring of marine pelagic microbiota by
732 the Australian Marine Microbial Biodiversity Initiative. *Scientific data, 5*(180130.
733 Bucklin A, Lindeque PK, Rodriguez-Ezpeleta N, Albaina A, Lehtiniemi M (2016) Metabarcoding of
734 marine zooplankton: prospects, progress and pitfalls. *Journal of Plankton Research, 38*(393-
735 400.
736 Bucklin A, Yeh HD, Questel JM, Richardson DE, Reese B, Copley NJ, Wiebe PH (2019) Time-series
737 metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES*
738 *Journal of Marine Science, 76*(4), 1162-1176.
739 Bylemans J, Gleeson DM, Hardy CM, Furlan E (2018) Toward an ecoregion scale evaluation of eDNA
740 metabarcoding primers: A case study for the freshwater fish biodiversity of the Murray–
741 Darling Basin (Australia). *Ecology and Evolution, 8*(17), 8697-8712.
742 Chao A, Gotelli NJ, Hsieh T, Sander EL, Ma K, Colwell RK, Ellison AM (2014) Rarefaction and
743 extrapolation with Hill numbers: a framework for sampling and estimation in species
744 diversity studies. *Ecological Monographs, 84*(1), 45-67.
745 Cheung WW, Watson R, Pauly D (2013) Signature of ocean warming in global fisheries catch. *Nature,*
746 *497*(7449), 365-368.
747 Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure.
748 *Australian Journal of Ecology, 18*(1), 117-143.
749 Clarke LJ, Beard JM, Swadling KM, Deagle BE (2017) Effect of marker choice and thermal cycling
750 protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution, 7*(3), 873-883.
751 Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, . . . Mariani S (2019) Non-
752 specific amplification compromises environmental DNA metabarcoding with COI. *Methods in*
753 *ecology and evolution, 10*(11), 1985-2001.

754 Constable AJ, Melbourne-Thomas J, Corney SP, Arrigo KR, Barbraud C, Barnes DK, . . . Ziegler P (2014)
755 Climate change and Southern Ocean ecosystems I: how changes in physical habitats directly
756 affect marine biota. *Glob Chang Biol, 20*(10), 3004-3025.
757 Cordier T, Alonso-Sáez L, Apothéloz-Perret-Gentil L, Aylagas E, Bohan DA, Bouchez A, . . . Keck F
758 (2020) Ecosystems monitoring powered by environmental genomics: a review of current
759 strategies with an implementation roadmap. *Molecular Ecology*.
760 de Broyer C, Danis B (2014) SCAR-MarBIN: the Antarctic marine biodiversity information network.
761 De Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, . . . Probert I (2015) Eukaryotic plankton
762 diversity in the sunlit ocean. *Science, 348*(6237), 1261605.
763 Deagle BE, Clarke LJ, Kitchener JA, Polanowski AM, Davidson AT (2018) Genetic monitoring of open
764 ocean biodiversity: An evaluation of DNA metabarcoding for processing continuous plankton
765 recorder samples. *Molecular Ecology Resources, 18*(3), 391-406.
766 Deiner K, Walser J-C, Mächler E, Altermatt F (2015) Choice of capture and extraction methods affect
767 detection of freshwater biodiversity from environmental DNA. *Biological Conservation,*
768 *183*(53-63.
769 Djurhuus A, Closek CJ, Kelly RP, Pitz KJ, Michisaki RP, Starks HA, . . . Hubbard K (2020) Environmental
770 DNA reveals seasonal shifts and potential interactions in a marine community. *Nature*
771 *Communications, 11*(1), 1-9.
772 Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics, 26*(19),
773 2460-2461.
774 Egerton J, Johnson A, Turner J, LeVay L, Mascareñas-Osorio I, Aburto-Oropeza O (2018)
775 Hydroacoustics as a tool to examine the effects of Marine Protected Areas and habitat type
776 on marine fish communities. *Scientific reports, 8*(1), 47.
777 Frederiksen M, Edwards M, Richardson AJ, Halliday NC, Wanless S (2006) From plankton to top
778 predators: bottom-up control of a marine food web across four trophic levels. *Journal of*
779 *Animal Ecology, 75*(6), 1259-1268.
780 Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, . . . Cornman RS (2016)
781 Critical considerations for the application of environmental DNA methods to detect aquatic
782 species. *Methods in ecology and evolution, 7*(11), 1299-1307.
783 Gordon DP, Beaumont J, MacDiarmid A, Robertson DA, Ahyong ST (2010) Marine biodiversity of
784 aotearoa New Zealand. *PLoS One, 5*(8).
785 Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S (2012) Assessing biodiversity of a freshwater
786 benthic macroinvertebrate community through non-destructive environmental barcoding of
787 DNA from preservative ethanol. *BMC Ecology, 12*(1), 28.
788 Hirai J, Kuriyama M, Ichikawa T, Hidaka K, Tsuda A (2015) A metagenetic approach for revealing
789 community structure of marine planktonic copepods. *Molecular Ecology Resources, 15*(1),
790 68-80.
791 Hoegh-Guldberg O, Bruno JF (2010) The impact of climate change on the world's marine ecosystems.
792 *Science, 328*(5985), 1523-1528.
793 Hosie GW, Fukuchi M, Kawaguchi S (2003) Development of the Southern Ocean continuous plankton
794 recorder survey. *Progress in Oceanography, 58*(2-4), 263-283.
795 Hosie GW, Schultz MB, Kitchener JA, Cochran TG, Richards K (2000) Macrozooplankton community
796 structure off East Antarctica (80-150 degrees E) during the Austral summer of 1995/1996.
797 *Deep-Sea Research Part Ii-Topical Studies in Oceanography, 47*(12-13), 2437-2463.
798 Hunt BP, Hosie GW (2006) Continuous Plankton Recorder flow rates revisited: clogging, ship speed
799 and flow meter design. *Journal of Plankton Research, 28*(9), 847-855.
800 Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, . . . Tappu R (2016) MEGAN community
801 edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS*
802 *Computational Biology, 12*(6), e1004957.
803 Hutchings P (2019) An advocate for taxonomic research in Australia. *Pacific conservation biology,*
804 *25*(1), 34-36.

Jarman SN, Berry O, Bunce M (2018) The value of environmental DNA biobanking for long-term biomonitoring. *Nature ecology & evolution, 2*(8), 1192.

Jarman SN, Deagle BE, Gales NJ (2004) Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology, 13*(5), 1313-1322.

Jeunen GJ, Knapp M, Spencer HG, Lamare MD, Taylor HR, Stat M, . . . Gemmell NJ (2019a) Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Molecular Ecology Resources, 19*(2), 426-438.

Jeunen GJ, Lamare MD, Knapp M, Spencer HG, Taylor HR, Stat M, . . . Gemmell NJ (2019b) Water stratification in the marine biome restricts vertical environmental DNA (eDNA) signal dispersal. *Environmental DNA*.

Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, . . . Edwards FA (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters, 16*(10), 1245-1257.

Kholia B, Fraser-Jenkins C (2011) Misidentification makes scientific publications worthless–save our taxonomy and taxonomists. *Current Science, 100*(4), 458-461.

Kloser R, Ryan T, Sakov P, Williams A, Koslow J (2002) Species identification in deep water using multiple acoustic frequencies. *Canadian Journal of Fisheries and Aquatic Sciences, 59*(6), 1065-1077.

Koubbi P, Ozouf-Costaz C, Goarant A, Moteki M, Hulley P-A, Causse R, . . . Tavernier E (2010) Estimating the biodiversity of the East Antarctic shelf and oceanic zone for ecoregionalisation: example of the ichthyofauna of the CEAMARC (Collaborative East Antarctic Marine Census) CAML surveys. *Polar Science, 4*(2), 115-133.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution, 35*(6), 1547-1549.

Lacoursière-Roussel A, Howland K, Normandeau E, Grey EK, Archambault P, Deiner K, . . . Bernatchez L (2018) eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecology and Evolution, 8*(16), 7763-7777.

Leduc N, Lacoursière-Roussel A, Howland KL, Archambault P, Sevellec M, Normandeau E, . . . Simard N (2019) Comparing eDNA metabarcoding and species collection for documenting Arctic metazoan biodiversity. *Environmental DNA, 1*(4), 342-358.

Leray M, Knowlton N (2016) Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1702), 20150331.

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, . . . 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*(1), 34.

Letunic I, Bork P (2006) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics, 23*(1), 127-128.

Lo W-T, Shih C-T, Hwang J-S (2004) Diel vertical migration of the planktonic copepods at an upwelling station north of Taiwan, western North Pacific. *Journal of Plankton Research, 26*(1), 89-97.

López-Escardó D, Paps J, De Vargas C, Massana R, Ruiz-Trillo I, Del Campo J (2018) Metabarcoding analysis on European coastal samples reveals new molecular metazoan diversity. *Scientific reports, 8*(1), 1-14.

Macher JN, Vivancos A, Piggott JJ, Centeno FC, Matthaei CD, Leese F (2018) Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources, 18*(6), 1456-1468.

Madden T (2013) The BLAST sequence analysis tool. In: *The NCBI Handbook*. National Center for Biotechnology Information (US), Bethesda (MD).

855 Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome
856 assemblies. *Bioinformatics, 27*(21), 2957-2963.
857 Maschette D, Sumner M, Raymond B (2019) SOmap: Southern Ocean maps. Version 0.5.0.
858 McLeod DJ, Hosie GW, Kitchener JA, Takahashi KT, Hunt BP (2010) Zooplankton atlas of the Southern
859 Ocean: the SCAR SO-CPR survey (1991–2008). *Polar Science, 4*(2), 353-385.
860 McQuatters-Gollop A, Atkinson A, Aubert A, Bedford J, Best M, Bresnan E, . . . Johns DG (2019)
861 Plankton lifeforms as a biodiversity indicator for regional-scale assessment of pelagic
862 habitats for policy. *Ecological Indicators, 101*(913-925.
863 Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, . . . Araki H (2015) MiFish, a set of universal
864 PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230
865 subtropical marine species. *Royal Society open science, 2*(7), 150088.
866 Morgan M, Anders S, Lawrence M, Aboyoun P, Pages H, Gentleman R (2009) ShortRead: a
867 bioconductor package for input, quality assessment and exploration of high-throughput
868 sequence data. *Bioinformatics, 25*(19), 2607-2608.
869 O'Hara TD, Williams A, Althaus F, Ross AS, Bax NJ (2020) Regional-scale patterns of deep seafloor
870 biodiversity for conservation assessment. *Diversity and Distributions*.
871 Ohman MD, Frost BW, Cohen EB (1983) Reverse diel vertical migration: an escape from invertebrate
872 predators. *Science, 220*(4604), 1404-1407.
873 Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests M (2007) The vegan
874 package. *Community ecology package, 10*(631-637.
875 Orsi AH, Whitworth III T, Nowlin Jr WD (1995) On the meridional extent and fronts of the Antarctic
876 Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers, 42*(5), 641-
877 673.
878 Pawlowski J, Kelly-Quinn M, Altermatt F, Apothéloz-Perret-Gentil L, Beja P, Boggero A, . . . Domaizon
879 I (2018) The future of biotic indices in the ecogenomic era: Integrating (e) DNA
880 metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total*
881 *Environment, 637*(1295-1310.
882 Piredda R, Tomasino M, D'erchia A, Manzari C, Pesole G, Montresor M, . . . Zingone A (2017)
883 Diversity and temporal patterns of planktonic protist assemblages at a Mediterranean Long
884 Term Ecological Research site. *FEMS Microbiology Ecology, 93*(1), fiw200.
885 Poloczanska ES, Burrows MT, Brown CJ, García Molinos J, Halpern BS, Hoegh-Guldberg O, . . .
886 Schoeman DS (2016) Responses of marine organisms to climate change across oceans.
887 *Frontiers in Marine Science, 3*(62.
888 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, . . . Glöckner FO (2012) The SILVA
889 ribosomal RNA gene database project: improved data processing and web-based tools.
890 *Nucleic Acids Research, 41*(D1), D590-D596.
891 R Core Team (2018) R: A language and environment for statistical computing. R Foundation for
892 Statistical Computing, Vienna, Austria.
893 Raes EJ, Bodrossy L, Van De Kamp J, Bissett A, Ostrowski M, Brown MV, . . . Waite AM (2018)
894 Oceanographic boundaries constrain microbial diversity gradients in the South Pacific Ocean.
895 *Proceedings of the National Academy of Sciences, 115*(35), E8266-E8275.
896 Ratcliffe FC, Webster TMU, Barreto DR, O'Rorke R, de Leaniz CG, Consuegra S (2020) Rapid
897 quantitative assessment of fish larvae community composition using metabarcoding.
898 *bioRxiv*.
899 Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System
900 (http://www.barcodinglife.org). *Molecular Ecology Notes, 7*(3), 355-364.
901 Rees HC, Maddison BC, Middleditch DJ, Patmore JR, Gough KC (2014) The detection of aquatic
902 animal species using environmental DNA–a review of eDNA as a survey tool in ecology.
903 *Journal of Applied Ecology, 51*(5), 1450-1459.

Reid P, Colebrook J, Matthews J, Aiken J, Team CPR (2003) The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. *Progress in Oceanography, 58*(2-4), 117-173.

Robson HL, Noble TH, Saunders RJ, Robson SK, Burrows DW, Jerry DR (2016) Fine-tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems. *Molecular Ecology Resources, 16*(4), 922-932.

Ruppert KM, Kline RJ, Rahman MS (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*, e00547.

Sassoubre LM, Yamahara KM, Gardner LD, Block BA, Boehm AB (2016) Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environmental Science & Technology, 50*(19), 10456-10464.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, . . . Robinson CJ (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology, 75*(23), 7537-7541.

Sepulveda AJ, Schabacker J, Smith S, Al-Chokhachy R, Luikart G, Amish SJ (2019) Improved detection of rare, endangered and invasive trout in using a new large-volume sampling method for eDNA capture. *Environmental DNA, 1*(3), 227-237.

Sigsgaard EE, Nielsen IB, Bach SS, Lorenzen ED, Robinson DP, Knudsen SW, . . . Willerslev E (2017) Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature ecology & evolution, 1*(1), 0004.

Sikes E, Howard W, Samson C, Mahan T, Robertson L, Volkman J (2009) Southern Ocean seasonal temperature and Subtropical Front movement on the South Tasman Rise in the late Quaternary. *Paleoceanography, 24*(2).

Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, . . . Bunce M (2017) Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific reports, 7*(1), 1-11.

Stat M, John J, DiBattista JD, Newman SJ, Bunce M, Harvey ES (2019) Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology, 33*(1), 196-205.

Suter L, Polanowski AM, Clarke LJ, Kitchener JA, Deagle BE (2020) Data for manuscript "Capturing open ocean biodiversity: comparing environmental DNA metabarcoding to the continuous plankton recorder". Dryad, Dataset. https://doi.org/10.5061/dryad.w3r2280mq.

Taberlet P, Bonin A, Zinger L, Coissac E (2018) *Environmental DNA: For Biodiversity Research and Monitoring*: Oxford University Press.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology, 21*(8), 2045-2050.

Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D (2012) The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proceedings of the National Academy of Sciences, 109*(40), 16208-16212.

Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One, 7*(8).

Thomsen PF, Møller PR, Sigsgaard EE, Knudsen SW, Jørgensen OA, Willerslev E (2016) Environmental DNA from seawater samples correlate with trawl catches of subarctic, deepwater fishes. *PLoS One, 11*(11), e0165252.

Thomsen PF, Willerslev E (2015) Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183*(4-18.

953 Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA
954 sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology,*
955 *73*(16), 5261-5267.
956 Wickham H (2016) *ggplot2: Elegant graphics for data analysis.* New York, USA: Springer.
957 Yamahara KM, Preston CM, Birch JM, Walz KR, Marin III R, Jensen S, . . . Ryan J (2019) In-situ
958 autonomous acquisition and preservation of marine environmental DNA using an
959 autonomous underwater vehicle. *Frontiers in Marine Science, 6*(373.
960 Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, . . . Miya M (2017) Environmental DNA
961 metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific reports,*
962 *7*(1), 1-12.

963

964

Data accessibility statement

The raw sequencing data, sequence processing pipeline, R markdown code for all analyses performed in R as well as all processed data is openly available on Dryad at https://doi.org/10.5061/dryad.w3r2280mq (Suter et al., 2020).

Author contributions

L.S. contributed to study design, led bioinformatics and data analysis and wrote the manuscript. A.M.P. contributed to study design, sample collection and data analysis and performed genetic laboratory work. L.J.C. contributed to bioinformatics and data analysis. J.A.K. contributed to sample collection and performed morphological data analysis. B.E.D. led study design and sample collection and advised on laboratory work, bioinformatics and data analysis. All authors contributed to writing and editing the manuscript.

Figure legends

Figure 1: **A**) Map of the transects sampled between Tasmania and Macquarie island in this study. **B**) Number of sequencing reads from the COI marker classified as "metazoan", "non-metazoan" and "not classified", separately for transects and sampling methods (eDNA: metabarcoding of environmental DNA samples; bulk DNA: DNA metabarcoding of bulk continuous plankton recorder samples; LV-eDNA: metabarcoding of large-volume eDNA samples). The map was drawn using the R package "SOmap" (Maschette, Sumner & Raymond, 2019) and the R package "orsifronts" to add the subtropical front (STF) (Orsi, Whitworth III & Nowlin Jr, 1995).

Figure 2: Comparison of metazoan alpha-diversity detected in transects with different sampling methods. **A**) Alpha-diversity detected in genetic transects (Gen-A and Gen-B) with metabarcoding of environmental DNA ("eDNA") samples and DNA metabarcoding of bulk continuous plankton recorder samples ("bulk DNA"), both characterised with the COI marker. Hill numbers were calculated for q = 0 (equivalent to species richness), q = 1 (equivalent to Shannon diversity) and q = 2 (equivalent to Simpson diversity), taking species abundance increasingly into account. **B**) Comparison of species richness (Hill numbers calculated with q = 0) detected in morphological transects (Morph-A and Morph-B) with metabarcoding of eDNA samples and morphologically analysed continuous plankton recorder samples ("morph. CPR"). **C**) Diversity profiles for genetic transects, summarising alpha diversity detected across all eDNA and bulk DNA samples of one transect with increasing q values, as well as showing the diversity of the single large-volume eDNA ("LV-eDNA") sample spanning the whole transect.

Figure 3: Comparison of metazoan beta-diversity detected across transects with different sampling methods; all genetic data come from metabarcoding using the COI marker. **A**) Principal coordinate analysis (PCoA) for genetic transects (Gen-A and Gen-B) based on binary Jaccard distances for environmental DNA samples ("eDNA"), genetically analysed bulk continuous plankton recorder samples ("bulk DNA") and large-volume eDNA samples ("LV-eDNA"). **B**) Binary Jaccard PCoA for morphological transects (Morph-A and Morph-B), for eDNA samples and morphologically analysed continuous plankton recorder samples ("morph. CPR"). **C**) Binary Jaccard PCoA of eDNA samples across all four transects. **D**) Number of species detected per metazoan phyla for genetic transects (Gen-A and Gen-B) with eDNA metabarcoding only, bulk metabarcoding only, or both methods. **E**) Number of species detected per metazoan phyla for morphological transects (Morph-A and Morph-B) with eDNA only, morphologically analysed CPR samples only or both methods.

Figure 4: Impact of taxonomic level and taxonomic group on the comparison of the sampling methods; all genetic data come from metabarcoding using the COI marker. **A**) Proportion of sequence reads of metazoan phyla, per transect (Gen-A, Gen-B, Morph-A, Morph-B) and sampling method (eDNA: environmental DNA metabarcoding; bulk DNA: bulk metabarcoding; LV-eDNA: large-volume eDNA metabarcoding;morph. CPR: morphologically analysed continuous plankton recorder samples. **B**) Proportion of sequence reads of calanoid copepod genera per transect and sampling method. **C**) Proportion of sequence reads of chordate species per transect and sampling method.

Figure 5: Persistence of eDNA signal in the water column. Species of the copepod genus *Clausocalanus* are known to make diel vertical migrations (DVM), i.e. they forage in surface waters at night and move to deeper layers during the day to avoid predation. In transects Morph-A and

Morph-B, *Clausocalanus* species in morphological continuous plankton recorder (mCPR) samples were predominantly detected during the night (bottom), clearly expressing DVM. In paired environmental DNA (eDNA) samples, two *Clausocalanus* species show the same DVM pattern, while a third species (*C. ingens*) is mainly detected during the day, potentially expressing reverse DVM.

Figure 6: Comparison of metazoan and non-metazoan diversity detected with the COI marker and the 18S marker at class level resolution in eDNA samples of all four transects (Gen-A, Gen-B, Morph-A, Morph-B) to establish what components of non-metazoan diversity were captured using COI.