

# A 3-year plankton DNA metabarcoding survey reveals marine biodiversity patterns in Australian coastal waters

Tina E. Berry<sup>1,2</sup>  | Megan L. Coghlan<sup>1</sup> | Benjamin J. Saunders<sup>3</sup>  |  
 Anthony J. Richardson<sup>4,5</sup>  | Matthew Power<sup>1</sup>  | Euan Harvey<sup>3</sup>  | Simon Jarman<sup>1</sup>  |  
 Oliver Berry<sup>6</sup>  | Claire H. Davies<sup>7</sup>  | Michael Bunce<sup>1,8</sup> 

<sup>1</sup>Trace and Environmental DNA (TrEnD) Laboratory, School Molecular and Life Sciences, Curtin University, Bentley, Western Australia, Australia

<sup>2</sup>eDNA Frontiers, Curtin University, Bentley, Western Australia, Australia

<sup>3</sup>Fish Ecology Laboratory, School Molecular and Life Sciences, Curtin University, Bentley, Western Australia, Australia

<sup>4</sup>CSIRO Environment, Queensland BioSciences Precinct (QBP), St Lucia, Queensland, Australia

<sup>5</sup>School of Environment, The University of Queensland, St Lucia, Queensland, Australia

<sup>6</sup>CSIRO Environomics Future Science Platform, Indian Ocean Marine Research Centre, The University of Western Australia, Perth, Western Australia, Australia

<sup>7</sup>CSIRO Oceans and Atmosphere, Hobart, Tasmania, Australia

<sup>8</sup>Department of Conservation, Wellington, Aotearoa, New Zealand

## Correspondence

Tina E. Berry, Trace and Environmental DNA (TrEnD) Laboratory, School Molecular and Life Sciences, Curtin University, Bentley, WA, Australia.  
 Email: [tina.berry@curtin.edu.au](mailto:tina.berry@curtin.edu.au)

## Funding information

Australian Research Council, Grant/Award Number: LP160100839 and LP160101508

Editor: Trishna Dutta

## Abstract

**Aim:** To use a long-term collection of bulk plankton samples to test the capacity of DNA metabarcoding to characterize the spatial and seasonal patterns found within a range of zooplankton communities, and investigate links with concurrent abiotic data collected as part of Australia's Integrated Marine Observing System (IMOS) programme.

**Location:** Samples were sourced seasonally for 3 years from nine Pan-Australian marine sites (n = 90).

**Methods:** Here, we apply a multi-assay metabarcoding approach to environmental DNA extracted from bulk plankton samples. Six assays (targeting 16S rRNA and COI genes) were used to target, amplify and sequence the zooplankton diversity found within each sample. The data generated from each assay were filtered and clustered into OTUs prior to analysis. Abiotic IMOS data collected contemporaneously enabled us to explore the physical and chemical drivers of community composition.

**Results:** From over 25 million sequences, we identified in excess of 500 distinct taxa and detected clear spatial differences. We found that site and sea surface temperature are the most consistent predictors of differences between zooplankton communities. We detected endangered and invasive species such as the bryozoan *Membranipora membranacea* and the mollusc *Maoricolpus roseus*, and seasonal occurrences of species such as humpback whales (*Megaptera novaeangliae*). We also estimated the number of samples required to detect any significant seasonal changes. For OTU richness, this was found to be assay dependent and for OTU assemblage, a minimum of nine samples per season would be required.

**Main Conclusion:** Our results demonstrate the ability of DNA to capture and map zooplankton community changes in response to seasonal and spatial stressors and provide vital evidence to environmental stakeholders. We confirm that a metabarcoding method offers a practical opportunity for an ecosystem-wide approach to long-term biomonitoring and understanding marine biomes where morphological analysis is not feasible.

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.

© 2023 The Authors. *Diversity and Distributions* published by John Wiley & Sons Ltd.

## KEYWORDS

bulk plankton, environmental DNA, marine biodiversity, metabarcoding, zooplankton

## 1 | INTRODUCTION

Human use of terrestrial ecosystems has accelerated rates of animal extinction and habitat loss throughout the Holocene (Di Marco et al., 2018; Jackson et al., 2001; McCauley et al., 2015). However, until recently, the inaccessibility of many marine environments has limited the exploitation of many marine faunas—particularly in deep seas and oceanic regions (see McCauley et al. (2015)). This protection ended with industrialization and the advent of commercial fisheries (Lotze et al., 2006; McCauley et al., 2015). Anthropogenic climate change also threatens marine species, and the ongoing transformation of marine environments in response to rising ocean temperatures is expected to have profound ecological (Edwards & Richardson, 2004; Gattuso et al., 2015; Griffiths et al., 2017) and economic and societal consequences (Beaugrand et al., 2002; Wernberg et al., 2016). Ocean acidification, caused by increased atmospheric CO<sub>2</sub> concentrations, further compounds these effects (Gattuso et al., 2015; Richardson, 2009). Taken together, these human impacts demonstrate the pressing need for effective marine management, including monitoring the state and trends of ecosystems and protecting of marine life for future generations (Edwards et al., 2010; Gattuso et al., 2015).

For practical reasons, many marine monitoring programmes concentrate on larger organisms, such as fish, seabirds and marine mammals (Lenanton et al., 2017; Richardson, 2009; Richardson et al., 2012). Yet these animals, for the most part, have long lifespans and comprise the higher end of the marine food chain, meaning they respond relatively slowly to ecosystem change. By contrast, zooplankton communities respond rapidly to changing conditions and are orders of magnitude more abundant than higher trophic levels, and make a more significant contribution to ocean nutrient cycling (Richardson, 2009). Zooplankton communities comprise mostly microscopic and short-lived animals that drift in ocean currents. They are a critical trophic link between the marine autotrophs (phytoplankton) and higher trophic levels that are typically harvested or valued by humans. Because of the essential role zooplankton have in marine ecosystems, their sensitivity to environmental change and relatively easy collection, zooplankton surveys have been used in marine monitoring for almost 100 years (Batten et al., 2019; Edwards et al., 2010; Richardson, 2009).

Through necessity, morphology has historically been the chief method for zooplankton species-level identification, but the process is time-consuming, prone to error and is difficult for juvenile life stages and damaged specimens (Deagle et al., 2017; Lindeque et al., 2013) and requires considerable taxonomic expertise. New automated methods such as ZooScan (Gorsky et al., 2010; Rohner et al., 2015) and Laser Optical Plankton Counters (Herman et al., 2004) have been developed, but neither of these methods can

identify zooplankton to a species level (Everett et al., 2017). DNA-based species identification has been used for many years as a tool to identify individual taxa. This barcoding process is accurate, relatively rapid and negates the need for taxonomic expertise. However, this approach is not feasible for taxa as small and varied as zooplankton. The development of metabarcoding, advanced computer technologies and reference sequence databases, together with progressive analytics, allows for the identification of a multitude of taxa from just one sample of DNA extracted from an environmental substrate. This metabarcoding approach is much more suitable for targeting zooplankton taxa and can be used for endangered, cryptic, reclusive or invasive taxa.

As a highly diverse complex mixture, genomic DNA extracted from a bulk sample (such as plankton) is but one type of environmental DNA (Taberlet et al., 2018). While the majority of the DNA extracted from a bulk plankton sample is organismal, much of it is extra-organismal. This is especially the case where the plankton is collected by filtering a large amount of water through a drop net or filtering across a transect, a process which has since been described as large-volume eDNA sampling (Suter et al., 2020). Ongoing research using eDNA metabarcoding technology has demonstrated the ability to obtain useful genetic data from a range of marine environmental substrates such as seawater (Alexander et al., 2019; Andruszkiewicz et al., 2017; Stat et al., 2017), scat (Berry et al., 2017; Casper et al., 2007; Deagle et al., 2005; Hardy et al., 2017) and sediment (Guardiola et al., 2015; Morard et al., 2017). However, only a relatively small number of metabarcoding studies of biodiversity use plankton as a substrate (Alberti et al., 2017; Berry et al., 2019; Deagle et al., 2017; Gimmler et al., 2016; Lindeque et al., 2013; Richardson et al., 2019; Suter et al., 2020). This is, in part, because many existing barcode markers lack the balance between the taxonomic coverage and the resolution required to thoroughly characterize this highly phylogenetically diverse substrate (Brown et al., 2015; Clarke et al., 2017; Lindeque et al., 2013); zooplankton samples commonly include specimens from a dozen phyla and high species diversity (Richardson et al., 2019).

In 2019, Berry et al. demonstrated that a multi-assay approach to metabarcoding could be used to map seasonal changes and heat-wave responses across a 5-year period using collections of zooplankton from a single site in Western Australia. Here, we extend the spatial extent of analysis and take six metabarcoding assays (Table S1) and apply them to 90 plankton samples taken from nine Australian National Reference Stations (NRS) over a 2-year period. The NRS network has been maintained by the Integrated Marine Observing System (IMOS) since 2008, as an integral part of a national programme of environmental monitoring (Lynch et al., 2014). The locations of the NRS range from tropical to temperate areas, and encompass the six marine biogeographical regions that surround

Australia (Figure 1). The NRS range also includes Australia's two principal poleward-flowing currents; the East Australian Current and the Leeuwin Current, on the western side of the country (Lynch et al., 2014). Our aim is to further test the capacity of DNA (using a multi-assay approach) to characterize the spatial and seasonal patterns found within the zooplankton communities across Australia and investigate links with concurrent abiotic data collected as part of the IMOS programme. In doing so, we seek to test the efficacy of the method for use in the ongoing monitoring of zooplankton communities.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling details and subsequent DNA extraction

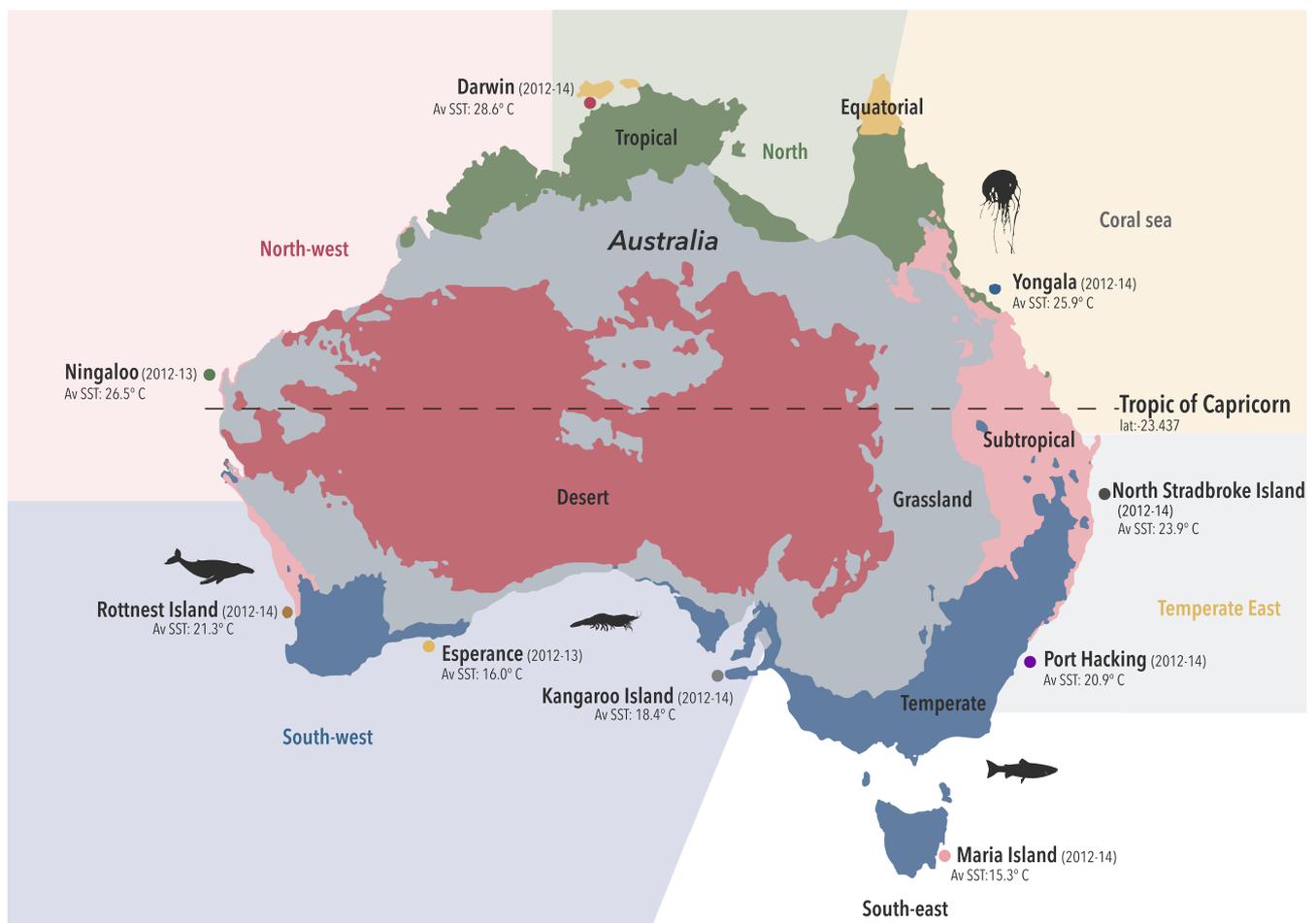
Samples were collected from nine IMOS National Reference Stations (Figure 1), with at least one site to represent each Australian marine biogeographical area. Samples were collected and stored with the forethought that they would be used in the future for genetic testing. Three sites are situated north of the Tropic of Capricorn and

six are south. The marine environments sampled vary from tropical/equatorial (Darwin—mean SST 28.6°C) to temperate (Maria Island—mean SST 15.3°C—Figure 1). Since mid-2010, sampling has taken place monthly at most of these sites and seasonally at Ningaloo and Esperance (Lynch et al., 2014). We selected 90 seasonal samples from the IMOS collection from 2012 to 2014, averaging 6–12 samples per site (a maximum of three samples from each season per site).

Plankton samples were collected using a 0.6-m-wide, 3-m-long drop net (Eriksen et al., 2019) with a 100µm mesh, which falls at 1 ms<sup>-1</sup>. The seabed depth at each site varied between 20 and 100 m (Lynch et al., 2014) and the drop net samples to a depth of 5 m above the seabed. Plankton is collected on the downward drop only—the net is closed as it is recovered to the surface.

Seawater was used to wash the samples into the cod end of the drop net for transfer to the sample jar. After collection, samples were kept on ice until filtered to remove excess water and stored at -80°C immediately upon return to the lab. After subsampling, they were kept at -20°C. Between sampling, the nets were cleaned in freshwater, before drying and storage.

Thawed plankton samples were homogenized using a handheld blender (OMNI Tip™ Homogenizer) and a hard tissue probe. About 40µL of each sample was then digested and extracted with the



**FIGURE 1** The Australian Integrated Marine Observing System (IMOS) sampling sites. The map includes adapted Köppen climate classifications, marine biogeographical areas and, for each site, the sampling years (in brackets) and mean sea surface temperatures (Av SST).

tissue protocol from the DNAeasy Blood and Tissue kit (Qiagen). Modifications included twofold quantities of the reagents used before the washing stage and a double elution of AE buffer (200  $\mu$ L total). Four extraction controls were created during the extraction stages to be used throughout data production. DNA extracts were kept at  $-20^{\circ}\text{C}$ . All samples were extracted and processed within laboratories designed for trace and environmental DNA.

## 2.2 | Amplification of barcodes, library build and sequencing

Six metabarcoding assays (Table S1) were applied to 1 in 10 dilutions of each DNA extract. For each assay, DNA extracts were assigned fusion-tagged primers incorporating assay-specific primers, Illumina adaptor sequences and unique combinations of six to eight base MID (Multiplex Identifier) tags—over 500 unique combinations. The MID tags identify sequences to a particular assay and sample. To avoid cross-contamination, the MID tag combinations had not previously been used for marine samples. Fusion tags were used to limit the risk of tag jumping by the use of only one round of PCR (Schnell et al., 2015). Each PCR-tagging reaction comprised: 1  $\times$  Taq Gold buffer (Applied Biosystems, ABI, USA), 2 mM  $\text{MgCl}_2$  (ABI, USA), 0.4 mg/mL BSA (Fisher Biotech, Australia), 0.25 mM dNTPs (Astral Scientific, Australia), 0.4  $\mu$ M each of forward and reverse primers (Integrated DNA Technologies, Australia), 0.6  $\mu$ L of 1/10,000 SYBR Green dye (Life Technologies, USA) and 1 U of Taq polymerase Gold (ABI, USA), 2  $\mu$ L of DNA and made up to 25  $\mu$ L with PCR grade water. PCR master mixes with tagged primers were prepared within an ultraclean environment to prevent contamination. DNA was added in a pre-PCR laboratory and the PCR reactions occurred in a third dedicated laboratory. Non-template and extraction controls were included on each PCR plate, only those controls which amplified were sequenced. PCR reactions were duplicated (to allow for stochasticity) and cycled at  $95^{\circ}\text{C}$  for 5 min followed by 50 cycles of  $95^{\circ}\text{C}$  for 30 s; the primer-specific Ta (annealing temperature; Table S1) for 30 s and  $72^{\circ}\text{C}$  for 45 s; and a final extension of  $72^{\circ}\text{C}$  for 10 min. Tagged amplicons were combined in roughly equimolar concentrations to produce several sequencing libraries. None of the fusion tag combinations on each of these libraries were 'saturated'. Meaning that significant unused tag combinations could be detected and the tagging process repeated if required. The libraries were filtered for size using a Pippin Prep (Sage Sciences, MA, USA) instrument (2% Agarose cassette with ethidium bromide) and quantified with a Qubit fluorometer (Thermo Fisher, Aus). All sequencing was completed on Illumina's MiSeq® (CA, USA) using the manufacturer's protocol (300 cycle V2 reagents and standard flow cell).

## 2.3 | Production of operational taxonomic units (OTUs)

The process of creating OTUs was repeated for each assay. Sequences were separated into their sample of origin by their MID

tags using Geneious R8 (Kearse et al., 2012). Initial filtering was conducted by ensuring the MID tags, gene-specific primers and sequencing adapters were all present in each sorted sequence without error. Sequences failing this step were removed from future analyses. Primers, adaptors and MID tags were detached from the remaining sequences. Each sequence was renamed to reflect its sample of origin. The filtered, trimmed and renamed sequences were then combined into one large assay-specific file for further quality filtering and clustering into OTUs.

USEARCH v8 (Edgar, 2010) was used for quality filtering and the clustering of OTUs. Sequences were filtered based on the quality (Q) scores inherent within fastq sequences, using a fastq filter—E\_max  $>0.5$ —and formed into groups of identical sequences, removing those with an abundance of  $<0.1\%$  of the total number of unique sequences across all samples. OTUs were clustered for each assay using a 97% similarity. USEARCH's cluster-otus command also identifies chimeric sequences, enabling them to be located and removed from the data. Low-abundance filtered sequences were then mapped back onto existing OTUs to ensure the inclusion of all relevant data. Any low-abundance unmapped amplicons were discarded from further analysis. This process, while potentially eliminating some scarce taxa, ensures the removal of possible erroneous amplicons. OTUs were then assigned to samples and converted to presence/absence data ready for potential identification and statistical analysis. The quality filtered sequence data are available on Data Dryad (<https://doi.org/10.5061/dryad.931zcrjns>).

## 2.4 | Taxonomic assignment of OTUs

OTU consensus sequences were searched against the National Center for Biotechnology Information (NCBI) GenBank nucleotide database (Nov 3 2017; Benson et al., 2014) using BLASTn (Basic Local Alignment Search Tool [Altschul et al., 1990]) with the parameters: num\_descriptions 50—num\_alignments 50—dust no and a reward of 1 (for aligned pairs of identical nucleotides). Output files from the GenBank search were imported into MEGAN v5 (MEtaGenome Analyser; [Huson et al., 2011]), restricting possible matches to those with a Lowest Common Ancestor (LCA) minimum score of 100. Reports were limited to the best 5% matches of sequences to the assigned taxon. MEGAN allows the visualization of taxonomic assignments and the option to inspect each match before an assignment is made. OTUs were assigned only where the reference sequence covered the whole of the OTU query sequence. The taxa identified by each assay were combined into one table to allow for comparisons between assays and further analysis. However, where taxa were identified more than once in a sample (by OTUs from more than one assay), these particular taxa were included in the table only once. Some of the lineages for identified taxa were erroneous, this was particularly evident where the same species appeared twice with differing lineages. It was decided to be consistent with assignments and use WoRMS (world register of marine species [Horton et al., 2018]) as the authority for this. Contaminant

(non-marine) OTUs were removed prior to analysis. Taxa were reported to the most informative taxonomic level using the following match parameters: species (97 < 100%); genus (95 < 97%); family (93 < 95%); order (91 < 93%); and class (89 < 91%). Where a sequence matches more than one taxon equally, the assignment was dropped to the lowest common taxonomic level.

## 2.5 | Statistical analysis of OTUs

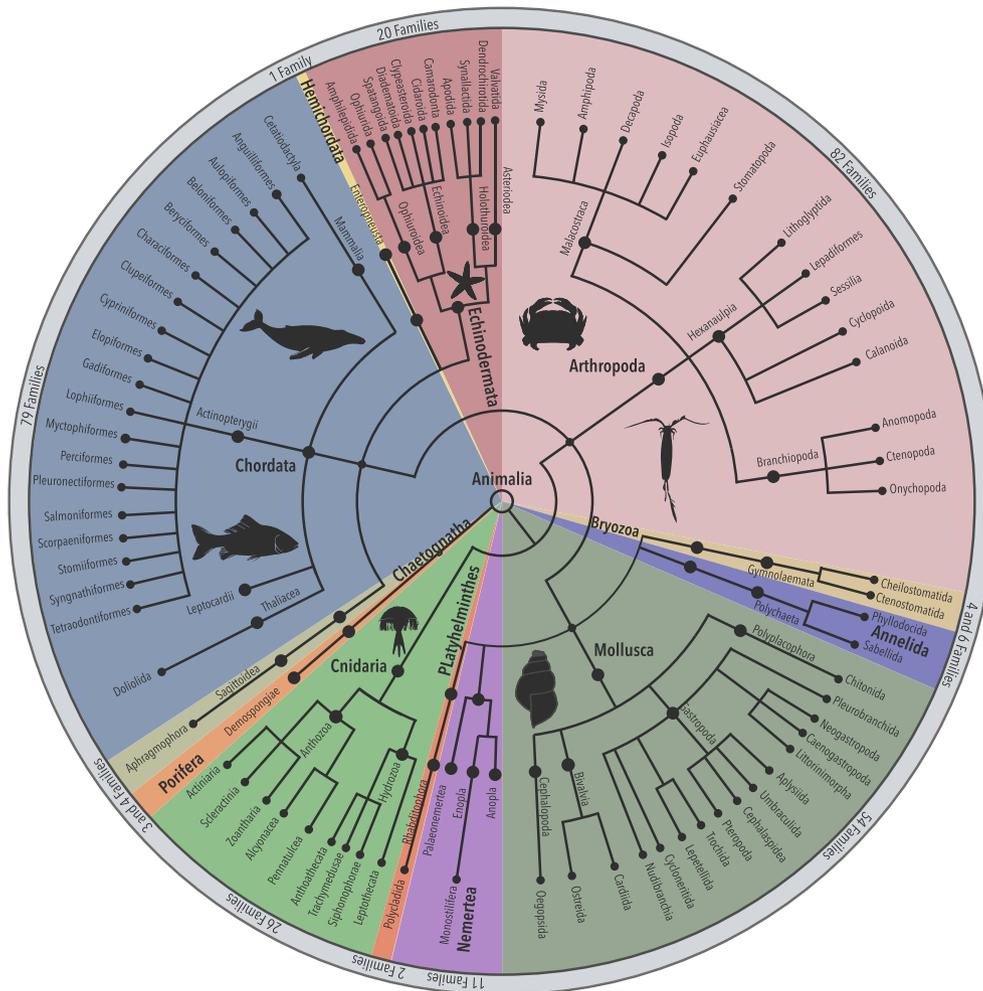
All statistical analyses were performed on the presence/absence of OTU data for sequences obtained for each assay on all 90 samples. For each sample, two measures were calculated: richness, a count of the number of OTUs; and assemblage, the presence/absence composition of the OTUs in the sample. Statistical analyses were performed using PERMANOVA+ (Anderson et al., 2008) add-on for Primer 7 (Clarke & Gorley, 2015) and R packages labdsv (Roberts, 2016) and vegan (Oksanen et al., 2016).

## 2.6 | Univariate analyses

To explore the univariate relationships between the OTUs from each assay and abiotic variables, we used richness as the response variable in generalized linear models (GLMs). We used a negative binomial error structure using the function `glm.nb` in the MASS package in R (Venables & Ripley, 2002) with a log-link function (Zuur et al., 2009). Abiotic variables in the initial model included site, sea surface temperature (SST), salinity, silicate, nitrate and phosphate. Model residuals were visually inspected for homogeneity of variance and normality. The resulting model was checked with an analysis of variance (ANOVA) and visualized with R package `visreg` (Breheny & Burchett, 2017).

## 2.7 | Multivariate analyses

We used PERMANOVA (Anderson, 2001) to investigate site, annual and seasonal effects on richness and assemblage measures. To



**FIGURE 2** Total Animalia detected in DNA extracted from 90 bulk plankton samples using multi-assay metabarcoding. Each pie slice represents, as a proportion, the number of families detected within an individual phylum (the inner nodes). The terminal nodes represent the orders found and the mid-inner nodes the classes. The rim provides an estimate of the number of families found within each phylum.

illustrate site-based patterns, we used non-metric multi-dimensional scaling (nMDS) in the R package *vegan*. To further test the site-based patterns found, we applied a cross-validation test (leave-one-out allocation of observations to groups [canonical analysis of principal coordinates (CAP): Anderson et al., 2008]) among sites to the OTU assemblages from each assay. The indicator species that were characteristic (i.e., present in the majority of the samples from one site but mostly absent in the others) of each site were identified using the function *indval* (indicator value analysis) in the R package *labdsv* (Roberts, 2016)—this analysis was applied to the presence/absence of taxa identified. We also used PERMANOVA (Anderson, 2001) to explore seasonal changes in the assemblage of OTUs. This analysis used a repeated-measures design with factor season (four seasons) nested within site (nine sites) with sampling years as replicates.

## 2.8 | Power to detect seasonal change

Power analyses for richness (R; R-Core-Team, 2015) and serial PERMANOVA analysis for assemblage were conducted on data obtained from Berry et al. (2019) to estimate the number of samples that might be required to detect seasonality. These analyses were done for the same assays as used in this study. The power analysis calculated the number of samples required to achieve a given power given the known effect sizes (F statistics) from Berry et al. (2019). As it is not possible to conduct a power analysis on multivariate data (Anderson Marti et al., 2014), serial PERMANOVA analyses were done instead. The serial analysis looked at the number of samples required to detect significant changes between each season. Increasing numbers of samples from Berry et al. (2019) were chosen randomly from each season of each year to provide input for the analysis, and the levels of statistical significance of the results were compared.

Multivariate analyses were performed using distance-based linear models (DistLM) in PERMANOVA+ using the same abiotic explanatory variables as in the univariate analysis. Bray–Curtis similarity matrices were created from the presence/absence of OTU data. The model that best explained the variation in the OTU assemblage for each assay was selected by the adjusted  $R^2$ .

To provide an indication of the average contribution of a sample to each site and each site to the OTUs produced by an assay, the number of OTUs found in each sample was converted to a percentage of the total number of OTUs found at each site, and the number of OTUs found at each site was converted to a percentage of the total number of OTUs created using each assay.

## 3 | RESULTS

### 3.1 | Overall biodiversity in Australian zooplankton communities

The six metabarcoding assays applied to 90 samples generated >25 million sequences and these produced >2000 individual OTUs. Of these, ~64% could be assigned to an appropriate taxonomic level

(Table S2), and over half of those assigned were to a genus or species level (Tables S3–S9). The average number of sequences per sample varied between assays. The 16S Universal assay produced an average of 3200 filtered sequences per sample which, across all samples, never the less produced 370 OTUs. In contrast, the Fish Assay produced an average of close to 50,000 filtered sequences per sample which formed only 202 OTUs across all samples (Table S2). However, the exploratory sequence number and OTU accumulation curves tested for each sample and assay all reached asymptote.

### 3.2 | Families present

In total, >1000 OTUs were assigned to >500 distinct taxa in Kingdom Animalia including almost 300 unique families (Figure 2). In Arthropoda alone, we identified over 60 families, primarily belonging to Hexanauplia and Malacostraca. In Chordata, we identified over 70 families, with 90% belonging to Actinopterygii. Over 40 Mollusca families were identified; of these, >80% were Gastropoda. In Cnidaria, two-thirds of the 17 families identified belong to Hydrozoa. In Echinodermata, 10 families were identified.

### 3.3 | Species identified

Using our multi-assay metabarcoding approach, we detected a broad range of taxa, with almost 250 assigned at the species level. While many assigned species (e.g., *Parvocalanus crassirostris*; a copepod detected in Darwin, and *Xiphonectes tenuipes*; a crab detected in Yongala) were found only at one site, many species were common to almost all sites, such as *Clausocalanus furcatus* (a copepod, detected at all sites except Darwin) and *Lensia subtiloides* (a Siphonophorae Hydrozoan, detected at all sites except Maria and Kangaroo Islands). Other species detections were restricted to a region. For example, the predatory sea snail *Epitonium replicatum* was found north of the tropic of Capricorn, and the genus of anchovy *Engraulis* was found only to the south of this latitude.

The indicator species analysis (using *indval*, Table S10) showed that the species most characteristic of particular IMOS stations belonged to many phyla, but Arthropoda and Mollusca featured strongly. For example, the top indicator species for Ningaloo was *Metalpheus paragracilis*, a small shrimp-like decapod; for Maria Island, it was *Maoricolpus roseus*, an invasive mollusc (see Table 4); for Port Hacking, it was the offshore copepod *Haloptilus longicornis*, suggesting an oceanic influence in the area; and for Kangaroo Island, it was *Membranipora membranacea*, a bryozoan and another invasive species (although this assignment may be contentious; see Table 4).

### 3.4 | Contamination

Overall, some non-marine contamination was noted in some of the samples (Tables S3 and S4): terrestrial insect (21 samples); house gecko (2 samples); dog (4 samples); gull (2 samples); and human

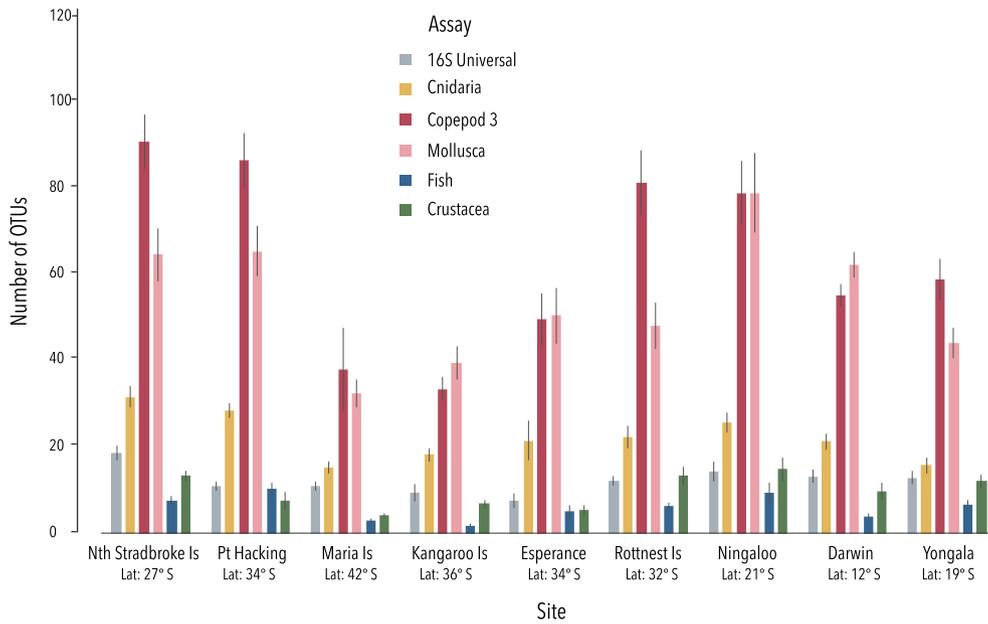


FIGURE 3 Differences in richness (number of OTUs) showing the mean number of OTUs per metabarcoding assay at each site.

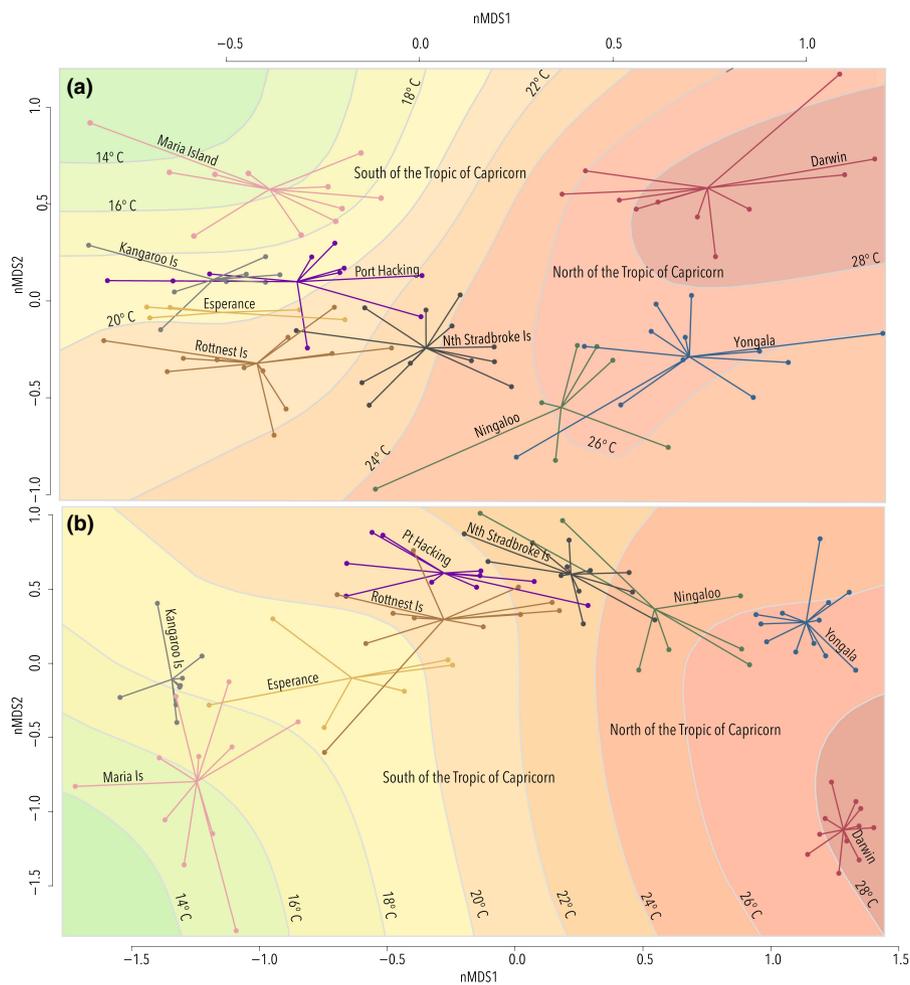


FIGURE 4 Spatial patterns in the planktonic assemblages detected by DNA analysis from bulk plankton samples. Non-metric multi-dimensional scaling plots from the 16S Universal (a) and Mollusca (b) assays showing delineation of OTU assemblages between both the sampling sites and the areas north and south of the Tropic of Capricorn (stress=0.11 & 0.10, respectively, both  $k=3$  &  $p \leq .001$ ). The relationship of the samples to mean sea surface temperature at the time of sampling is indicated by the temperature gradient. Clusters of the same colour represent samples from each site and the junction of the coloured lines indicates the centroids for each site.

TABLE 1 The best-performing models showing the variance in the zooplankton community explained by abiotic factors.

Assay used	OTU diversity test	Variable	SST	Salinity	Silicate	Nitrate	Phosphate	Site	Deviance explained
16S Universal	Assemblage	BM Sequential Probability	0.005	0.012	0.427	0.005	-	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.054	0.043	0.044	0.027	0.033	0.302	R <sup>2</sup>
		BM Probability	0.01	-	-	-	-	0.002	R <sup>2</sup>
Cnidaria	Assemblage	BM Sequential Probability	0.001	0.097	-	0.017	0.33	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.083	0.078	0.081	0.034	0.048	0.394	R <sup>2</sup>
		BM Probability	-	-	-	-	-	<0.001	R <sup>2</sup>
Copepod 3	Assemblage	BM Sequential Probability	0.002	0.037	0.012	0.029	-	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.13	0.094	0.099	0.051	0.065	0.535	R <sup>2</sup>
		BM Probability	-	-	0.049	-	-	<0.001	R <sup>2</sup>
Crustacea	Assemblage	Best Model Sequential Probability	0.02	0.173	-	0.11	-	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.052	0.042	0.045	0.023	0.028	0.292	R <sup>2</sup>
		BM Probability	-	-	-	-	-	<0.001	R <sup>2</sup>
Fish	Assemblage	BM Sequential Probability	-	-	-	0.259	0.186	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.034	0.033	0.03	0.024	0.027	0.197	R <sup>2</sup>
		BM Probability	0.020	-	-	-	-	<0.001	R <sup>2</sup>
Mollusca	Assemblage	BM Sequential Probability	0.002	0.136	0.222	0.37	-	0.001	Adj R <sup>2</sup>
		Marginal Proportion	0.113	0.085	0.093	0.045	0.053	0.46	R <sup>2</sup>
		BM Probability	-	-	-	-	0.024	<0.001	R <sup>2</sup>

Note: For the taxa detected by each assay and based on OTUs, DISTLM was used to relate the zooplankton assemblage (Anderson et al., 2008) to environmental variables, and nbGLM to relate the zooplankton richness (Venables & Ripley, 2002) to the same environmental variables. Only those factors included in the best model have probability values.

(21 samples). The source of these could be from the samples (i.e., terrestrial input from the sea surface), contamination of gear or sample handling. Regardless of source, identified non-marine OTUs were removed from any downstream analysis. Controls that amplified during the PCR stage were sequenced but contained no sequences.

### 3.5 | Spatial variation in zooplankton richness

Overall, the nested PERMANOVA showed significant differences among sites in the number of individual OTUs present in each sample (richness—univariate measurement) for all assays. However, the pairwise PERMANOVA analysis illustrated that not all sites differed, and predominantly the differences in richness were between the southern coldest sites of Kangaroo and Maria Islands and the more northern and warmer sites (Figure 3 and Table S11).

### 3.6 | Spatial variation in zooplankton assemblages

Every metabarcoding assay showed significant differences among sites in the overall analysis of the OTU composition (assemblage—multivariate measurement; PERMANOVA). The pairwise analysis of assemblage within each sample also showed many significant pairwise differences among all sites (PERMANOVA;  $p \leq .001$ —Table S11) and across almost all assays. The exception to this was the Fish Assay. While Fish assemblages from Ningaloo, Darwin, Yongala and

North Stradbroke Island were significantly different from all other sites, assemblages from Esperance and those of the southern-most sites of Port Hacking, Maria, Kangaroo and Rottneest Islands were not significantly different from each other.

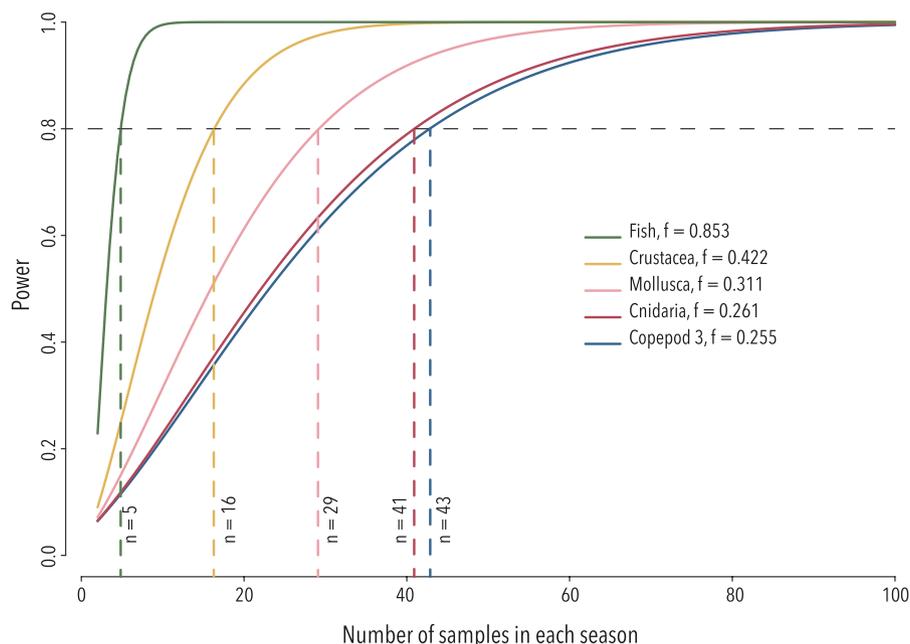
The strong site differences found within the assemblage were exemplified by two of the metabarcoding assays; 16S Universal (Figure 4a) and Mollusca (Figure 4b). These figures illustrate the delineation between assemblages from north and south of the Tropic of Capricorn and the clear relationship of sites with sea surface temperature. An illustration of the assemblage separation between sites for the other assays may be found in Figure S1.

The cross-validation test supported the assemblage differences between sites for each assay. Here, the OTU data from a single sample could consistently be allocated—using composition alone—to the correct site 80%–90% of the time. The exception was the Fish assay, which was allocated to the correct site 56% of the time (Table S12).

### 3.7 | Seasonal variation in zooplankton richness and assemblage

Seasonal changes in richness were observed in data from all assays, except the Crust and 16S universal assays (PERMANOVA). However, there was little significant seasonal change in richness within each site (Table S13).

While all assays (except Fish) showed overall significant differences in assemblage for seasonality, the analysis of seasonality within each site showed little difference. There were some



**FIGURE 5** Power curves from each assay indicating the number of samples required to detect seasonal changes in richness. Thirteen samples per season sampled from Rottneest Island (Berry et al., 2019) were used to create the curves (created in R - R-Core-Team, 2015), and effect sizes  $F$  calculated from PERMANOVA results. Power levels of 0.8 are indicated with the corresponding number of samples per season required to achieve this.

significant seasonal differences, particularly for the taxa detected by the Copepod 3, Mollusca and Cnidaria assays, but the remainder of the assays showed little to no significant response (Table S13).

### 3.8 | Drivers of zooplankton communities

#### 3.8.1 | Richness: GLM

The general linear modelling (nbGLM) revealed that site was most strongly related to biotic richness across all assays. Sea surface temperature (SST) had a significant impact on the 16S Universal and Fish assays, and for Copepod 3 and Mollusca, silicate and phosphate (respectively) were significant drivers of richness

(Table 1 and Figure S2). The variation explained by the best model varied from 26% for the 16S Universal assay—detecting meroplankton—to 69% for the Copepod 3 assay—detecting primarily holoplankton.

#### 3.8.2 | Assemblages: DISTLM

Distance-based linear modelling showed that site and nitrate were the most important predictors of assemblage (Table 1). SST and salinity were included in the models for all assays with the exception of Fish, and silicate was important for the 16S Universal, Copepod 3 and Mollusca assays. The concentration of phosphate was absent from the most parsimonious model for all assays except the Fish and

TABLE 2 The effect of increasing the sample size on the ability of each assay to detect seasonal differences in the assemblage at Rottneest Island.

Assay	Overall significance	Summer spring	Summer winter	Summer autumn	Spring winter	Spring autumn	Winter autumn
<i>One sample per season over 3 years—Three samples per season</i>							
Cnidaria	***	–	–	–	–	–	–
Copepod 3	*	–	–	–	–	–	–
Crustacea	–	–	–	–	–	–	–
Fish	–	–	–	–	–	–	–
Mollusca	**	–	–	–	–	–	–
<i>One sample per season over 5 years—Five samples per season</i>							
Cnidaria	*	*	**	–	–	–	–
Copepod 3	**	–	** HB	–	** HB	–	–
Crustacea	***	** HB	** HB	–	*	–	–
Fish	**	–	** HB	–	*	*	–
Mollusca	***	–	** HB	–	** HB	–	*
<i>Two samples per season over 5 years—Ten samples per season</i>							
Cnidaria	***	** HB	*** HB	** HB	*	** HB	–
Copepod 3	***	** HB	*** HB	** HB	*** HB	** HB	–
Crustacea	***	** HB	*** HB	** HB	** HB	** HB	* HB
Fish	***	–	*** HB	*** HB	*** HB	*** HB	** HB
Mollusca	***	** HB	*** HB	** HB	*** HB	** HB	** HB
<i>Three samples per season over 5 years—Fifteen samples per season</i>							
Cnidaria	***	* HB	*** HB	** HB	*** HB	*** HB	** HB
Copepod 3	***	*	*** HB	* HB	*** HB	*** HB	–
Crustacea	***	** HB	*** HB	* HB	** HB	*** HB	* HB
Fish	***	** HB	*** HB	** HB	*** HB	*** HB	*** HB
Mollusca	***	* HB	*** HB	** HB	*** HB	*** HB	*** HB
<i>Three samples per season over 3 years—Nine samples per season</i>							
Cnidaria	***	* HB	*** HB	** HB	** HB	** HB	–
Copepod 3	***	* HB	*** HB	** HB	*** HB	*** HB	** HB
Crustacea	***	–	** HB	–	** HB	*	–
Fish	***	** HB	*** HB	** HB	* HB	** HB	* HB
Mollusca	***	** HB	*** HB	*** HB	** HB	*** HB	** HB

Note: \*\*\* ( $p \leq .001$ ), \*\* ( $p \leq .01$ ), \* ( $p \leq .05$ ) and <sup>HB</sup> Holm–Bonferroni correction.

TABLE 3 The total OTU diversity detected by each assay for each site and each sample within the site.

Assay	16S Universal (365 OTUs)		Cnidaria (285 OTUs)		Copepod 3 (604 OTUs)		Crustacea (322 OTUs)
	Site contribution to total OTUs (%)	Sample contribution to site OTUs (%)	Site contribution to total OTUs (%)	Sample contribution to site OTUs (%)	Site contribution to total OTUs (%)	Sample contribution to site OTUs (%)	Site contribution to total OTUs (%)
Ningaloo (7)	17	23	26	34	40	33	20
Darwin (12)	23	16	28	27	23	39	22
Yongala (12)	20	18	22	25	27	36	19
Nth Stradbroke Is (12)	26	19	29	38	39	39	26
Port Hacking (10)	15	21	27	37	34	42	12
Maria Is (11)	14	21	18	30	32	20	6
Kangaroo Is (8)	8	31	19	34	12	45	7
Esperance (6)	8	26	21	35	21	39	5
Rottneest Is (12)	15	23	29	27	33	41	17
Range	8–26	16–31	18–29	25–35	12–40	33–45	6–26
Average total %	16	22	24	32	29	37	15

Cnidarian assays. The best models used combinations of the measured abiotic variables to explain between 10% and 49% of the variation in the measured assemblages; 10% for the Fish assay (detecting only fish) and 49% for the Copepod 3 assay (detecting primarily holoplankton).

### 3.9 | Power analysis of the number of samples needed to detect spatial differences in richness

The power analysis based on data from Berry et al. (2019) taken over 5 years showed that for most assays, the effect size for seasonal changes was very low, so substantially more sampling than was conducted in this study would be required to detect seasonal changes in richness; the exception to this was the Fish assay. As a result of larger seasonal effects in fish numbers, sufficient power could be achieved to detect seasonal change using the Fish assay with just five samples from each season (Figure 5). By contrast, Crustacea required 16 samples per season and the remaining assays ranged from 29 to 43 samples per season.

### 3.10 | Assemblage analysis of the number of samples needed to detect spatial differences

Serial analysis of the Rottneest Island Assemblage (Berry et al., 2019) showed an increase in significant pairwise seasonal differences with the number of samples used (PERMANOVA). While five samples per season provided some significant results (second row; Table 2), we estimate a minimum of nine samples, from each season, are required to track seasonal changes effectively (final row; Table 2).

### 3.11 | Contribution of each sample to OTU diversity

The contribution of each sample to the total OTU diversity detected by each assay at each site is shown in Table 3, as well as the site contribution to the total OTU diversity detected by the assay. Copepod 3 produced the most reproducible results from sample to sample, with an average sample contribution to each site from 33% to 45%. The least reproducible was the Fish assay, where the average sample contribution ranged from 13% to 27%.

On average, each sample contained between 24% and 34% of the OTUs detected at each site depending on the assay. However, the contribution of each site to the total number of OTUs detected by an assay was between 12% and 29%.

## 4 | DISCUSSION

### 4.1 | Biodiversity of Australian zooplankton communities

A high diversity of taxa were detected within this temporal and spatial study. Given the size of the dataset, it was not practical to check the biological credibility of each assignment, an issue with large datasets that have been recognized previously (Deagle et al., 2017). However, an examination using the Atlas of Living Australia (ALA, 2017) revealed that ~76% of the assignments were detected within their expected habitats, ~8% were detected within both known and unexpected areas and ~90% of all assignments were known to inhabit Australian waters, whether or not they were detected in their usual locale. Despite this, there were many animals we did not detect. Ctenophores, salps, larvaceans and ostracods should all be present

Sample contribution to site OTUs (%)	Fish (202 OTUs)		Mollusca (553 OTUs)		Total	
	Site contribution to total OTUs (%)	Sample contribution to site OTUs (%)	Site contribution to total OTUs (%)	Sample contribution to site OTUs (%)	Mean site contribution to total OTUs (%)	Mean sample contribution to site OTUs (%)
23	24	19	41	35	28	28
14	16	13	32	35	24	24
19	24	14	22	36	22	25
15	25	15	30	39	29	28
21	33	15	30	39	25	29
24	7	20	21	28	16	24
31	3	27	20	36	12	34
33	12	21	25	37	15	32
25	15	21	27	32	23	28
14–33	3–33	13–27	20–41	28–39	12–29	24–34
23	18	18	28	35	22	28

within a zooplankton community but were not detected by the range of assays deployed. Given that ~36% of all OTUs were not assigned to the taxonomic framework, it is likely that many of these OTUs are representative of sequences that are not yet included in reference databases or are entered with errors that might confuse their taxonomic assignment. A recent search of GenBank (accessed Dec 2021) for Salpidae and Appendicularia COI revealed sequences for only five and three species respectively. The gaps in the reference databases with regard to plankton have been recognized previously; in 2016, Chain et al. identified only 61% of plankton reads from an 18S assay and more recently Wangenstein et al. (2018) were unable to identify 36% of their marine COI MOTUs. However, we identified taxa that are not routinely identified morphologically in plankton samples.

One of the more curious taxa was humpback whale (*Megaptera novaeangliae*). The whale was found in six samples and four sites by two different assays on both the east and west coasts of Australia. While this baleen whale clearly does not have a planktonic stage to its lifecycle, it does feed on krill, small fish and plankton (Johnson & Wolman, 1984), and the DNA detected could originate either from particulate matter such as faeces or shed cells, or from secondary predation where detritivores, such as krill, feed upon whale waste. *M. novaeangliae* has a known near-coastal migration pattern that passes near the sampling sites where it was found and all detections occurred during the whales' regular spring migration time, and there were no detections outside of these times. We noted that one of the sites was at Rottneest Island. In a previous study, Berry et al. (2019) used the same Rottneest Island samples and the Copepod 3 assay but did not report this taxon. We re-examined the unfiltered Copepod 3 sequences from the Rottneest samples amplified in both studies to find that the whale sequences were present in Berry et al. (2019), but at an abundance too low to pass the (stringent) filtering process. Finding whale DNA in 6 of the 90 samples, at time when this taxon is known to be in the area, provides evidence for the use of plankton samples as an option to detect species much larger than zooplankton in the water column.

Another interesting result was the three-spot swimming crab (*Portunus sanguinolentus*). This commercially harvested sand crab was the top indicator species at North Stradbroke Island. This finding may be a reflection of the crab fisheries in nearby Moreton Bay. However, while ALA (2017) indicates that the crab's range is across northern half of Australia, North Stradbroke was the only site we detected it. A number of other taxa were found that are of economic, conservational and ecological interest (see Table 4 for a selection).

## 4.2 | Diversity of zooplankton OTUs across Australia

### 4.2.1 | Richness

While there were few significant differences in richness among sites, there were some noted among the colder sites of Maria Island and Kangaroo Island, and the more northern sites. On average, Darwin was 15°C warmer than Maria Island. Given that some previous studies have identified SST as a main driver of species richness (Benedetti et al., 2021), it was expected that the differences noted in richness were reflective of this disparity in mean sea surface temperature. However, our analysis of the relationship between richness and abiotic factors found that site was the most influential abiotic factor (Table 1). While this may be the result of correlation between site and SST, the finding is in accordance with Chaudhary et al. (2021). The authors in that study undertook the task of reviewing research data as far back as 1955, from over 48,000 marine animal species to determine (among other findings) that, while SST is an important variable, latitude explains more variation in species richness. Here, SST was significant only for the Fish and 16S Universal assays; both of which detect predominately meroplankton.

In this study, we used a maximum of 12 samples taken across 3 years. These resulted in finding little or no seasonal changes in

richness across this time. However, samples from a previous study (Berry et al., 2019), which used 12 samples per year from one site for 5 years, were used to determine how much sampling would be required to detect seasonality in richness for each assay (Figure 5). As expected, the most responsive assay was the Fish assay (5 samples from each season), this was followed by the Crustacea assay (16 samples), both of which detect predominately meroplankton. This indicates that meroplankton are more likely to show seasonal changes in OTU numbers, as they quite often emerge in summer and only spend a portion of their life within the zooplankton community, usually as juveniles or eggs. While we acknowledge that seasonal patterns will vary in intensity depending on the sites, a similar pattern was noted off the coast of Plymouth by Highfield et al. (2010), where a morphological study examined weekly samples from 1988 to 2007 to show that the occurrence of meroplankton peaked in spring and then again in late summer.

#### 4.2.2 | Assemblage

Given the climatic and spatial distinctions between the sites, it was unsurprising that we found highly significant differences in OTU assemblages between all sites and across all assays (except the Fish assay). The spatial differences noted are consistent with findings from previous zooplankton studies (e.g., Chain et al., 2016). The incorporation of the abiotic factors showed both site and SST to be important drivers for the assemblage of OTUs. However, it was the assays detecting primarily holoplankton (Copepod 3, Mollusca and Cnidaria) that produced the models most responsive to the abiotic factors. This pattern suggests that obligate zooplankton are most affected by environmental factors, at least in the short term. While the assemblages from the Fish assay did not differ between all sites, the pairwise analysis showed significant differences between the Southern and Northern sites. This is a likely consequence of fishes being larger and more mobile than species detected by the other five assays.

There were few significant seasonal differences noted in OTU assemblage across all assays. Our study used a maximum of three samples per season across 3 years (one sample per season per year). Post hoc analysis demonstrated that this experimental design was not adequate sampling to show seasonal differences for any of the metabarcoding assays (Table 2). Once again, we examined extended data from the Rottneest Island site (Berry et al., 2019) to determine how much sampling might be needed to detect seasonal changes in assemblage using each assay. In that study, assemblage differences between summer and winter in all assays could be reliably detected using one sample per season across 5 years. However, our study indicates that for the detection of significant differences between all seasons, 9 to 12 samples per season are required. These findings are supported by a recent study that used 48 filtered water samples (12 samples per season) taken over a 3-year period from a single site in the Gulf of Naples. In that study, Di Capua et al. (2021)

used a V4-18S RNA fragment to demonstrate clear seasonality in the community composition of the zooplankton sampled. Chain et al. (2016) also found significant seasonal changes in their data. This study used six independent transects from each port over two seasons (May and December). This result is analogous to our data (Table 2), which shows that five samples per season will produce significant differences between summer and winter. Our data add to the growing evidence and provide important recommendations as to how future environmentally based DNA surveys might be designed and implemented into long-term monitoring programmes.

While the DISTLM results (Table 1), particularly with respect to site and SST, are not unexpected, there are limited studies linking abiotic factors to a wide variety of zooplankton (Mackas et al., 2012) and fewer still that employ a multi-assay metabarcoding approach. Almost all studies that have related zooplankton to environmental conditions use taxonomy based on microscopy (e.g., Everaert et al., 2018; Labuce et al., 2020; Shi et al., 2020). Here, we confirm that DNA data from bulk zooplankton samples can also be used as a measure of the zooplankton community.

#### 4.3 | Sampling design

This Australia-wide audit has implications for the design of future large-volume eDNA studies, whether they be spatial or temporal in nature. The question is how much sampling is necessary to provide an indication of zooplankton diversity at a particular location. The answer is not straightforward and varies depending on whether the aim is to examine assemblage or richness or both (Figure 5 and Table 2). While it might be easy to advocate for many samples and replicates, the costs and logistics of sampling at sea make a more considered approach necessary. Batten et al. (2019) advocate the globally coordinated use of continuous plankton recorders to monitor plankton populations and provide data for the sustainable management of our oceans. Our position is that an environmental metagenomic approach using bulk plankton samples could provide much of this information without the need for taxonomic expertise. Further, a valuable synergy would be to use DNA analysis of continuous plankton recorder samples, which has been shown to be an efficient, cost-efficient approach to identifying biodiversity in the epipelagic zone (Deagle et al., 2017).

The examination of the proportion of the contribution of individual samples and sites provides a glimpse of the percentage of OTUs in a single sample that contributes to the OTU diversity found within a site (Table 3). While these results could be partly attributed to the disparity in sample number within the sites, Ningaloo had only 7 samples and was the second largest contributor of OTU diversity (28%), and Maria Island had 11 samples and was the third smallest contributor (16%). There is also an overlap of OTUs between samples, preventing simple overall conclusions. The difference between the sample contributions of the Copepod 3 assay (33%–45%) and the Fish assay (13%–27%) also supports

TABLE 4 A selection of commercially and environmentally important taxa detected within the plankton samples.

Taxon	Site	Detections	Assays	Importance	Notes
<i>Maoricolpus roseus</i> (New Zealand screwshell)	Maria Is	8 of 11	Copepod 3 Mollusca	Invasive	This was the top indicator species for Maria Island. It lives in high densities and the accumulation of living and dead shells makes the habitat incompatible with the survival of native species (Probst & Crawford, 2008). It was introduced to Tasmania in the 1920s (Probst & Crawford, 2008) but is now found on the east coast of Australia (ALA, 2017). The shell has a planktonic larval stage, providing a method for dispersal
<i>Membranipora membranacea</i> (Lacy-crust bryozoan)	Maria Is Port Hacking Kangaroo Is Esperance	5 of 11 1 of 10 8 of 8 2 of 6	Copepod 3 Mollusca Crustacea	Invasive	This was the top indicator species for Kangaroo Island. It is a colonial bryozoan that grows on the surface of kelp blades causing them to become brittle and break (Saunders & Metaxas, 2008). It causes defoliation and reduces survival of the host kelp, which then allows the infiltration of other opportunistic species (Saunders & Metaxas, 2008). While the sequences were robust matches to the reference sequences across two genes, many bryozoan species are not found on the GenBank database (Benson et al. (2014)—although ALA (2017) currently provides no alternatives at that site)
<i>Megaptera novaeangliae</i> (Humpback whale)	Rottneest Is Ningaloo Port Hacking Nth Stradbroke Is	2 of 12 2 of 7 1 of 10 1 of 12	Copepod 3 16S Universal	Tourism	This whale is found all around Australia (ALA, 2017). It feeds on krill, small fish and zooplankton (Johnson & Wolman, 1984). The DNA detected is likely present in the samples from shed skin cells, saliva, urine and/or faeces. The detection times correlate with known migration patterns
<i>Pegasus volitans</i> (Slender sea moth)	Darwin	1 of 12	Fish Copepod 3	Potentially threatened	An IUCN report on this species states there is insufficient data to determine its conservation status (Pollom, 2016). It is commonly caught as by-catch from trawl fishing and sold for use in traditional Chinese medicine to replace the increasingly costly and rarer seahorse, and is thought to be susceptible to overexploitation and has a planktonic stage in its lifecycle (Vincent Amanda, 2003)
<i>Portunus sanguinolentus</i> (three-spot swimming crab)	Nth Stradbroke Is	5 of 12	Crustacea 16S Universal	Commercial	The top indicator species for North Stradbroke Island. This is a commercially harvested edible crab that is found across the Indo-Pacific region (Sumpton et al., 1989)

the results of the power analysis (Figure 5), which showed fish needed the least number of samples and Copepod 3 the most to exhibit seasonality in the richness data.

Generally, the southern, colder, sites (Maria Island, Kangaroo Island and Esperance) are those that had the fewest OTUs when using assays that favour the larger meroplankton—Fish, Crustacea and 16S Universal. Although this might be an effect of bottom depth, it might also be that the prevalence of meroplankton species could increase in more tropical waters, although this is a yet-to-be-answered question. However, the warmer sites (especially Darwin) and those situated within Australia's large coastal

currents—the East Australian (Yongala, Nth Stradbroke Is and Port Hacking) and Leeuwin (in the West: Ningaloo and Rottneest Is)—showed more OTU diversity both per sample and contributing to Australia-wide diversity across all assays, as would be expected in warmer regions.

Given the results provided by this, and the post hoc analysis using the extended Rottneest data (Figure 5 and Table 2), it is evident that continuous temporal and spatial sampling is needed to produce a thorough understanding of zooplankton communities. The variation in community composition over time at each of these sites demonstrates the limits of single-point sampling.

## 5 | CONCLUSIONS

Long-term biomonitoring of marine environments is essential to the understanding of the biotic networks they support. It is well recognized that an ecosystem-wide approach to this issue is needed, however, to attempt this using traditional morphological approach is not realistic—multi-assay DNA metabarcoding of bulk plankton samples makes this sizeable task more attainable.

The data attained here highlight the breadth of knowledge that can be achieved from extended spatial and temporal studies and illustrate the ability of multi-gene metabarcoding to reveal clear spatial differences among highly complex zooplankton communities. While the sampling regime was not conducive to detecting significant seasonal changes, the combined data from Rottneest Island (Berry et al., 2019) provide a benchmark for experimental design for future DNA-based monitoring programmes using plankton as a substrate.

This study provides insight into the importance of collecting and maintaining accessible data and sample archives; both biotic and abiotic data. Time-stamped data on biotic composition provide a unique resource to reveal community composition patterns and allow changes in them to be measured in light of associated abiotic factors. Broad-scale spatial and temporal studies have become critical to provide much needed evidence for the management and mitigation of the anthropogenic impacts experienced by our oceans. Further, the increasing speed (automation) and portability of DNA sequencers make the incorporation of metagenomic methods into the marine monitoring toolkit a matter of when, not if.

### ACKNOWLEDGEMENTS

None.

### FUNDING INFORMATION

I thank the Integrated Marine Observing System (IMOS; supported by the Australian Government) for providing access to the samples and abiotic data used during this study. The resources provided by the Pawsey Supercomputing Centre, which is funded by the Governments of Australia and Western Australia, also supported this work. M.B., E.S.H. and M.S. acknowledge the support of ARC Linkage projects (LP160100839 and LP160101508) to explore marine metabarcoding applications.

### CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

### DATA AVAILABILITY STATEMENT

The quality filtered sequence data are available on Data Dryad (<https://doi.org/10.5061/dryad.931zcrjns>).

### ORCID

Tina E. Berry  <https://orcid.org/0000-0002-7203-2437>

Benjamin J. Saunders  <https://orcid.org/0000-0003-1929-518X>

Anthony J. Richardson  <https://orcid.org/0000-0002-9289-7366>

Matthew Power  <https://orcid.org/0000-0002-9009-8966>

Euan Harvey  <https://orcid.org/0000-0002-9069-4581>

Simon Jarman  <https://orcid.org/0000-0002-0792-9686>

Oliver Berry  <https://orcid.org/0000-0001-7545-5083>

Claire H. Davies  <https://orcid.org/0000-0002-0424-1835>

Michael Bunce  <https://orcid.org/0000-0002-7721-4790>

### REFERENCES

- Alberti, A., Poulain, J., Engelen, S., Labadie, K., Romac, S., Ferrera, I., Albin, G., Aury, J.-M., Belser, C., & Bertrand, A. (2017). Viral to metazoan marine plankton nucleotide sequences from the Tara oceans expedition. *Scientific Data*, 4, 170093.
- Alexander, J. B., Bunce, M., White, N., Wilkinson, S. P., Adam, A. A. S., Berry, T., Stat, M., Thomas, L., Newman, S. J., Dugal, L., & Richards, Z. T. (2019). Development of a multi-assay approach for monitoring coral diversity using eDNA metabarcoding. *Coral Reefs*, 39, 159–171.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215, 403–410.
- Anderson, M., Gorley, R. N., & Clarke, R. K. (2008). *Permanova+ for primer: Guide to software and statistical methods*. Primer-E Limited.
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26, 32–46.
- Anderson Marti, J., Santana-Garcon, J., & Chase, J. (2014). Measures of precision for dissimilarity-based multivariate analysis of ecological communities. *Ecology Letters*, 18, 66–73.
- Andruszkiewicz, E. A., Starks, H. A., Chavez, F. P., Sassoubre, L. M., Block, B. A., & Boehm, A. B. (2017). Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLoS One*, 12, e0176343.
- Atlas of Living Australia website [Internet]. (2017). Retrieved September, 2016, from <http://www.ala.org.au2017>
- Batten, S. D., Abu-Alhaja, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., Kitchener, J. A., Koubbi, P., McQuatters-Gollop, A., Muxagata, E., Ostle, C., Richardson, A. J., Robinson, K. V., Takahashi, K. T., Verheye, H. M., & Wilson, W. (2019). A global plankton diversity monitoring program. *Frontiers in Marine Science*, 6, 321.
- Beaugrand, G., Reid, P. C., Ibanez, F., Lindley, J. A., & Edwards, M. (2002). Reorganization of North Atlantic marine copepod biodiversity and climate. *Science*, 296, 1692–1694.
- Benedetti, F., Vogt, M., Elizondo, U. H., Righetti, D., Zimmermann, N. E., & Gruber, N. (2021). Major restructuring of marine plankton assemblages under global warming. *Nature Communications*, 12, 5226.
- Benson, D. A., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2014). GenBank. *Nucleic Acids Research*, 42, D32–D37.
- Berry, T. E., Osterrieder, S. K., Murray, D. C., Coghlan, M. L., Richardson, A. J., Greal, A. K., Stat, M., Bejder, L., & Bunce, M. (2017). DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (*Neophoca cinerea*). *Ecology and Evolution*, 7, 5435–5453.
- Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., Davies, C. H., Berry, O., Harvey, E. S., & Bunce, M. (2019). Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. *PLoS Genetics*, 15, e1007943.
- Breheeny, P., & Burchett, W. (2017). Visualization of regression models using Visreg. *The R Journal*, 9, 56–71.
- Brown, E. A., Chain, F. J., Crease, T. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution*, 5, 2234–2251.
- Casper, R. M., Jarman, S. N., Deagle, B. E., Gales, N. J., & Hindell, M. A. (2007). Detecting prey from DNA in predator scats: A comparison with morphological analysis, using *Arctocephalus* seals fed a known diet. *Journal of Experimental Marine Biology and Ecology*, 347, 144–154.

- Chain, F. J. J., Brown, E. A., MacIsaac, H. J., Cristescu, M. E., & Cowie, R. (2016). Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Diversity and Distributions*, *22*, 493–504.
- Chaudhary, C., Richardson, A. J., Schoeman, D. S., & Costello, M. J. (2021). Global warming is causing a more pronounced dip in marine species richness around the equator. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(15), e2015094118.
- Clarke, K., & Gorley, R. (2015). *Getting started with PRIMER v7*. PRIMER-E: Plymouth Marine Laboratory.
- 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*, 873–883.
- Deagle, B. E., Clarke, L. J., Kitchener, J. A., Polanowski, A. M., & Davidson, A. T. (2017). Genetic monitoring of open ocean biodiversity: An evaluation of DNA metabarcoding for processing continuous plankton recorder samples. *Molecular Ecology Resources*, *18*(3), 391–406.
- Deagle, B. E., Tollit, D. J., Jarman, S. N., Hindell, M. A., Trites, A. W., & Gales, N. J. (2005). Molecular scatology as a tool to study diet: Analysis of prey DNA in scats from captive Steller Sea lions. *Molecular Ecology*, *14*, 1831–1842.
- Di Capua, I., Piredda, R., Mazzocchi, M. G., Zingone, A., & Fields, D. (2021). Metazoan diversity and seasonality through eDNA metabarcoding at a Mediterranean long-term ecological research site. *ICES Journal of Marine Science*, *78*, 3303–3316.
- Di Marco, M., Venter, O., Possingham, H. P., & Watson, J. E. M. (2018). Changes in human footprint drive changes in species extinction risk. *Nature Communications*, *9*, 4621.
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*, 2460–2461.
- Edwards, M., Beaugrand, G., Hays, G. C., Koslow, J. A., & Richardson, A. J. (2010). Multi-decadal oceanic ecological datasets and their application in marine policy and management. *Trends in Ecology & Evolution*, *25*, 602–610.
- Edwards, M., & Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, *430*, 881–884.
- Eriksen, R. S., Davies, C. H., Bonham, P., Coman, F. E., Edgar, S., McEnulty, F. R., McLeod, D., Miller, M. J., Rochester, W., Slotwinski, A., Tonks, M. L., Uribe-Palomino, J., & Richardson, A. J. (2019). Australia's long-term plankton observations: The integrated marine observing system national reference station network. *Frontiers in Marine Science*, *6*, 161.
- Everaert, G., Deschutter, Y., De Troch, M., Janssen, C. R., & De Schampelaere, K. (2018). Multimodel inference to quantify the relative importance of abiotic factors in the population dynamics of marine zooplankton. *Journal of Marine Systems*, *181*, 91–98.
- Everett, J. D., Baird, M. E., Blanchard, J., Buchanan, P., Bulman, C., Davies, C., Downie, R., Griffiths, C., Heneghan, R., Laiolo, L., Lopez, A. L., Lozano-Montes, H., Matear, R., McEnulty, F., Robson, B., Rochester, W., Skerratt, J., Smith, J. A., Strzelecki, J., ... Richardson, A. J. (2017). Modeling what we sample and sampling what we model: Challenges for zooplankton model assessment. *Frontiers in Marine Science*, *4*, 19.
- Gattuso, J. P., Magnan, A., Billé, R., Cheung, W. W. L., Howes, E. L., Joos, F., Allemand, D., Bopp, L., Cooley, S. R., Eakin, C. M., Hoegh-Guldberg, O., Kelly, R. P., Pörtner, H. O., Rogers, A. D., Baxter, J. M., Laffoley, D., Osborn, D., Rankovic, A., Rochette, J., ... Turley, C. (2015). Contrasting futures for ocean and society from different anthropogenic CO<sub>2</sub> emissions scenarios. *Science*, *349*(6243), aac4722.
- Gimmler, A., Korn, R., de Vargas, C., Audic, S., & Stoeck, T. (2016). The Tara oceans voyage reveals global diversity and distribution patterns of marine planktonic ciliates. *Scientific Reports*, *6*, 33555.
- Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemann, L., Romagnan, J. B., Cawood, A., Pesant, S., Garcia-Comas, C., & Prejger, F. (2010). Digital zooplankton image analysis using the ZooScan integrated system. *Journal of Plankton Research*, *32*(3), 285–303.
- Griffiths, H. J., Meijers, A. J. S., & Bracegirdle, T. J. (2017). More losers than winners in a century of future Southern Ocean seafloor warming. *Nature Climate Change*, *7*, 749–754.
- Guardiola, M., Uriz, M. J., Taberlet, P., Coissac, E., Wangensteen, O. S., & Turon, X. (2015). Deep-sea, deep-sequencing: Metabarcoding extracellular DNA from sediments of marine canyons. *PLoS One*, *10*, e0139633.
- Hardy, N., Berry, T., Kelaher, B. P., Goldsworthy, S. D., Bunce, M., Coleman, M. A., Gillanders, B. M., Connell, S. D., Blewitt, M., & Figueira, W. (2017). Assessing the trophic ecology of top predators across a recolonisation frontier using DNA metabarcoding of diets. *Marine Ecology Progress Series*, *573*, 237–254.
- Herman, A. W., Beanlands, B., & Phillips, E. F. (2004). The next generation of optical plankton counter: The laser-OPC. *Journal of Plankton Research*, *26*(10), 1135–1145.
- Highfield, J. M., Eloire, D., Conway, D. V. P., Lindeque, P. K., Attrill, M. J., & Somerfield, P. J. (2010). Seasonal dynamics of meroplankton assemblages at station L4. *Journal of Plankton Research*, *32*, 681–691.
- Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N., & Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN 4. *Genome Research*, *21*, 1552–1560.
- Jackson, J. B. C., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., Bradbury, R. H., Cooke, R., Erlandson, J., Estes, J. A., Hughes, T. P., Kidwell, S., Lange, C. B., Lenihan, H. S., Pandolfi, J. M., Peterson, C. H., Steneck, R. S., Tegner, M. J., & Warner, R. R. (2001). Historical overfishing and the recent collapse of coastal ecosystems. *Science*, *293*, 629–637.
- Johnson, J. H., & Wolman, A. A. (1984). The humpback whale, *Megaptera novaeangliae*. *Marine Fisheries Review*, *46*, 30–37.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thieler, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649.
- Labuce, A., Dimante-Deimantovica, I., Tunens, J., & Strake, S. (2020). Zooplankton indicator-based assessment in relation to site location and abiotic factors: A case study from the Gulf of Riga. *Environmental Monitoring and Assessment*, *192*, 147.
- Lenanton, R. C. J., Dowling, C. E., Smith, K. A., Fairclough, D. V., & Jackson, G. (2017). Potential influence of a marine heatwave on range extensions of tropical fishes in the eastern Indian Ocean—Invaluable contributions from amateur observers. *Regional Studies in Marine Science*, *13*, 19–31.
- Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., & Atkinson, A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS One*, *8*, e81327.
- Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., Kidwell, S. M., Kirby, M. X., Peterson, C. H., & Jackson, J. B. C. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science*, *312*, 1806–1809.
- Lynch, T. P., Morello, E. B., Evans, K., Richardson, A. J., Rochester, W., Steinberg, C. R., Roughan, M., Thompson, P., Middleton, J. F., Feng, M., Sherrington, R., Brando, V., Tilbrook, B., Ridgway, K., Allen, S., Doherty, P., Hill, K., & Moltmann, T. C. (2014). IMOS National Reference Stations: A continental-wide physical, chemical and biological coastal observing system. *PLoS One*, *9*, e113652.
- Mackas, D. L., Greve, W., Edwards, M., Chiba, S., Tadokoro, K., Eloire, D., Mazzocchi, M. G., Batten, S., Richardson, A. J., Johnson, C., Head, E., Conversi, A., & Peluso, T. (2012). Changing zooplankton seasonality in a changing ocean: Comparing time series of zooplankton phenology. *Progress in Oceanography*, *97–100*, 31–62.
- McCauley, D. J., Pinsky, M. L., Palumbi, S. R., Estes, J. A., Joyce, F. H., & Warner, R. R. (2015). Marine defaunation: Animal loss in the global ocean. *Science*, *347*, 1255641.
- Morand, R., Lejzerowicz, F., Darling, K. F., Lecroq-Bennet, B., Winther Pedersen, M., Orlando, L., Pawlowski, J., Munitza, S., de Vargas, C.,

- & Kucera, M. (2017). Planktonic foraminifera-derived environmental DNA extracted from abyssal sediments preserves patterns of plankton macroecology. *Biogeosciences*, *14*, 2741–2754.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Szoecs, E. (2016). *Vegan: Community ecology package* (R package version 2.3-0). <http://CRAN.R-project.org/package=vegan>
- Pollom, R. (2016). *Pegasus volitans* (errata version published in 2017). The IUCN Red List of Threatened Species 2016: e.T16476A115133968. [Internet] <https://doi.org/10.2305/IUCN.UK.2016-3.RLTS.T16476A46624265.en>
- Probst, T. A., & Crawford, C. M. (2008). Population characteristics and planktonic larval stage of the New Zealand screwshell *Maoricolpus roseus*. *Journal of Molluscan Studies*, *74*, 191–197.
- R: A language and environment for statistical computing. (2015). R Foundation for Statistical Computing, Vienna, Austria [Internet]. <http://www.R-project.org/>
- Richardson, A. (2009). Plankton and climate. In *Elements of Physical Oceanography: A derivative of the Encyclopedia of Ocean Sciences*, 397. Academic Press.
- Richardson, A. J., Brown, C. J., Brander, K., Bruno, J. F., Buckley, L., Burrows, M. T., Duarte, C. M., Halpern, B. S., Hoegh-Guldberg, O., Holding, J., Kappel, C. V., Kiessling, W., Moore, P. J., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Schoeman, D. S., Schwing, F., Sydeman, W. J., & Poloczanska, E. S. (2012). Climate change and marine life. *Biology Letters*, *8*, 907–909.
- Richardson, A. J., Uribe-Palomino, J., Slotwinski, A., Coman, F., Miskiewicz, A. G., & Suthers, I. M. (2019). Coastal and marine zooplankton: Identification and biology. In M. Suthers Iain, D. Rissik, & J. Richardson Anthony (Eds.), *Plankton: A guide to their ecology and monitoring for water quality* (pp. 141–208). CSIRO Publishing.
- Roberts, D. W. (2016). *labdsv: Ordination and multivariate analysis for ecology*. (R package version 1.8-0).
- Rohner, C. A., Armstrong, A. J., Pierce, S. J., Prebble, C., Cagua, E. F., Cochran, J. E. M., Berumen, M. L., & Richardson, A. J. (2015). Whale sharks target dense prey patches of sergestid shrimp off Tanzania. *Journal of Plankton Research*, *37*(2), 1–11.
- Saunders, M., & Metaxas, A. (2008). High recruitment of the introduced bryozoan *Membranipora membranacea* is associated with kelp bed defoliation in Nova Scotia, Canada. *Marine Ecology Progress Series*, *369*, 139–151.
- Schnell, I. A., Bohm, K., & Gilbert, T. P. (2015). Tag jumps illuminated—Reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, *15*, 1289–1303.
- Shi, Y., Niu, M., Zuo, T., Wang, J., Luan, Q., Sun, J., Yuan, W., Shan, X., & Pakhomov, E. A. (2020). Inter-annual and seasonal variations in zooplankton community structure in the Yellow Sea with possible influence of climatic variability. *Progress in Oceanography*, *185*, 102349.
- Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., Harvey, E. S., & Bunce, M. (2017). Ecosystem bio-monitoring with eDNA: Metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, *7*, 12240.
- Sumpton, W. D., Smith, G. S., & Potter, M. A. (1989). Notes on the biology of the Portunid crab, *Portunus sanguinolentus* (Herbst), in subtropical Queensland waters. *Marine and Freshwater Research*, *40*, 711–717.
- Suter, L., Polanowski, A. M., Clarke, L. J., Kitchener, J. A., & Deagle, B. E. (2020). Capturing open ocean biodiversity: Comparing environmental DNA metabarcoding to the continuous plankton recorder. *Molecular Ecology*, *30*(13), 3140–3157.
- Taberlet, P., Brown, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA for biodiversity research and monitoring*. Oxford University Press.
- Venables, W. N., & Ripley, B. D. (2002). Random and mixed effects. In *Modern applied statistics with S* (pp. 271–300). Springer.
- Vincent Amanda, C. J. (2003). Trade in pegasid fishes (sea moths), primarily for traditional Chinese medicine. *Oryx*, *31*, 199–208.
- Wangensteen, O. S., Palacin, C., Guardiola, M., & Turon, X. (2018). DNA metabarcoding of littoral hard-bottom communities: High diversity and database gaps revealed by two molecular markers. *PeerJ*, *6*, e4705.
- Wernberg, T., Bennett, S., Babcock, R. C., de Bettignies, T., Cure, K., Depczynski, M., Dufois, F., Fromont, J., Fulton, C. J., Hovey, R. K., Harvey, E. S., Holmes, T. H., Kendrick, G. A., Radford, B., Santanagarcon, J., Saunders, B. J., Smale, D. A., Thomsen, M. S., Tuckett, C. A., ... Wilson, S. (2016). Climate-driven regime shift of a temperate marine ecosystem. *Science*, *353*, 169–172.
- World Register of Marine Species (WoRMS) [Internet]. (2018). *WoRMS Editorial Board*. [cited 2018-04-11]. Retrieved from <http://www.marinespecies.org>
- Zuur, A., Ieno, E., Walker, N., Saveliev, A., & Smith, G. (2009). *Mixed effects models and extensions in ecology with R*. Springer. 574 p.

## BIOSKETCHES

**Tina E. Berry** has a background in developing molecular tools for the analysis of environmental DNA. While she now works with a wide variety of environmental substrates, Tina has a particular interest in the biodiversity and conservation of both freshwater and marine ecosystems. The remaining authors have diverse backgrounds that converge; with common interests in the biodiversity and conservation of marine ecosystems. **Anthony J. Richardson** and **Claire H. Davies** have a particular interest in zooplankton research. **Benjamin J. Saunders** and **Anthony J. Richardson** have backgrounds in statistics, and **Megan L. Coghlan**, **Matthew Power**, **Simon Jarman**, **Euan Harvey**, **Oliver Berry** and **Michael Bunce** are all interested in the development of environmental DNA for ecosystem monitoring.

Author contributions: tEB was involved in all stages of the research as presented, and in the preparation of the manuscript. MB & AJR supervised the project. OB, SJ, AJR & MB were instrumental in conceptualising and directing the research. AJR & CHD provided access to the samples and collated abiotic data for analysis. MC & MP were involved in the production of the data. BJS, AJR & EH all aided with the data analysis. All authors were involved in drafting and editing the manuscript.

## 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:** Berry, T. E., Coghlan, M. L., Saunders, B. J., Richardson, A. J., Power, M., Harvey, E., Jarman, S., Berry, O., Davies, C. H., & Bunce, M. (2023). A 3-year plankton DNA metabarcoding survey reveals marine biodiversity patterns in Australian coastal waters. *Diversity and Distributions*, *29*, 862–878. <https://doi.org/10.1111/ddi.13699>