MISS KATRINA M. WEST (Orcid ID : 0000-0002-9026-5058)
DR. JOSEPH D. DIBATTISTA (Orcid ID : 0000-0002-5696-7574)
DR. MICHAEL BUNCE (Orcid ID : 0000-0002-7721-4790)

Article type : Original Article

Corresponding author mail id:- katrina.m.west@postgrad.curtin.edu.au

eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem

Katrina M. West[1], Michael Stat[1,2], Euan S. Harvey[1], Craig L. Skepper[3], Joseph D. DiBattista[1,4], Zoe T. Richards[1], Michael J. Travers[3], Stephen J. Newman[3] and Michael Bunce[1,5]

[1] Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Bentley, WA, 6102, Australia.

[2] School of Environmental and Life Sciences, The University of Newcastle, Callaghan, NSW, 2308, Australia.

[3] Western Australian Fisheries and Marine Research Laboratories, Department of Primary Industries and Regional Development, Government of Western Australia, P.O. Box 20, North Beach, WA, 6920, Australia.

[4] Australian Museum Research Institute, Australian Museum, Sydney, NSW, 2010, Australia.

[5] Environmental Protection Authority, 215 Lambton Quay, Wellington 6011, New Zealand.

Abstract

Environmental DNA (eDNA) metabarcoding, a technique for retrieving multi-species DNA from environmental samples, can detect a diverse array of marine species from filtered seawater samples. There is a growing potential to integrate eDNA alongside existing monitoring methods in order to establish or improve the assessment of species diversity. Remote island reefs are increasingly vulnerable to climate-related threats and as such there is a pressing need for cost-effective whole-ecosystem surveying to baseline biodiversity, study assemblage changes and ultimately develop sustainable management plans. We investigated the utility of eDNA metabarcoding as a high-resolution, multi-trophic biomonitoring tool at the Cocos (Keeling) Islands, Australia (CKI) – a remote tropical coral reef atoll situated within the eastern Indian Ocean. Metabarcoding assays targeting the mitochondrial 16S rRNA and CO1 genes, as well as the 18S rRNA nuclear gene, were applied to 252 surface seawater samples collected from 42 sites within a 140 km^2 area. Our assays successfully detected a wide range of bony fish and elasmobranchs (244 taxa), crustaceans (88), molluscs (37) and echinoderms (7). Assemblage composition varied significantly between sites, reflecting habitat partitioning across the island ecosystem and demonstrating the localisation of eDNA signals, despite extensive tidal and oceanic movements. In addition, we document putative new occurrence records for 46 taxa and compare the efficiency of our eDNA approach to visual survey techniques at CKI. Our study demonstrates the utility of a multi-marker metabarcoding approach in capturing multi-trophic biodiversity across an entire coral reef atoll and sets an important baseline for ongoing monitoring and management.

Keywords: environmental DNA, whole-ecosystem, multi-trophic, island reef, biodiversity, monitoring.

1. Introduction

Tropical islands harbour at least 40% of the world's coral reefs (Spalding, Ravilious, & Green, 2001) and support a diverse range of marine ecosystems and communities. Such high island reef diversity results from a complex interplay of biogeographic and oceanographic processes, along with historical and contemporary factors such as disturbance history (Briggs, 1966; Cortés et al., 2017; Graham et al., 2018; Houk, Benavente, Iguel, Johnson, & Okano, 2014). Isolated archipelagos and peripheral islands at the borders of biogeographic provinces, such as those of the Hawaiian archipelago, Mascarene Islands, Galapagos Islands, Pitcairn Island group and Easter Island, are a source of evolutionary novelty and high endemism (Allen, 2008; Budd & Pandolfi, 2010; Friedlander et al., 2013; Hourigan & Reese, 1987). Other peripheral island reef ecosystems such as Christmas Island, the Cocos (Keeling) Islands and the Socotra Archipelago lie in suture zones where marine biogeographical provinces collide, leading to regions of secondary contact that drive hybridisation and diversification (DiBattista et al., 2015; Hobbs, Frisch, Allen, & Van Herwerden, 2008). Documenting and monitoring biodiversity is particularly important at remote reefs because in many cases, these locations can provide baseline information about how 'natural' systems are structured and composed with fewer (or less intense) anthropogenic influences (Knowlton & Jackson, 2008).

The establishment of island reef baselines and subsequent monitoring, that can inform both local and global management decisions, is intrinsically reliant upon cost-effective and robust surveying techniques. The emergence of environmental DNA (eDNA) metabarcoding, an innovative detection technique profiling multi-species genetic material in an environmental sample, is revolutionising our approach to non-invasive, efficient, whole-ecosystem surveying (Bista et al., 2017; Port et al., 2016; Stat et al., 2017). After widespread global testing, this technique has demonstrated its ability to detect a wide variety of biota from aquatic ecosystems, including dark diversity – taxa that are not easily identified using observation-based approaches (Boussarie et al., 2018; Jerde, Mahon, Chadderton, & Lodge, 2011; Vörös, Márton, Schmidt, Gál, & Jelić, 2017). As such, eDNA metabarcoding is anticipated to become a key tool in marine biodiversity discovery and monitoring, as recently outlined in a decadal plan for taxonomy and biosystematics in Australia and New Zealand (Taxonomy Decadal Plan Working Group, 2018). Likewise, the method is capable of responding to UNESCO's 'Decade of Ocean Science for Sustainable

Development (2021-2030) which calls for improved ways to monitor the world's oceans (IOC-UNESCO, 2018).

With high biodiversity values, varied marine environments and a remote geographical position, the Cocos (Keeling) Islands (CKI; 12° 12″ S, 96° 54″ E) are an ideal model system for testing the utility of eDNA metabarcoding in surveying the multi-trophic biodiversity of an entire island reef system. The CKI are one of Australia's external territories, located approximately 1000km southwest of Java, Indonesia, in the tropical eastern Indian Ocean. This remote island group is comprised of a southern and northern coral atoll; the southern atoll (considered the main CKI atoll) is comprised of 24 islets (including the inhabited isles of West Island and Home Island), whilst the northern atoll (referred to as North Keeling) is a singular island atoll protected as a national park. The CKI lie in a suture zone along the marine biogeographic boundaries of the Indian and Pacific Oceans and as such are comprised of species of Indo-West Pacific origin, with little endemism (Hobbs *et al.*, 2014; Woodroffe & Berry, 1994). The islands are renowned for their marine biodiversity; with a total of 602 recorded fish species (Hobbs, Newman, et al., 2014), over 100 species of hard coral (Richards & Hobbs, 2014), 700 mollusc species (Tan & Low, 2014) and 200 crustacean species (Mendoza, Lasley Jr., & Ng, 2014; Morgan, 1994).

Over the last two decades, disturbance events such as mass fish and coral die-offs from deoxygenation, coral bleaching, crowns-of-thorns starfish (*Acanthaster planci*) outbreaks, and cyclonic storm surges (Bunce, 1988; Hobbs & McDonald, 2010; Woodroffe & Berry, 1994), have collectively impacted marine communities at the CKI. The susceptibility of the marine life at the CKI to increasing climate-related threats is exacerbated by the isolation of the island group, as recruitment is restricted from external locations such as from neighbouring Christmas Island (approximately 980 km distance) and Indonesia (approximately 1,600 km), to species that have a long-lived pelagic larval phase (Hourston, 2010). The application of eDNA methods, which has the potential to simultaneously characterise a variety of reef communities, provides an efficient approach for documenting the increasingly vulnerable marine biodiversity of the CKI.

In this study, we conduct a multi-marker eDNA metabarcoding survey of the various marine environments surrounding the main CKI atoll in order to: (i) evaluate the spatial sensitivity of an eDNA signal retrieved from seawater, by assessing its ability to delineate adjacent habitats amidst tidal and oceanic movements within and around an entire island ecosystem, (ii) compare the efficiency of our eDNA approach against extensive visual/video, line fishing and ichthyocide survey techniques previously employed at the CKI, and (iii) provide the first molecular baselines for CKI bony fish, elasmobranch, crustacean, mollusc and echinoderm diversity and community composition, which we anticipate will provide the basis for inclusion into longer term monitoring programs. Lastly, in order to provide a roadmap for surveying tropical island ecosystems, we also explore assay performance, the required levels of replication and the taxonomic resolution that can be achieved at present with our chosen assays and publicly accessible reference databases. The overarching aim was to better evaluate the overall utility of a multi-marker metabarcoding approach towards whole-ecosystem surveying of island reef ecosystems.

2. Materials and Methods

2.1 Field sampling

Six one-litre seawater replicates were sampled from 42 sites in and around the main CKI atoll (Figure 1, SI Table 1) in April 2017, totalling 252 samples across a 140 km^2 area. Samples were obtained from six different habitats zones based on habitat classifications generated from prior environmental surveys (Australian Government, 2005; Hender, McDonald, & Gilligan, 2001; Williams, 1994). Thirteen sites were located within the designated low complexity lagoon area, nine within the medium complexity lagoon, four within the high complexity lagoon, four on the outer reef terrace, 11 within the intertidal reef and one within a lagoonlet site on the West Island (see SI Section 1 for further habitat zone information). In addition, sample sites were further classified based on whether they were in the outer lagoon, inner lagoon or within one of the major surge channels, where oceanic water from the outside of the atoll is flushed through the lagoon at high tide, in a south-east to north-west direction (Hobbs and Macrae, 2012; SI Section 1).

Seawater samples were collected at the surface of each site using bleach sterilised Nalgene bottles and then immediately stored on ice. Each sample was individually filtered across Pall 0.2µm Supor® polyethersulfone membranes using a Pall Sentino® Microbiology pump (Pall Corporation, Port Washington, USA), within three hours of collection. Samples of the 10% bleach solution - used to clean filtration equipment over the course of each sampling day - were included as filtration controls. These filtration controls serve to detect any potential cross-contamination in water filtering between successive sites and during handling post-filtering. The filter membranes were immediately frozen and stored at -20°C prior to their transportation to a quarantine facility within the Trace & Environmental DNA (TrEnD) Laboratory in Perth, Western Australia; filters were stored long-term at -80°C.

2.2 Laboratory processing

DNA was extracted from the filter membranes, within two weeks of collection, using a DNeasy Blood and Tissue Kit (Qiagen; Venlo, Netherlands) with the following modifications: 540µl of ATL lysis buffer and 60µl of Proteinase K during the cell digestion phase. Negative controls, containing no sample (or filters), were extracted and processed alongside all samples in order to detect any cross-contamination. Previously published primers were sourced to amplify bony fish, elasmobranchs, crustaceans and other eukaryotes (including molluscs and echinoderms), respectively, from mixed environmental samples. The four applied PCR assays, herein referred to as 16S Fish, COI Elasmobranch, 16S Crustacean and 18S Universal (Table 1), were incorporated into our eDNA metabarcoding workflow. The multiplex COI Elasmobranch assay (forward primers: Fish F2 and V52 (Ward, Zemlak, Innes, Last, & Hebert, 2005) and reverse primer: Shark COI-MINIR (Fields, Abercrombie, Eng, Feldheim, & Chapman, 2015) had two degenerate positions (Ns) added into each primer to increase the variability of elasmobranch amplicons. Quantitative PCR (qPCR) protocols are described in depth in SI Section 2.

Pooled libraries were then size-selected using a Pippin Prep (Sage Science, Beverly, USA), purified with a Qiaquick PCR Purification Kit (Qiagen, Venlo, Netherlands), quantified using a Qubit 4.0 Fluorometer (Invitrogen, Carlsbad, USA), and loaded onto either a 300 cycle (for unidirectional sequencing) or 500 cycle (for paired-end sequencing) MiSeq® V2 Standard Flow Cell on an Illumina MiSeq platform (Illumina, San Diego, USA), housed in the TrEnD Laboratory at Curtin University (SI Section 2).

2.3 Bioinformatics & taxonomic assignments

Sequences were merged, demultiplexed and filtered using a combination of AdapterRemoval (v2; Schubert, Lindgreen, & Orlando, 2016), Geneious® (10.0.6; Kearse *et al.*, 2012), USEARCH (v9.0.2132; Edgar, 2010) and OBITools (v1.2.9; Boyer, Mercier, Bonin, Taberlet, & Coissac, 2014). Quality filtering parameters included a minimum length of 100bp, a maximum expected error rate of 0.5, no ambiguous bases (N's), a minimum read count of 5, a read threshold ratio of 0.05% and no chimeras (see SI Section 3 for pipeline details). Quality filtered dereplicated sequences were then queried against the National Centre for Biotechnology Information's (NCBI) GenBank nucleotide database (accessed in 2017/18; Benson *et al.*, 2005) and an in-house 16S rDNA fish database (SI Table 6) via Zeus, an SGI cluster, based at the Pawsey Supercomputing Centre in Kensington, Western Australia. Taxonomic assignments were made in MEGAN (MEtaGenome Analyzer v5.11.3; (Huson *et al.*, 2007; see SI Section 3) and classified based on geographic distribution and habitat association data obtained from FishBase (Froese & Pauly, 2018) and the World Register of Marine Species (WoRMS; Horton *et al.*, 2018).

To present putative new occurrence records at the CKI, we required that all global congeneric taxa (of the putative species in question) had been barcoded for the gene region of the respective assay; this removes the possibility that a new occurrence record is actually that of a closely-related (not yet barcoded) species. Rarefaction analyses were conducted using QIIME (v1; Caporaso *et al.*, 2010), and assay datasets subsampled in USEARCH (SI Section 3). Taxa lists were converted into presence/absence matrices for taxonomic groups of interest (Fish &

Elasmobranchs, Crustaceans, Molluscs, and Echinoderms) using fuzzySim (Barbosa, 2015) in RStudio (v1.1.423; RStudio Team, 2015).

2.4 Statistical analyses

Subsampled datasets were employed for univariate and multivariate statistics comparing diversity and community composition between sites/habitats at the CKI (see SI Section 4). Observed taxonomic diversity (equivalent to species richness) at each site was calculated using the package 'BAT' (Biodiversity Assessment Tools; Cardoso, Rigal and Carvalho, 2015) and graphed using ggplot2 (Wickham, 2016) in RStudio. Community composition was visualised by non-metric multi-dimensional scaling (nMDS) and tested for the amount of variation explained by spatial predictor variables (distance - latitude/longitude, depth, lagoon and marine habitat) by distance-based linear modelling (DistLM) in the PERMANOVA+ add-on (Anderson, Gorley, & Clarke, 2008) of PRIMER v7 (Clarke & Gorley, 2015). Permutational analysis of multivariate dispersions (PERMDISP; Anderson et al., 2008) was used to test homogeneity of variance within marine habitat zones – there were no significant differences for any of the taxonomic groups. Similarity percentage analyses (SIMPER) were also conducted in PRIMER, and an indicator species analysis performed using the R package 'indicspecies' (De Caceres & Jansen, 2016).

3. Results

3.1 Sampling and sequencing statistics

The four metabarcoding assays run across 252 seawater eDNA extracts yielded a total of 61,219,212 sequencing reads. The mean number of quality filtered sequences per sample was 35,931 for the 16S Fish assay; 30,390 for the 16S Crustacean assay; 42,476 for the COI Elasmobranch assay; and 28,123 for the 18S Universal assay (SI Table 7). The rarefaction analyses determined suitable subsampling cut-offs of 8000 reads for the 16S Fish assay, 4000 for the 16S Crustacean assay, 4000 for the COI Elasmobranch assay, and 8000 for the 18S Universal assay (SI Figure 1). Assay performance is reported in SI Section 5 and SI Figure 2.

After examining taxa accumulation based on the addition of sampling replicates at each site, it was observed that six water replicates (chosen a priori) was not quite sufficient to fully sample the fish and elasmobranch diversity, as well as the crustacean and mollusc diversity from each of the respective sites at the CKI (SI Figure 3). On fitting polynomial curves to the taxa accumulation data, it was extrapolated that an average of 8.38 ± 1.33, 11.75 ± 8.73 and 8.23 ± 3.32 water replicates would be required to sample the fish and elasmobranch, crustacean and mollusc diversity at each site, respectively. However, in using six replicates, we were still able to detect on average 90.1% of fish and elasmobranch, 79.2% of crustacean and 89.0% of mollusc diversity at each site. Six water replicates were adequate to sample echinoderms with the applied assays, with an average of 5.91 ± 2.39 water replicates required.

Two filtration controls amplified (with qPCR C_T values lower than 40 cycles) for the 18S Universal assay; all other assays did not amplify controls. The controls were subsequently sequenced and produced detection hits for algae (order: Pleurostomatida and infrakingdom: Heterokonta), fungi (genus: Cladosporium), insects (superfamily: Membracoidea), demosponges (order: Haplosclerida) and plants (clade: Rosids). Given that these groups were not targeted or of interest for this project, they were omitted from further analyses for all samples. One sample contained a detection hit for the subfamily Salmoninae; this taxon is sporadically detected as a reagent contamination in our workflows and has been reported elsewhere as fish contamination (Thomsen et al., 2016) - accordingly, it was removed from the fish dataset.

3.2 Overall diversity

The four metabarcoding assays collectively yielded 372 identifiable taxa, representing 110 families within 40 orders (Figure 2, SI Table 8) across the targeted taxonomic groups at the CKI. The following diversity and community composition data provides new baselines for bony fish, elasmobranch, crustacean, mollusc and echinoderm groups across the sampled CKI habitat zones.

Two hundred and forty fish and elasmobranch taxa were detected (219 at a genus or species level of assignment) from 48 families within 17 orders (SI Table 8). Of these, 84.2% of the bony fish and elasmobranch taxa are widely distributed across the Indo-Pacific region, 2.1% distributed in the Pacific Ocean only, 2.1% distributed in the Indian Ocean only, 11.2% are circumglobal and 0.4% have uncertain distributions. There were no species detected at the CKI that are endemic. The majority of detected bony fish and elasmobranch species (90%) are associated with coral reefs (1-60m depth), 0.8% with deep reef (>60m depth), 3.8% with coral and deep reef, 4.6% are pelagic, 0.4% with coral and deep reef and are also pelagic, and 0.4% with coral and intertidal reef. Predominant bony fish families included Labridae (32 taxa), Pomacentridae (23), Blenniidae (15), Gobiidae (14) and Muraenidae (13). The only elasmobranch families detected were Mobulidae (1), Myliobatidae (1) and Carcharhinidae (4). Particular taxa of interest that have subsistence, recreation or conservation importance at the CKI are presented in Table 2 and Figure 3. In addition, we report 18 putative new fish occurrence records at the CKI (see SI Table 9 for sequencing statistics and evaluation of the new occurrence records), including goldband snapper (*Pristipomoides multidens*) and a number of deep-sea fishes including lanternfish (*Lampanyctus*, *Lampadena*, order Myctophiformes) and dragonfish (order Stomiiformes).

Within the Crustacea, we focussed on decapods, krill (order Euphausiacea) and mantis shrimp (order Stomatopoda) for this study. This resulted in a total of 88 crustacean taxa (77 assigned at a genus or species level) from 28 families within the orders of Decapoda, Euphausiacea and Stomatopoda (SI Table 8). Of these crustacean taxa, 60% are widely distributed across the Indo-Pacific region, 6% distributed in the Pacific Ocean only and 34% are circumglobal. There were no taxa solely of Indian Ocean origin or endemic to CKI. Predominate families include Xanthidae (34 taxa), Alpheidae (8), Diogenidae (6) and Grapsidae (4). Taxa of interest (Table 2) include subsistence species such as the mangrove swimming crab (*Thalamita crenata*) and the fourspine rock lobster (*Panulirus penicillatus*), in addition to prawns (Penaeidae), which were recorded for the first time at CKI. Overall, we detected 25 crustaceans not previously recorded at the CKI, largely xanthid crabs and snapping shrimp (SI Table 9).

Thirty-seven mollusc taxa were detected (23 at a genus or species level of assignment) from 29 families within 17 orders (SI Table 8). Of these mollusc taxa, 13.5% are widely distributed across the Indo-Pacific region, 5.4% distributed in the Pacific Ocean only, 8.1% distributed in the Indian Ocean only and 73% are circumglobal. No taxa were endemic to CKI. No singular family dominated the diversity of molluscs detected at CKI; all families were represented with only one or two taxa. This included the gastropod families Bullidae (2 taxa), Littorinidae (2), Aplysiidae (2) and Strombidae (2), and the bivalve mollusc families such as Limidae (2) and Mytilidae (2). Taxa of interest (Table 2) included subsistence species such as the spider conch (*Lambis lambis*) and oysters (family Ostreidae), as well as pearl oysters (*Pinctada sp.*). There were two putative new mollusc occurrence records at the CKI (SI Table 9), including the sea slug genus *Stiliger* and the bivalve genus *Basterotia.*

Only 7 echinoderm taxa were detected (5 at a genus or species level of assignment) from 5 families within 3 orders (SI Table 8). Of these echinoderm taxa, 14.3% are associated with an Indian Ocean distribution; the remaining 85.7% are circumglobal. Similar to the molluscs, no single echinoderm family dominated the biodiversity. The five families detected were the brittle star families Ophiocomidae (3), Ophiodermatidae (1) and Amphiuridae (1), and the sea cucumber families Synaptidae (1) and Holothuriidae (1). There was one putative new echinoderm occurrence record at the CKI, with the detection of *Ophioderma*, a genus of brittle star (SI Table 9).

3.3 Community composition

Taxon richness varied across habitat zones (SI Figure 4) with significant differences in diversity detected between intertidal reef and high complexity lagoon habitats; the latter with the lowest taxon richness ($P < \alpha$ H-B; SI Tables 10 and 11). The highest diversity of bony fish/elasmobranchs and crustaceans was detected in the intertidal reef (average of 47.2 ± 5.4 bony fish/elasmobranchs and 9.7 ± 1.7 crustaceans per site), however this was not significantly different for bony fish/elasmobranchs ($P > \alpha$ H-B; SI Figure 5; SI Tables 10, 12 and 13).

Echinoderm richness was highest in the low complexity lagoon (average of 3.6 ± 0.4 taxa per site; SI Figure 4; SI Tables 10 and 14), whilst mollusc richness revealed no significant differences between habitat types (SI Table 10). The proportion of the major taxonomic groups amplified and represented at each site was consistent between marine habitat zones. The only deviations were a lack of echinoderms in the high complexity lagoon area and an increased proportion of crustaceans in the lagoonlet compared to all other marine habitat zones (SI Figure 5).

Overall taxonomic richness was also investigated for variation between sites within the outer lagoon, surge channels and inner lagoon (SI Figure 6 & 7). There were no significant differences in richness between lagoon zones, except when examining echinoderms independently (SI Table 15). A significant difference was detected in echinoderm richness between the surge channel (4.1 ± 0.3 taxa) and the inner lagoon zone (1.9 ± 0.5 taxa per site; $P < \alpha$ H-B; SI Table 16).

The taxa composition of bony fish, elasmobranch and crustacean communities is clearly seen to transition across the habitat and lagoon zones (nMDS; Figure 4). Within the mollusc and echinoderm ordinations however, very little clustering based on lagoon and marine habitat zones was observed. Distance based linear model (DistLM) analyses confirmed that community variation based on marine habitat was highly significant between sites ($P < 0.05$) and explained the highest proportion of fitted variance for all taxonomic groups (Table 3; SI Tables 17-20), except echinoderms where habitat was a non-significant predictor variable. Conversely, the lagoon zones (outer lagoon, surge channel, inner lagoon) explained the highest proportion of fitted variance for echinoderms; additionally, it was also a highly significant predictor variable for all taxonomic groups. Distance between sites (latitude/longitude) formed significant predictor variables for bony fish and elasmobranchs as well as molluscs and echinoderms; however, it only explained a small proportion of the variability for these taxonomic groups. For crustaceans, distance between sites did not influence their community composition. Finally, depth explained a small, yet significant, proportion of fish and elasmobranch community composition. Cumulatively, the spatial predictor variables tested in the DistLM analyses formed models that explained between 30-44% of total fitted variance between sites. Where the predictor variables of

marine habitat and lagoon zones were significant in the DistLM analyses, similarity percentage analyses (SIMPER) were used to identify prominent taxa contributing most to the similarity between sites within each zone. These assignments are reported in SI Section 6 and in SI Tables 21-27, in addition to a variety of indicator species in SI Tables 28 and 29.

3.4 Comparison of eDNA to previous CKI marine monitoring techniques

An extensive set of fish collections and site surveys conducted at the CKI between 1973-2017, have recorded in total, 626 bony fish and elasmobranch species from 85 families (Hobbs, Newman, et al., 2014; Harvey et al. (unpublished)). Our eDNA metabarcoding survey from a single sampling point identified 240 bony fish and elasmobranch taxa from 48 families, including 18 putative new fish occurrence records. This raises the total fish record at the CKI to 644 taxa from 88 families, with our single eDNA survey capturing 37.3% of the total CKI fish fauna. This level of detection is comparable to recent baited remote underwater video (BRUV) surveys conducted in 2016 and 2017 (Table 4) which captured 41.5% and 38.0% of the total CKI fish fauna, respectively (Harvey et al. (unpublished); see SI Tables 30 and 31 for BRUV species lists). Our eDNA survey however, was not as effective in capturing CKI bony fish and elasmobranch diversity as previously used invasive techniques, such as rotenone, explosives, quinaldine, angling, dipnets and spearing, which were collectively applied in two successive surveys in 1973 and 1989 (Table 4; Allen & Smith-Vaniz, 1994). Although, it is worth noting only a single (16S) bony fish assay was employed and additional assays would have the potential to detect further fishes.

A total of 198 species of Decapoda (e.g. crabs, shrimps, lobsters) and 13 species of Stomatopoda (mantis shrimp) have previously been recorded at the CKI (Ahyong, 2014; Morgan, 1994). Our eDNA survey detected a total of 88 crustacean taxa; 83 Decapoda, 4 Stomatopoda and the first official record of Euphausiacea (krill) at the CKI. The detection and verification of 25 putative new crustacean occurrence records from the eDNA survey raises the total Decapoda recorded at the CKI to 221 taxa, the Stomatopoda recorded to 14 taxa and the Euphausiacea recorded to one taxon. However, an above and underwater observation survey for crustaceans (Morgan, 1994)

still detected a greater fraction of crustacean groups (89.6% of recorded Decapoda) than eDNA metabarcoding (37.3%) at the CKI (see Table 4).

4. Discussion

4.1 eDNA metabarcoding baseline of a tropical island reef ecosystem

This study is the first to conduct a multi-trophic eDNA survey of an entire coral reef atoll and demonstrates the utility of multi-assay Tree-of-life (ToL) eDNA metabarcoding approaches in assessing whole-ecosystem biodiversity (Stat et al., 2017). This non-destructive technique exhibited a high detection capability that is particularly suitable for characterising island reef habitats. There is a pressing need to efficiently survey remote, island ecosystems given that they support an exceptionally high level of marine biodiversity and/or endemism (Roberts et al., 2002), yet through their isolation are increasingly vulnerable to climate-related threats (Whittaker & Fernández-Palacios, 2007). ToL metabarcoding approaches (Stat et al., 2017) are of particular use for environmental management in Small Island Developing States (SIDS), which typically share resource challenges, yet are at the forefront of the downstream effects of global climate change (United Nations, 2012). While our survey is restricted to a single time-point, the data provides novel information on the biodiversity present at the CKI, demonstrates the value of non-destructive approaches to assess community composition, simultaneously assesses multitrophic diversity, establishes a roadmap for ongoing monitoring, and raises new biogeographic hypotheses regarding range extensions and habitat health at the CKI. The value of long-term eDNA chronosequences was recently demonstrated (Berry et al., 2019) and this study will establish a baseline upon which regular eDNA biomonitoring can take place.

4.2 Bony fish and elasmobranch diversity and community composition

The community composition of the fish and elasmobranch taxa detected by eDNA was largely congruent with the previously established fish diversity profile at the CKI. The predominant fish families of the eDNA dataset were all in the top ten recorded families with greatest species

richness at the CKI and also neighbouring Indonesia (Hobbs, Newman, et al., 2014). The highest taxonomic diversity of fish was observed in the intertidal reef and the lowest in the high complexity lagoon. This disparity was surprising given previous surveys have documented a high diversity and abundance of butterflyfish and herbivorous fish species in the outer reef terrace, as well as high and medium complexity lagoon zones, which incidentally were noted in previous surveys to feature the most developed and healthy reefs of the CKI (Australian Government, 2005; Hender et al., 2001). Recent surveys of coral reef health and distribution at the CKI however, noted a 50-60% reduction of hard coral cover and genera abundance in the lagoon between 2010-2014 (Evans, Konzewitsch, & Bellchambers, 2016) and substantial numbers of dead branching and plating coral across the high complexity lagoon zone in 2018 (Z Richards, personal communication, December 2018).

Habitat heterogeneity at CKI, as defined by Hender, McDonald and Gilligan (2001) in their classification of habitat zones (SI Section 1), was a significant spatial predictor variable in our survey, explaining the highest proportion of fish and elasmobranch variability. The position of each site either in the outer lagoon, surge channel or inner lagoon, alongside the distance and depth between sites were also found to be significant predictors, yet explained a lower proportion of variability in community composition. This indicates that fish and elasmobranch community composition across the CKI is largely shaped by marine habitat preferences, however the small, yet significant effect of the lagoonal zones, does indicate that species move across habitat zones. This is likely influenced by tidal changes, whereby at low tide, the low complexity area at the southern end of the inner lagoon partially dries (Woodroffe & Biribo, 2011). Likewise, in the outer lagoon, species would be expected to migrate off the intertidal reef amidst exposure with daily tidal changes.

There is a high level of daily water movement via surge channels at CKI, hence it is also possible that an eDNA signal from the inner lagoon may contain DNA shed in the outer lagoon and subsequently transported; thus, the habitat profile of inner lagoon sites would be confounded. However, there was no significant difference between the taxonomic diversity of fish in the three lagoon zones. In addition, we observed a clear transition of fish community composition between

sites situated in the outer lagoon, surge channel and inner lagoon (see SI Section 7 for further fish community composition information). This indicates that eDNA signals from seawater are localised enough to discriminate between sites over small spatial scales and can provide a reproducible profile of their community composition. This data adds to a growing body of evidence that there is strong spatial fidelity in eDNA signals, attributed to dispersion and degradation in seawater (Jeunen et al., 2018; Koziol et al., 2018; Murakami et al., 2019; Stat et al., 2019).

The detection of unresolved deep-sea lanternfish and dragonfish assignments (order: Myctophiformes and Stomiiformes) illustrate the potential of using eDNA metabarcoding to identify cryptic biodiversity, particularly of taxonomic groups that inhabit pelagic zones and are largely inaccessible with other monitoring techniques. The expansion of reference sequence databases remains essential to resolve assignments to a high taxonomic resolution. Whilst it was not possible to elucidate the unresolved Myctophiformes and Stomiiformes sequences beyond order level (exempting the lanternfish *Lampadena luminosa*, *Lampanyctus* and *Bolinichthys*), it may be relevant that these taxonomic groups consist of small to medium-sized predatory fishes known to undertake diel vertical migration; whereby they inhabit meso- and bathypelagic zones during the day before migrating to surface waters at sunset to feed (Watanabe, Moku, Kawaguchi, Ishimaru, & Ohno, 1999). Additionally, there are reports of Stomiiforme larvae developing in near-surface layers, prior to descending to the meso- and bathypelagic zones (Fahay, 2007). It is therefore not unexpected that we detected these groups in the intertidal and adjacent medium complexity reefs. The detection of deep-sea fish, which we attribute to diel-vertical migration and the surfacing of larvae, does however, raise a potential issue with the crossover detection of eggs/larvae and mature fish, as at present there is no way to discriminate them using an eDNA metabarcoding approach. During spawning periods, there is likely a sizeable increase in (gamete derived) mitochondrial DNA, and to a greater extent nuclear DNA (Bylemans et al., 2017), which in conjunction with water movement could confound community profiles. Therefore, some caution must be taken when analysing eDNA metabarcoding data on small spatial scales that are subject to extensive tidal and current movement, or even cyclical spawning events.

Six of the nine elasmobranch taxa previously recorded at the CKI were detected using our eDNA method. The genus of *Mobula* (devil rays) was the only new elasmobranch occurrence record, however recent phylogenetic research indicates that *Mobula* is paraphyletic to *Manta* (manta rays) and subsequently it was recommended that *Manta* be revised as a subgroup under *Mobula* (White et al., 2018). *Manta birostris* is a known elasmobranch species at the CKI, commonly found in cleaning stations between Home Island and Direction Island (Z Richards, personal communication, December 2018) and on the southern edge of West Island (Brewer et al., 2009). This species was widely detected by eDNA at 11 sites across the atoll, including the intertidal reef and outer reef terrace.

4.3 Crustacean diversity and community composition

The Xanthidae are the predominant decapod group at the CKI, comprising approximately 38% (34 species) of crustaceans detected in our survey. This is congruent with previous surveys whereby xanthid taxa comprised 36% of Decapoda and Stomatopoda records at the CKI (Ahyong, 2014; Morgan, 1994) and validates that the 16S Crustacean primers (Berry et al., 2017) have the capability to accurately detect these crustacean groups. However, we only detected six of a possible 48 species in the Diogenidae family (marine hermit crabs), which after xanthid crabs, are thought to be one of the most diverse decapod groups at the CKI (Morgan, 1994). An *in-silico* analysis established that the 16S Crustacean primer assay has the ability to amplify the majority of taxa within Diogenidae. Thus, there seems to be no obvious explanation for this reduction in taxonomic diversity, as we surveyed extensively across their preferred habitats of sheltered, shallow intertidal reef and sandy areas. That the Diogenidae species were largely detected on the outer side of the lagoon is consistent with Morgan's (1994) survey indicating an absence of marine hermit crabs in the northern and southern part of the lagoon off West Island.

Another significant finding in the crustacean analysis, was the detection and first official occurrence record of Penaeidae (prawns) at the CKI. There is only one recorded sighting of Penaeidae at the CKI, with reference to an unidentified prawn being fed on by a *Thalamita*

crenata (McKillup & McKillup, 1996). The lack of Penaeidae in any formal survey at the CKI is surprising, given the widespread abundance of the family, particularly around Australia and Indonesia and its presence at neighbouring Christmas Island (Morgan, 2000). Interestingly, our detection could only be assigned at a family level (Penaeidae) because of a low identity percentage, indicating that we do not have a reference sequence for the detected taxa. This largely excludes the possibility that our detection is of imported prawns from mainland Australia for the purposes of human consumption, as Australian commercial prawn species (e.g. tiger prawns, king prawns and banana prawns) have numerous publicly-accessible reference sequences. The detection of Penaeidae at two sites (4, 30) at the CKI can now be used to guide further surveys in an effort to recover specimens for taxonomic and reference sequencing purposes.

The community composition of crustaceans is similar to that of the bony fish and elasmobranchs, with the habitat and lagoon zones explaining the most variation. The genus of *Thalamita* (Portunidae, Decapoda) was the most widespread crustacean taxa, representative in all habitat and lagoonal zones, except for the high complexity lagoon zone. This genus contains a plethora of species found at the CKI both in the shallows of the lagoon and on the outer edge of the atoll (Tweedie, 1950). The genus of *Cyclodius* (Xanthidae, Decapoda), in particular *Cyclodius ungulatus*, typified the high complexity lagoon zone and is largely associated with dead coral and coral rock, although has also been recorded among living *Pocillopora* coral (Tweedie, 1950). The omnivorous *Daira perlata* (Daridae, Decapoda) was widespread both in the low complexity lagoon and intertidal reef, dominating the latter. *Daira perlata* has been previously reported in crevices and under rocks on the outer edge of the CKI atoll (Tweedie, 1950). Lastly, *Chlorodiella barbata* (Xanthidae, Decapoda) was widespread in the low and medium complexity lagoon zones, found in both the surge channel and inner lagoon. Whilst its habitat distribution has not been formally reported at the CKI, *Chlorodiella barbata* has been reported in a number of lagoon, fringing reef and barrier reef sites in French Polynesia (Monteforte, 1987; Peyrot-Clausade, 1989).

4.4 Mollusc diversity and community composition

A total of 757 mollusc species from 126 families have been previously recorded at the CKI; 80.6% of the class Gastropoda (snails and slugs) and 18.4% Bivalvia (bivalve molluscs) (Tan & Low, 2014). Our mollusc analysis, however, only detected a small fraction of these reported species with a total of 37 mollusc taxa across 29 families. This low level of detection likely reflects the fact that we did not apply a specifically-designed mollusc assay, but rather extracted mollusc assignments from a 'universal' 18S assay. As discussed previously, the 18S universal assay amplified 72 classes with the majority of BLAST hits at a family level. Therefore, it is recommended that for an in-depth eDNA mollusc analysis or to monitor threatened species such as giant clams, a more specific primer assay must be developed and applied. From the molluscs that were amplified however, we were able to discriminate community composition variation between sites. This was again largely attributed to habitat variation, in conjunction with distance (latitude/longitude) between sites and the site's position within the inner lagoon, surge channel or outer lagoon.

Lambis lambis, known as the common spider conch or gong gong locally, is a gastropod mollusc that is extensively fished using artisanal techniques in the CKI where it is considered a delicacy. Preferring algal-dominated lagoon flats, the southeast area of the lagoon once supported a high density of *L. lambis*. However, recent surveys indicate a change in their distribution, which may be due to macroalgal habitat declines (Bellchambers, Meeuwig, Evans, & Legendre, 2011) and an overall reduction in stock abundance between 2008-2014 (Evans et al., 2016). Not many of the sites surveyed overlapped with those of Evans, Konzewitsch and Bellchambers (2016); nonetheless we only detected *L. lambis* at one site (20), in the medium complexity lagoon area. This may reflect low population size in this area, however further eDNA and visual surveying across their preferred habitats is necessary.

4.5 Echinoderm diversity and community composition

The last updated checklist of echinoderms at the CKI reported a total of 89 species across 5 classes (Crinoidea, Asteroidea, Ophiuroidea, Echinoidea and Holothuroidea; Marsh, 1994). We however, only detected 7 echinoderm taxa from 5 families across 2 classes. Limited detection is

again attributed to the use of the 18S universal assay, rather than applying a more specific echinoderm assay. Unlike the bony fish and elasmobranch, crustacean and mollusc communities, habitat, in addition to depth, was not a significant predictor variable for echinoderm community composition. This may reflect the ability of the detected echinoderm taxa to occupy mixed habitats, however, there was a lack of echinoderm DNA detected in the four sites located in the high complexity lagoon area of the inner lagoon. Whilst it could be argued that by sampling surface water we may be excluding benthic species, we did detect echinoderms on the outer reef terrace, which had the greatest depth profile by far (average depth ~ 13 m). Therefore, our lack of detection likely reflects a low abundance of echinoderms in the high complexity lagoon area.

4.6 Future integration in reef monitoring

As demonstrated in this survey, eDNA metabarcoding is a powerful detection tool, capable of producing localised profiles of marine ecosystem composition amidst tidal and oceanic movements, which are reflective of habitat partitioning on an island scale. The eDNA approach proved to be equally efficient in detecting fish and elasmobranch diversity to recent BRUV surveys at the CKI. Interestingly there was no overlap in new occurrence records between the BRUV and eDNA surveys, demonstrating their potential utility to provide complementary baseline data, as previously reported by Stat et al., (2019). Whilst both the eDNA and BRUV surveys fall short of the efficiency of line and spearfishing as well as ichthyocides (Allen & Smith-Vaniz, 1994) as a fish survey tool at the CKI, they are preferable in that they are non-lethal; the added benefit of eDNA metabarcoding is that it is completely non-invasive. However, like visual surveillance approaches, putative new occurrence records derived from eDNA will require further verification by specimen collection. Despite this, eDNA will be useful for directing targeted surveying, not only to verify new occurrence records, but also in an effort to physically recover unresolved taxa that may represent new species.

In capturing the crustacean diversity at the CKI, our eDNA crustacean primer assay fell short of traditional above and underwater observation surveys (Morgan, 1994). However, the application of this assay did contribute 25 new occurrence records for crustaceans at the CKI, demonstrating

that there remain cryptic and/or elusive taxa that have yet to be surveyed in this region. The low percentage of known crustacean diversity captured by the 16S Crustacean primers (Berry et al., 2017), in comparison to previous survey techniques, may reflect a need for further primer optimisation in addition to the continual development of reference databases, given that the majority proportion of 16S Crustacean amplicons were attributed as no hits (SI Section 5). Alternatively, this may reflect preliminary research indicating lower eDNA detection rates from crustaceans, in contrast to fish and amphibian species (Forsström & Vasemägi, 2016; Tréguier et al., 2014).

Whilst eDNA metabarcoding is now widely employed in a number of settings, there remains some uncertainties in regard to experimental design, assay choice and laboratory/bioinformatic workflows that need to be optimised before this technique can be fully integrated into monitoring toolkits (see Cristescu and Hebert, 2018). In testing the utility of eDNA metabarcoding to survey the marine biodiversity of the CKI, we encountered some of these uncertainties, which we wish to address here, in anticipation that our findings will guide further eDNA surveying in tropical islands and more broadly, coral reef ecosystems. Our sampling design (six water replicates per site) was determined *post hoc* to be suboptimal with eight samples being the optimal replication required to completely sample the bony fish, elasmobranch, crustacean and mollusc diversity at sites within the CKI. Given that coral reefs support an exceptionally high level of marine biodiversity (Roberts et al., 2002), it is recommended that where possible, a pilot study is undertaken prior to a large survey in order to evaluate the level of replication required (for the given question). Alternatively, preference should be given to collecting a high yet feasible number of water replicates in the field, not all of which might progress through metabarcoding workflows.

Whilst there was some overlap in the detection of taxonomic groups, the four assays applied in this eDNA survey each produced unique taxon hits, which further supports their use as stand-alone metabarcoding assays. It is therefore recommended that for metabarcoding, the 16S Fish assay is used to amplify a wide range of bony fish at a fine-scale level of taxonomic resolution (family to species level), the 16S Crustacean assay is used to amplify crustaceans, whilst the COI

Elasmobranch assay is primarily used for the amplification of elasmobranchs. The majority of BLAST hits extracted from the 18S Universal assay were at a family-level resolution; the power of this assay lies in its ability to anneal to a wide range of taxonomic groups. We therefore recommend that the 18S Universal assay is employed to explore the range of taxonomic groups present in an environmental sample, prior to using a secondary assay(s) to increase resolution of target taxa.

Global research into the use of eDNA metabarcoding widely supports its implementation as a marine detection tool alongside traditional, observation-based monitoring techniques (Kelly et al., 2017; Stat et al., 2019; Taberlet, Bonin, Zinger, & Coissac, 2018). All traditional techniques in current use have biases and limitations ranging from inabilities to detect small and cryptic species, requirements for the use of bait, invasiveness (e.g. lethal methods), as well as logistical and cost-constraints. Whilst eDNA metabarcoding bypasses many of these concerns, it is not without its own limitations. This is largely in regard to its reliance on incomplete reference databases, where many species have yet to be barcoded, and, at present, its limited use for quantification above and beyond presence-absence surveys (McInerney & Rees, 2018). It is therefore highly recommended that eDNA metabarcoding is used in conjunction with widely employed morphological assessments, such as through line fishing, trawling and visual/video surveillance, which additionally provide information on size profiles and biomass. Ultimately, the method employed will depend on the primary biomonitoring question at hand, but a combined approach with eDNA metabarcoding has proven to yield substantially greater taxonomic richness than with either a sole genetic or sole morphological method (Kelly et al., 2017; Stat et al., 2019). Building on rapid developments in eDNA metabarcoding, it is expected that this genetic approach will soon be integrated into best-practice marine resource management practices.

5. Conclusion

The application of a fine-tuned suite of eDNA metabarcoding assays can characterise whole marine ecosystem diversity, reveal fine-scale reef communities and provide putative new occurrence records that can efficiently establish or update marine baselines. This study

demonstrated the localisation of eDNA signals from seawater, providing community composition data that reflected habitat partitioning across an entire coral reef atoll. It provided high resolution data (i.e. to genus and species level) that is complementary to recent stereo-video surveillance research on bony fish and elasmobranchs at the CKI. An updated multi-trophic baseline, with the discovery of 46 new occurrence records - notably species of mullet, snapper and deep-sea fish - will provide the basis for inclusion into longer term monitoring programs. Repeated observations of eDNA measurements across the georeferenced sites herein will facilitate assessment of the status and ultimately trends in biodiversity, particularly in response to disturbance events at CKI (e.g. deoxygenation, coral bleaching). This can be integrated alongside traditional surveying methods to yield greater taxonomic richness and provide additional information in relation to size profiles and biomass. There is also the added benefit that eDNA samples can be archived, and as new primer assays for improved detection of echinoderms or molluscs, or any other taxa of interest are developed, samples can be retrospectively analysed. We strongly advocate for the ongoing development and application of multi-marker eDNA metabarcoding approaches to efficiently characterise and establish multi-trophic island reef baselines. In addition, continual barcoding of endemic and cryptic taxa, particularly deep-sea marine species, will aid in resolving ambiguous taxonomic assignments, which will be beneficial in future island reef metabarcoding projects.

6. Acknowledgements

We acknowledge the Department of Primary Industries and Regional Development (WA), BMT Oceanica and the Australian Research Council (LP160100839) for project funding. For fieldwork assistance we thank Chynna Cahill and Parks Australia. In the development of the 16S fish database we would like to thank R. Czarnik, E. Fenny, C. Wakefield, D. Fairclough, S. Payet, N. Jarvis, S. Moyle, D. Boddington, and K. Smith. For valuable assistance in the TrEnD laboratory we thank Megan Coghlan, Matt Power, Adam Koziol, Frederik Seersholm, Matthew Heydenrych, Kristen Fernandes, Alicia Grealy & Tina Berry. We acknowledge bioinformatic assistance from the Pawsey Supercomputing Centre (WA). For fish distribution data at the

CKI, we would like to thank J.P. Hobbs. Finally, we give thanks to Tane Sinclair-Taylor for providing photos taken in the CKI.

7. References

Ahyong, S. T. (2014). Stomatopod Crustacea of Christmas Island and the Cocos (Keeling) Islands. *Raffles Bulletin of Zoology*.

Allen, G. R. (2008). Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *18*(5), 541–556.

Allen, G. R., & Smith-Vaniz, W. F. (1994). Fisheries of the Cocos (Keeling) Islands.

Anderson, M., Gorley, R. N., & Clarke, R. K. (2008). *Permanova+ for Primer: Guide to Software and Statisticl Methods*. Primer-E Limited.

Australian Government. (2005). *Status of the coral reefs at the Cocos (Keeling) Islands: a report on the status of the marine community at Cocos (Keeling) Islands, east Indian Ocean, 1997-2005*. Australia.

Barbosa, A. M. (2015). fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*, *6*(7), 853–858.

Bellchambers, L. M., Meeuwig, J. J., Evans, S. N., & Legendre, P. (2011). Modelling habitat associations of the common spider conch in the Cocos (Keeling) Islands. *Marine Ecology Progress Series*, *432*, 83–90.

Bennett, S., Halford, A. R., Choat, J. H., Hobbs, J. A., Santana-Garcon, J., Ayling, A. M., … Newman, S. J. (2018). Geography and island geomorphology shape fish assemblage structure on isolated coral reef systems. *Ecology and Evolution*, *8*(12), 6242–6252.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). GenBank. *Nucleic Acids Research*, *33*(Database Issue), D34–D38. https://doi.org/10.1093/nar/gki063

Berry, T. E., Osterrieder, S. K., Murray, D. C., Coghlan, M. L., Richardson, A. J., Grealy, A. K., … Bunce, M. (2017). DNA metabarcoding for diet analysis and biodiversity: A case study using the endangered Australian sea lion (Neophoca cinerea). *Ecology and Evolution*, *7*(14), 5435–5453.

Berry, T. E., Saunders, B. J., Coghlan, M. L., Stat, M., Jarman, S., Richardson, A. J., … Bunce,

M. (2019). Marine environmental DNA biomonitoring reveals seasonal patterns in biodiversity and identifies ecosystem responses to anomalous climatic events. *PLOS Genetics*, *15*(2), e1007943. Retrieved from https://doi.org/10.1371/journal.pgen.1007943

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., … Creer, S. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications*, *8*, 14087.

Boussarie, G., Bakker, J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J.-B., … Robbins, W. D. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, *4*(5), eaap9661.

Boyer, F., Mercier, C., Bonin, A., Taberlet, P., & Coissac, E. (2014). OBITools: a Unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*.

Brewer, D. T., Skewes, T. D., Lyne, V. D., Davies, C., Taranto, T. J., Fuller, M., … Andersen, J. (2009). Conservation values in Commonwealth waters of the Christmas and Cocos (Keeling) Island remote Australian territories.

Briggs, J. C. (1966). Oceanic islands, endemism, and marine paleotemperatures. *Systematic Zoology*, *15*(2), 153–163.

Budd, A. F., & Pandolfi, J. M. (2010). Evolutionary novelty is concentrated at the edge of coral species distributions. *Science*, *328*(5985), 1558–1561.

Bunce, P. (1988). *The Cocos (Keeling) Islands: Australian Atolls in the Indian Ocean*. Jacaranda Press for the Cocos (Keeling) Islands Council.

Bylemans, J., Furlan, E. M., Hardy, C. M., McGuffie, P., Lintermans, M., & Gleeson, D. M. (2017). An environmental DNA-based method for monitoring spawning activity: A case study, using the endangered Macquarie perch (Macquaria australasica). *Methods in Ecology and Evolution*, *8*(5), 646–655.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., … Gordon, J. I. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, *7*(5), 335.

Cardoso, P., Rigal, F., & Carvalho, J. C. (2015). BAT–Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods in Ecology and Evolution*, *6*(2), 232–236.

744 Clarke, K. R., & Gorley, R. N. (2015). *PRIMER v7: User Manual/Tutorial*. *There is no*
745 *corresponding record for this reference*. Plymouth, UK.
746 Cortés, J., Enochs, I. C., Sibaja-Cordero, J., Hernández, L., Alvarado, J. J., Breedy, O., …
747 Hermosillo, A. (2017). Marine biodiversity of Eastern Tropical Pacific coral reefs. In *Coral*
748 *Reefs of the Eastern Tropical Pacific* (pp. 203–250). Springer.
749 Cristescu, M. E., & Hebert, P. D. N. (2018). Uses and Misuses of Environmental DNA in
750 Biodiversity Science and Conservation. *Annual Review of Ecology, Evolution, and*
751 *Systematics*, *49*, 209–230.
752 De Caceres, M., & Jansen, F. (2016). Package 'indicspecies.'
753 Deagle, B. E., Gales, N. J., Evans, K., Jarman, S. N., Robinson, S., Trebilco, R., & Hindell, M. A.
754 (2007). Studying seabird diet through genetic analysis of faeces: a case study on macaroni
755 penguins (Eudyptes chrysolophus). *PLoS One*, *2*(9), e831.
756 DiBattista, J. D., Rocha, L. A., Hobbs, J. A., He, S., Priest, M. A., Sinclair-Taylor, T. H., …
757 Berumen, M. L. (2015). When biogeographical provinces collide: hybridization of reef
758 fishes at the crossroads of marine biogeographical provinces in the Arabian Sea. *Journal of*
759 *Biogeography*, *42*(9), 1601–1614.
760 Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST.
761 *Bioinformatics*, *26*(19), 2460–2461. Retrieved from
762 http://dx.doi.org/10.1093/bioinformatics/btq461
763 Evans, S. N., Konzewitsch, N., & Bellchambers, L. . (2016). *An update of the Department of*
764 *Fisheries, Western Australia, Invertebrate and Reef Health Research and Monitoring at*
765 *Cocos (Keeling) Islands*.
766 Fahay, M. P. (2007). Stomiiformes and Ateleopodiformes. In *Early Stages of Fishes in the*
767 *Western North Atlantic Ocean:(Davis Strait, Southern Greenland and Flemish Cap to Cape*
768 *Hatteras)* (pp. 252–305). Northwest Atlantic fisheries organization.
769 Fields, A. T., Abercrombie, D. L., Eng, R., Feldheim, K., & Chapman, D. D. (2015). A novel
770 mini-DNA barcoding assay to identify processed fins from internationally protected shark
771 species. *PloS One*, *10*(2), e0114844.
772 Forsström, T., & Vasemägi, A. (2016). Can environmental DNA (eDNA) be used for detection
773 and monitoring of introduced crab species in the Baltic Sea? *Marine Pollution Bulletin*,

774 *109*(1), 350–355.
775 Friedlander, A. M., Ballesteros, E., Beets, J., Berkenpas, E., Gaymer, C. F., Gorny, M., & Sala,
776 E. (2013). Effects of isolation and fishing on the marine ecosystems of Easter Island and
777 Salas y Gómez, Chile. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *23*(4),
778 515–531.
779 Froese, R., & Pauly, D. (2018). FishBase.
780 Graham, N. A. J., Wilson, S. K., Carr, P., Hoey, A. S., Jennings, S., & MacNeil, M. A. (2018).
781 Seabirds enhance coral reef productivity and functioning in the absence of invasive rats.
782 *Nature*, *559*(7713), 250.
783 Hender, J., McDonald, C. ., & Gilligan, J. . (2001). *Baseline surveys of the marine environments*
784 *and stock size estimates of marine resources of the south Cocos (Keeling) Atoll (0-15m),*
785 *eastern Indian Ocean. Draft Report to Fisheries Resources Research Fund.* Canberra.
786 Hobbs, J.-P.A., & McDonald, C. A. (2010). Increased seawater temperature and decreased
787 dissolved oxygen triggers fish kill at the Cocos (Keeling) Islands, Indian Ocean. *Journal of*
788 *Fish Biology*, *77*(6), 1219–1229. https://doi.org/10.1111/j.1095-8649.2010.02726.x
789 Hobbs, J. P. A., & Macrae, H. (2012). Unusual weather and trapped coral spawn lead to fish kill
790 at a remote coral atoll. *Coral Reefs*, *31*(4), 961. https://doi.org/10.1007/s00338-012-0918-8
791 Hobbs, Jean-Paul A, Frisch, A. J., Allen, G. R., & Van Herwerden, L. (2008). Marine hybrid
792 hotspot at Indo-Pacific biogeographic border. *Biology Letters*, *5*(2), 258–261.
793 Hobbs, Jean-Paul A, Newman, S. J., Mitsopoulos, G. E. A., Travers, M. J., Skepper, C. L.,
794 Gilligan, J. J., … Ayling, A. M. (2014). Fishes of the Cocos (Keeling) Islands: new records,
795 community composition and biogeographic significance. *RAFFLES BULLETIN OF*
796 *ZOOLOGY*, (30), 203–219.
797 Horton, T., Kroh, A., Ahyong, S., Bailly, N., Boyko, C. B., Brandão, S. N., … Zhao, Z. (2018).
798 World Register of Marine Species (WoRMS). WoRMS Editorial Board. Retrieved from
799 http://www.marinespecies.org
800 Houk, P., Benavente, D., Iguel, J., Johnson, S., & Okano, R. (2014). Coral reef disturbance and
801 recovery dynamics differ across gradients of localized stressors in the Mariana Islands. *PloS*
802 *One*, *9*(8), e105731.
803 Hourigan, T. F., & Reese, E. S. (1987). Mid-ocean isolation and the evolution of Hawaiian reef

804 fishes. *Trends in Ecology & Evolution*, *2*(7), 187–191.
805 Hourston, M. (2010). Review of the exploitation of marine resources of the Australian Indian
806 Ocean Territories: the implications of biogeographic isolation for tropical island fisheries.
807 *Fisheries Research Report*, (208). Retrieved from
808 http://www.fisheries.wa.gov.au/Documents/research_reports/frr208.pdf
809 Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic
810 data. *Genome Research*, *17*(3), 377–386. https://doi.org/10.1101/gr.5969107
811 IOC-UNESCO. (2018). *The United Nations Decade of Ocean Science for Sustainable*
812 *Development (2021-2030)*. Retrieved from
813 http://unesdoc.unesco.org/images/0026/002619/261962e.pdf
814 Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight-unseen" detection
815 of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150–157.
816 Jeunen, G., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., … Gemmell, N.
817 J. (2018). Environmental DNA (eDNA) metabarcoding reveals strong discrimination among
818 diverse marine habitats connected by water movement. *Molecular Ecology Resources*.
819 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A.
820 (2012). Geneious Basic: An integrated and extendable desktop software platform for the
821 organization and analysis of sequence data. *Bioinformatics*, *28*(12), 1647–1649. Retrieved
822 from http://dx.doi.org/10.1093/bioinformatics/bts199
823 Kelly, R. P., Closek, C. J., O'Donnell, J. L., Kralj, J. E., Shelton, A. O., & Samhouri, J. F. (2017).
824 Genetic and manual survey methods yield different and complementary views of an
825 ecosystem. *Frontiers in Marine Science*, *3*, 283.
826 Knowlton, N., & Jackson, J. B. C. (2008). Shifting baselines, local impacts, and global change on
827 coral reefs. *PLoS Biology*, *6*(2), e54.
828 Koziol, A., Stat, M., Simpson, T., Jarman, S., DiBattista, J. D., Harvey, E. S., … Bunce, M.
829 (2018). Environmental DNA metabarcoding studies are critically affected by substrate
830 selection. *Molecular Ecology Resources*.
831 Marsh, L. M. (1994). Echinoderms of the Cocos (Keeling) Islands.
832 McInerney, P. J., & Rees, G. N. (2018). More (or less?) bounce for the ounce: a comparison of
833 environmental DNA and classical approaches for bioassessment. *Marine and Freshwater*

Research, *69*(6), 992–996.

McKillup, S. C., & McKillup, R. V. (1996). The feeding behaviour of Thalamita crenata (Portunidae, Decapoda), a cannibalistic marine scavenger. *Marine & Freshwater Behaviour & Phy*, *28*(4), 255–267.

Mendoza, J. C. E., Lasley Jr., R. M., & Ng, P. K. L. (2014). New rock crab records (Crustacea: Brachyura: Xanthidae) from Christmas and Cocos (Keeling) Islands, Eastern Indian Ocean. Retrieved from https://s3.amazonaws.com/academia.edu.documents/37851400/S30_C29_274_300.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1525158965&Signature=Eq%2BXlguGUW3pZERzDP5DNlMkbHk%3D&response-content-disposition=inline%3Bfilename%3DNew_rock_crab_records_Crusta

Monteforte, M. (1987). The decapod reptantia and stomatopod crustaceans of a typical high island coral reef complex in French Polynesia (Tiahura, Moorea Island): zonation, community composition and trophic structure.

Morgan. (2000). DECAPOD CRUSTACEA OF CHRISTMAS ISLAND. *History*, *10*(16), 629–637.

Morgan, G. J. (1994). Decapod crustaceans of the Cocos (Keeling) Islands. *Atoll Research Bulletin*, (414). Retrieved from https://repository.si.edu/bitstream/handle/10088/5875/00414.pdf

Murakami, H., Yoon, S., Kasai, A., Minamoto, T., Yamamoto, S., Sakata, M. K., … Yamashita, Y. (2019). Dispersion and degradation of environmental DNA from caged fish in a marine environment. *Fisheries Science*, 1–11.

Nations, U. (2012). The future we want. In *Outcome document of the United Nations Conference on Sustainable Development*. UN New York.

Peyrot-Clausade, M. (1989). Crab cryptofauna (Brachyura and Anomura) of Tikehau, Tuamotu Archipelago, French Polynesia. *Coral Reefs*, *8*(3), 109–117.

Pochon, X., Bott, N. J., Smith, K. F., & Wood, S. A. (2013). Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. *PloS One*, *8*(9), e73935.

Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J.,

… Kelly, R. P. (2016). Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, *25*(2), 527–541.

Richards, Z. T., & Hobbs, J.-P. A. (2014). The status of hard coral diversity at Christmas Island and Cocos (Keeling) Islands. *Raffles Bulletin of Zoology*, *30*(30), 376–398.

Roberts, C. M., McClean, C. J., Veron, J. E. N., Hawkins, J. P., Allen, G. R., McAllister, D. E., … Wells, F. (2002). Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, *295*(5558), 1280–1284.

Schubert, M., Lindgreen, S., & Orlando, L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes*, *9*, 88. https://doi.org/10.1186/s13104-016-1900-2

Spalding, M. D., Ravilious, C., & Green, E. P. (2001). *World Atlas of Coral Reefs*. Univ of California Press.

Stat, M., Huggett, M. J., Bernasconi, R., DiBattista, J. D., Berry, T. E., Newman, S. J., … Bunce, M. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports*, *7*(1), 12240.

Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M., & Harvey, E. S. (2019). Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology*.

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

Tan, S. K., & Low, M. E. Y. (2014). Checklist of the Mollusca of Cocos (Keeling) / Christmas Island ecoregion. *RAFFLES BULLETIN OF ZOOLOGY*, *30*. Retrieved from https://www.researchgate.net/profile/Martyn_Low2/publication/270162575_Checklist_of_the_Mollusca_of_Cocos_Keeling_Christmas_Island_ecoregion/links/54a20a790cf256bf8baf7c57/Checklist-of-the-Mollusca-of-Cocos-Keeling-Christmas-Island-ecoregion.pdf

Taxonomy Decadel Plan Working Group. (2018). *Discovering Diversity: A decadel plan for taxonomy and biosystematics in Australia and New Zealand 2018-2028*. https://doi.org/10.1111/j.1469-8137.2010.03389.x

Team, Rs. (2015). RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL Http://Www. Rstudio. Com*, *42*.

894 Thomsen, P. F., Rask, P., Sigsgaard, E. E., Knudsen, W., Ankjær, O., & Willerslev, E. (2016).
895 Environmental DNA from Seawater Samples Correlate with Trawl Catches of Subarctic ,
896 Deepwater Fishes, 1–22. https://doi.org/10.1371/journal.pone.0165252
897 Tréguier, A., Paillisson, J., Dejean, T., Valentini, A., Schlaepfer, M. A., & Roussel, J. (2014).
898 Environmental DNA surveillance for invertebrate species: advantages and technical
899 limitations to detect invasive crayfish P rocambarus clarkii in freshwater ponds. *Journal of*
900 *Applied Ecology*, *51*(4), 871–879.
901 Tweedie, M. W. F. (1950). The fauna of the Cocos-Keeling Islands. Brachyura and Stomatopoda.
902 In *Bulletin of the Raffles Museum*. Citeseer.
903 Vörös, J., Márton, O., Schmidt, B. R., Gál, J. T., & Jelić, D. (2017). Surveying Europe's only
904 cave-dwelling chordate species (Proteus anguinus) using environmental DNA. *PloS One*,
905 *12*(1), e0170945.
906 Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. N. (2005). DNA barcoding
907 Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological*
908 *Sciences*, *360*(1462), 1847–1857.
909 Watanabe, H., Moku, M., Kawaguchi, K., Ishimaru, K., & Ohno, A. (1999). Diel vertical
910 migration of myctophid fishes (Family Myctophidae) in the transitional waters of the
911 western North Pacific. *Fisheries Oceanography*, *8*(2), 115–127.
912 [dataset] West et al; 2020; Sequencing data for eDNA metabarcoding survey reveals fine-scale
913 coral reef community variation across a remote, tropical island ecosystem; Dryad Digital
914 Repository; doi:10.5061/dryad.4qrfj6q6h
915 White, W. T., Corrigan, S., Yang, L. E. I., Henderson, A. C., Bazinet, A. L., Swofford, D. L., &
916 Naylor, G. J. P. (2018). Phylogeny of the manta and devilrays (Chondrichthyes : m obulidae
917), with an updated taxonomic arrangement for the family, (1987), 50–75.
918 Whittaker, R. J., & Fernández-Palacios, J. M. (2007). *Island Biogeography: Ecology, Evolution,*
919 *and Conservation*. Oxford University Press.
920 Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.
921 Williams, D. G. (1994). Marine habitats of the Cocos (Keeling) Islands. *Atoll Research Bulletin*,
922 (406).
923 Woodroffe, C. D., & Berry, P. F. (1994). Scientific studies in the Cocos (Keeling) islands: an

introduction. *Atoll Research Bulletin*.

Woodroffe, C. D., & Biribo, N. (2011). Atolls. In D. Hopley (Ed.), *Encyclopedia of Modern Coral Reefs: Structure, Form and Process* (pp. 51–71). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-2639-2_4

8. Data Accessibility

Demultiplexed (unfiltered) metabarcoding sequencing data and taxonomic presence/absence matrices are available for download on Dryad Digital Repository (doi:10.5061/dryad.4qrfj6q6h).

9. Authors' contributions

Funding acquisition, M.B, M.S, E.S.H, J.D.D, S.J.N; conceptualisation, M.B, M.S, E.S.H, J.D.D, S.J.N, K.M.W; fieldwork, K.M.W, E.S.H, C.L.S; laboratory processing and analysis, K.M.W, M.S, J.D.D, Z.T.R, M.J.T, M.B; writing and editing of manuscript, K.M.W, M.S, E.S.H, J.D.D, S.J.N, Z.T.R, M.J.T, M.B.

10. Figures

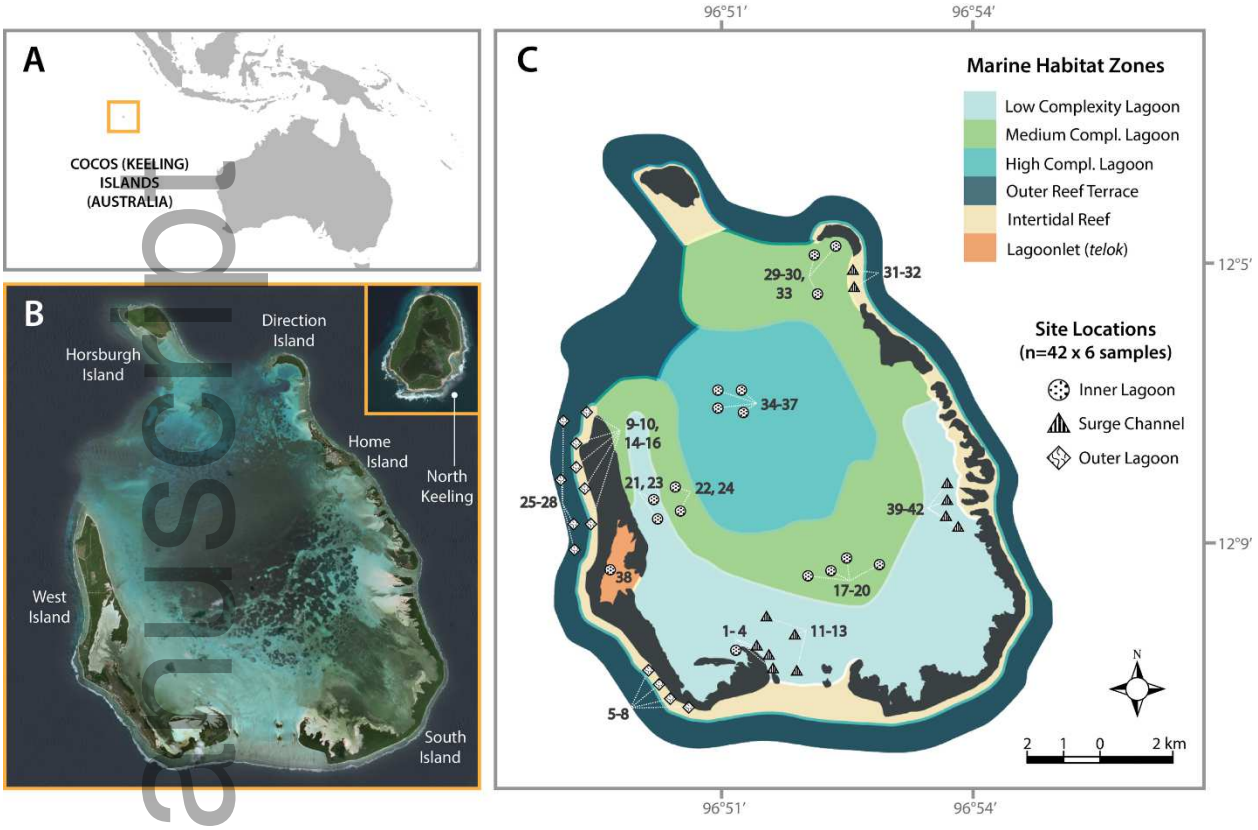

Figure 1. (A) Location of the Cocos (Keeling) Islands (CKI) in the eastern Indian Ocean. (B) Satellite imagery of the main CKI atoll with an inset of North Keeling atoll (not to scale). (C) Marine habitat zone map of the CKI, based on previous habitat strata divisions reported by Australian Government (2005), Hender et al. (2001) and Williams (1994). Site locations indicate where 6x1L water replicates were collected per site for eDNA surveying (for a total of 252 water samples across 42 sites) and were additionally classified as being situated either within the Inner Lagoon, Surge Channel or Outer Lagoon. Number of sites per marine habitat zone are as follows: Low Complexity n=13, Medium Complexity n=9, High Complexity n=4, Outer Reef Terrace n=4, Intertidal Reef n=11, and Lagoonlet (*telok*) n=1. Number of sites per lagoon zone as are follows: Inner Lagoon n=17, Surge Channel n=12, and Outer Lagoon n=13.

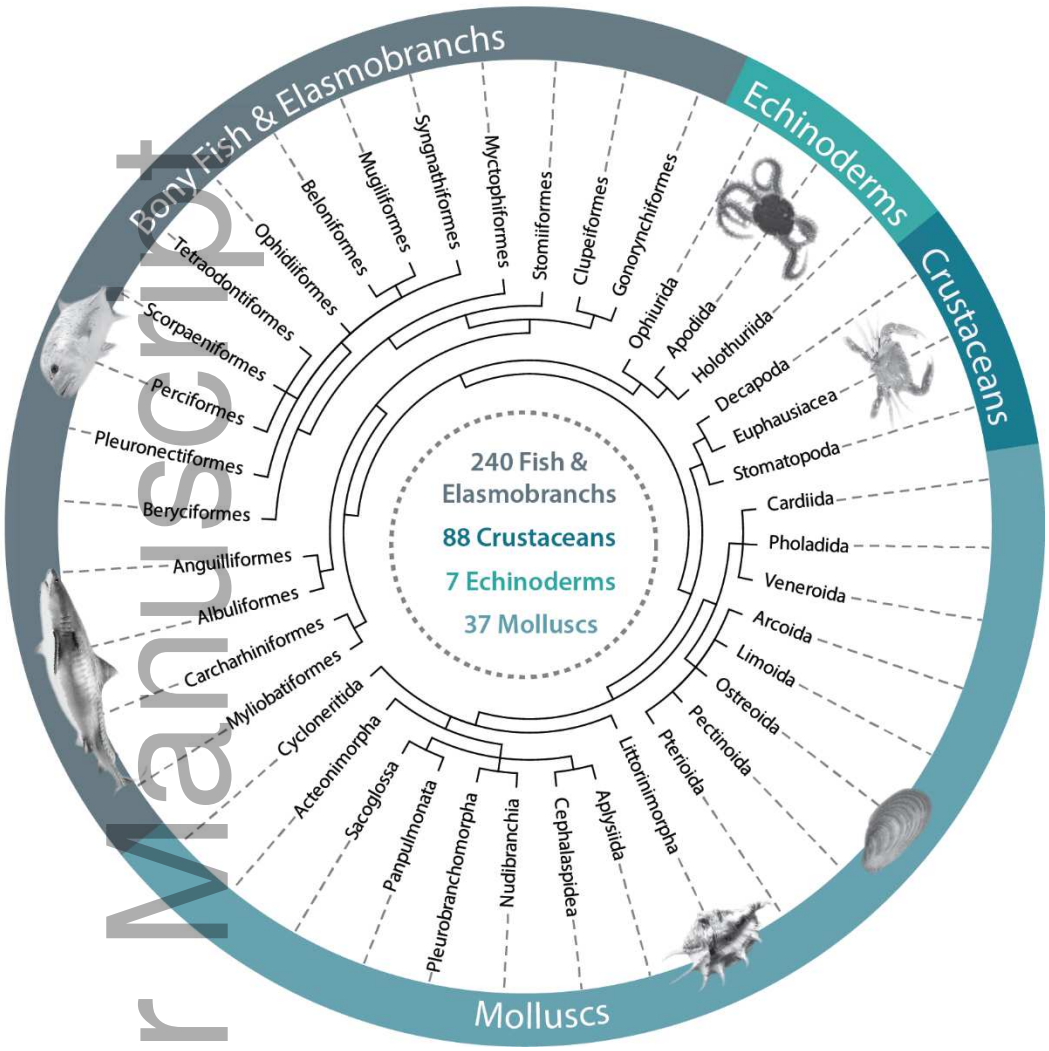

Figure 2. Order level dendrogram of bony fish, elasmobranch, crustacean, mollusc and echinoderm diversity detected at the CKI using eDNA metabarcoding. The number of taxa depicted in the central circle refers to the overall number of unique assignments in each taxonomic group (i.e. at a species, genus, family level etc).

968
969
970
971
972
973
974
975

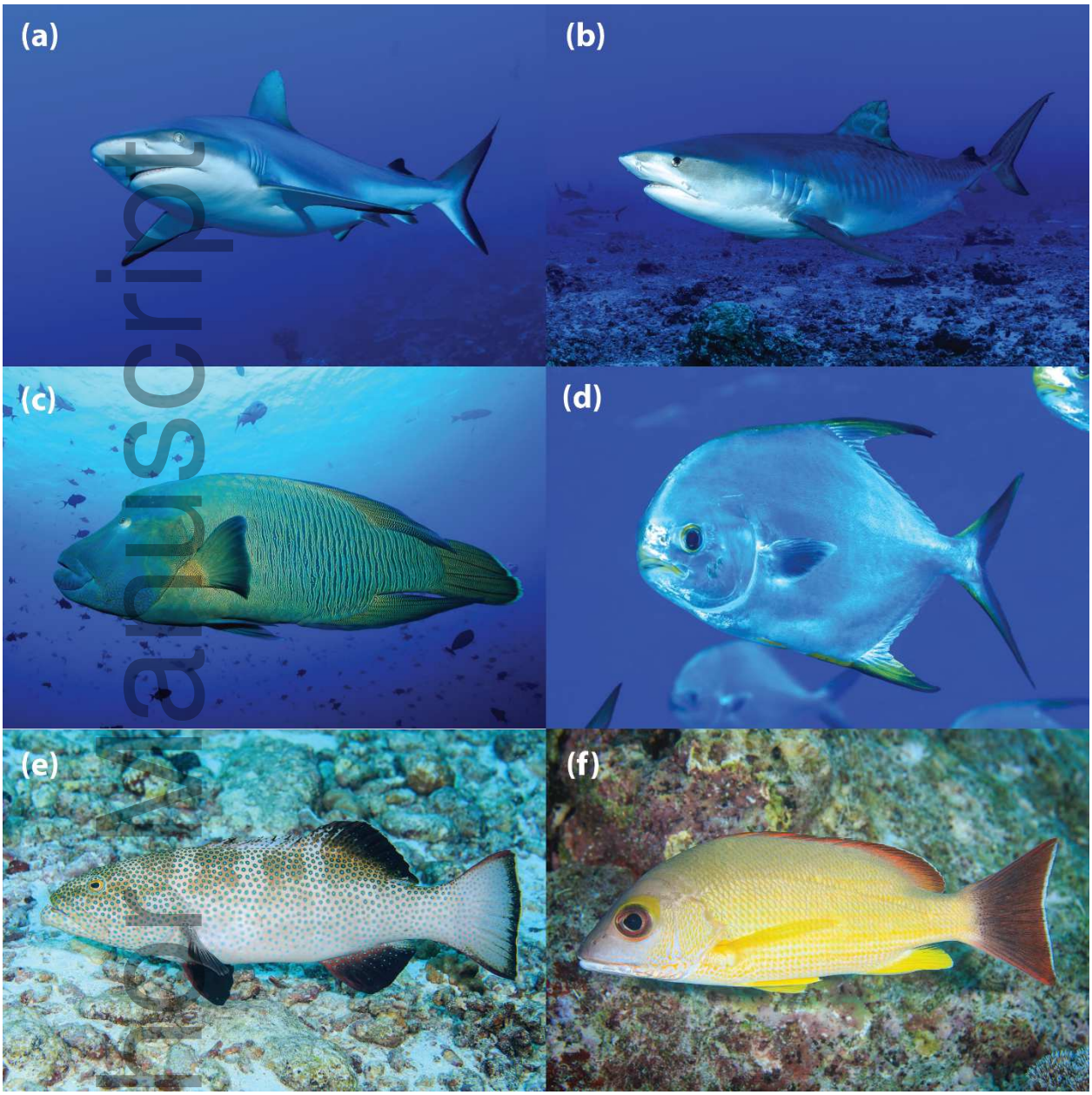

Figure 3. Taxa of interest detected by eDNA at the CKI. Shark species**: (a)** Grey reef shark (*Carcharhinus amblyrhynchos*) and **(b)** Tiger shark (*Galeocerdo cuvier*). Subsistence and recreational fishing species: **(c)** Humphead Maori wrasse (*Cheilinus undulatus*); **(d)** Snubnose dart (*Trachinotus blochii*); **(e)** Passionfruit coral trout (*Plectropomus areolatus*); **(f)** Blacktail snapper (*Lutjanus fulvus*). Photos: Tane Sinclair-Taylor.

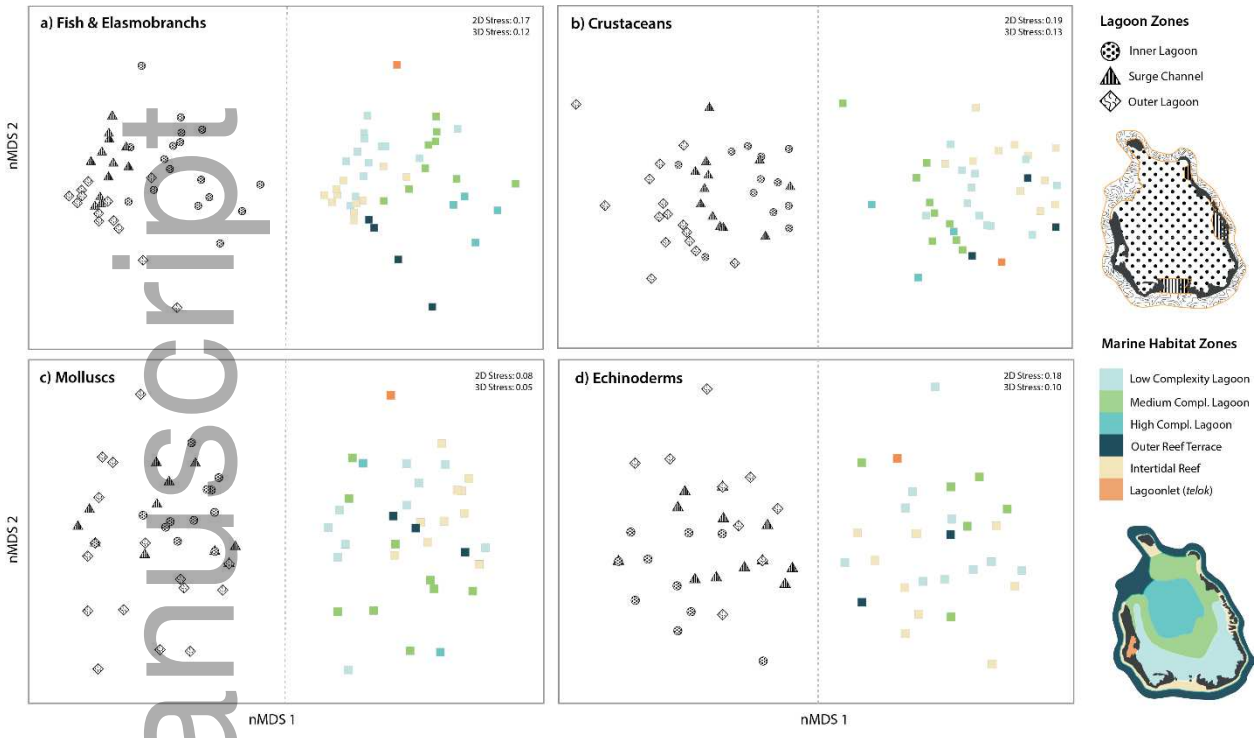

Figure 4. Non-metric multi-dimensional scaling (nMDS) ordination plots of a) bony fish and elasmobranch, b) crustaceans, c) molluscs and d) echinoderm community compositions. Each plot is divided into two parts with the left and right sides representing individual sites colour coded according to lagoon and marine habitat zone, respectively.

11. Tables

Table 1. PCR assay information for marine eDNA metabarcoding at CKI. Four primer sets: 16S Fish, 16S Crustacean, COI Elasmobranch and 18S Universal, corresponding to the following gene regions, respectively: mitochondrial 16S rDNA, mitochondrial COI, and nuclear 18S rDNA, were applied to all collected seawater samples. In the primer name, "F" refers to the forward primer and "R" to the reverse primer.

PCR Assay	Target taxa	Primer Name	Oligonucleotide Sequence (5'-3')	Target Length (bp)	Annealing temp (°C)	Primer Reference
16S Fish	Fish	16SF/D	GACCCTATGGAGCTTTAGAC	178-228	54	Berry et al., 2017
		16S2R-degenerate	CGCTGTTATCCCTADRGTAACT			Deagle et al., 2007
COI Elasmobranch	Sharks, Skates & Rays	FishF1-degenerate	ACCAACCACAAAGANATNGGCAC	110-241	52	**Original:** Ward et al., 2005 **Degenerate:** This study
		FishF2-degenerate	TCNACNAATCATAAAGATATCGGCAC			
		Shark COI-MINIR-degenerate	GATTATTACNAAAGCNTGGGC			**Original:** Fields et al., 2015 **Degenerate:**

						This study
16S Crustacean	Crustacean	Crust16S_F (short)	GGGACGATAAGACCCTATA	90-213	51	Berry et al., 2017
		Crust16S_R (short)	ATTACGCTGTTATCCCTAAAG			
18S Universal	Eukaryotes	18S_1F	GCCAGTAGTCATATGCTTGTCT	336-423	52	Pochon, Bott, Smith, & Wood, 2013
		18S_400R	GCCTGCTGCCTTCCTT			

1018

1019 **Table 2. Taxa of interest.** Marine habitat abbreviations are as follows: LCL, Low Complexity Lagoon; MCL, Medium Complexity
1020 Lagoon; HCL, High Complexity Lagoon; ORT, Outer Reef Terrace; IR, Intertidal Reef; L, Lagoonlet.

Common Name	Scientific Name	Importance	Total sites	Total reads	Marine Habitat Zone	Distribution
Bony Fish & Elasmobranchs						
Bonefish	*Albula*	Recreational fishing species	15	27,017	LCL, MCL, ORT, IR, L	Uncertain
White banded cod	*Anyperodon leucogrammicus*	Subsistence and recreational fishing species	1	157	HC	Indo-Pacific
Green humphead parrotfish	*Bolbometopon muricatum*	Subsistence fishing species	20	41,444	LCL, MCL, HCL, ORT, IR	Indo-Pacific
Lanternfish	*Bolinichthys*	New record at CKI. Genus comprised of epi- and mesopelagic species.	1	49	IR	Circumglobal
Goatsbeard brotula	*Brotula multibarbata*	Recreational fishing species	4	11,098	LCL, ORT, IR	Indo-Pacific
Giant trevally	*Caranx ignobilis*	Recreational fishing species	6	1,915	LCL, MCL, ORT, IR	Indo-Pacific
Bluefin trevally	*Caranx melampygus*	Recreational fishing species	1	177	MCL	Indo-Pacific
Grey reef shark	*Carcharhinus amblyrhynchos*	Requiem shark, not fished or targeted	12	120,664	LCL, MCL, HCL, IR	Indo-Pacific
Blacktip reef shark	*Carcharhinus melanopterus*	Requiem shark, not fished or targeted	19	49,427	LCL, MCL, HCL, ORT, IR, L	Indo-Pacific
Milkfish	*Chanos chanos*	Subsistence and recreational fishing species	18	108,492	LCL, MCL, ORT, IR, L	Indo-Pacific
Humphead Maori wrasse	*Cheilinus undulatus*	Subsistence fishing species	10	34,252	LCL, MCL, HCL, ORT, IR	Indo-Pacific
Rainbow runner	*Elagatis bipinnulata*	Recreational fishing species	1	132	LCL	Circumglobal
Squaretail mullet	*Ellochelon vaigiensis*	Subsistence fishing species	6	75,038	LCL, IR, L	Indo-Pacific
Grouper	*Epinephelus*	Subsistence and recreational fishing species	38	131,075	LCL, MCL, HCL, IR	Indo-Pacific
Tiger shark	*Galeocerdo cuvier*	Requiem shark, not fished or targeted	1	111	HCL	Circumglobal

Strongspine silver-biddy	*Gerres longirostris*	Subsistence fishing species	9	1527	LCL, IR, L	Indo-Pacific
Luminous lanternfish	*Lampadena luminosa*	New record at CKI. Bathypelagic species.	1	739	MCL	Circumglobal
Lanternfish	*Lampanyctus*	New record at CKI. Genus comprised of epi-, meso- and bathypelagic species.	1	40	MCL	Circumglobal
Emperors	*Lethrinus*	Subsistence fishing species. Historical concerns regarding overexploitation (Hender et al. 2001).	31	68,884	LCL, MCL, HCL, IR, L	Indo-Pacific
Two-spot red snapper	*Lutjanus bohar*	Subsistence and recreational fishing species	12	63,742	LCL, MCL, HCL, ORT, IR	Indo-Pacific
Blacktail snapper	*Lutjanus fulvus*	Subsistence and recreational fishing species	24	84,016	LCL, MCL, ORT, IR, L	Indo-Pacific
One-spot snapper	*Lutjanus monostigma*	Subsistence fishing species	1	940	HCL	Indo-Pacific
Giant oceanic manta ray	*Manta birostris*	Not fished or targeted	11	695	LCL, MCL, HCL, ORT, IR	Circumglobal
Devil rays	*Mobula*	New record at CKI, not fished or targeted	2	1,628	IR	Circumglobal
Lanternfish & blackchins	Myctophiformes	New record at CKI. Order comprised of bathypelagic species. Requires reference sequences to resolve assignments.	4	4,208	LCL, MCL, IR	Circumglobal
Hornlip mullet	*Oedalechilus labiosus*	New mullet record at CKI	1	218	LCL	Indo-Pacific
Passionfruit coral trout	*Plectropomus areolatus*	Historical concerns regarding overexploitation (Fletcher & Santoro 2009).	5	4,563	MCL, HCL	Indo-Pacific
Goldband snapper	*Pristipomoides multidens*	New snapper record at CKI	1	313	IR	Indo-Pacific
Dragonfishes & relatives	Stomiiformes	New record at CKI. Order comprised of epi-, meso- and bathypelagic species. Requires reference sequences to resolve assignments.	1	252	IR	Circumglobal
Snubnose dart	*Trachinotus blochii*	Recreational fishing species	5	3,764	LCL, MCL, L	Indo-Pacific
Whitetip reef shark	*Triaenodon obesus*	Requiem shark, not fished or targeted	9	29,617	LCL, MCL, HCL, IR	Indo-Pacific

Crustaceans						
Fourspine rock lobster	*Panulirus penicillatus*	Subsistence fishing species	1	5,144	LCL	Indo-Pacific
Prawns	Penaeidae	New record at CKI	2	51,845	LCL, MCL	Circumglobal
Mangrove swimming crab	*Thalamita crenata*	Subsistence fishing species	2	9,544	LCL, L	Indo-Pacific
Molluscs						
Spider conch	*Lambis lambis*	Subsistence fishing species	1	19	MCL	Indo-Pacific
Oysters	Ostreidae	Subsistence fishing species	2	54	MCL, IR	Circumglobal
Pearl oysters	*Pinctada*	Potential commercial interest	2	53	LCL	Circumglobal
Echinoderms						
Sea cucumbers	Holothuriidae	Ecologically important species	13	1,454	LCL, MCL, IR	Circumglobal

1021

1022

1023

Table 3. Summary table of the distance based linear model (DistLM) analyses. These were constructed using a sequential step-wise selection procedure and adjusted R^2 criterion. Under the spatial predictor variables, 'Lagoon' refers to the lagoon zones, whilst 'Habitat' refers to the marine habitat zones. Significant codes are as follows: 0 < 0.001 '***', 0.001 < 0.01 '**', 0.01 < 0.05 '*'. The variables highlighted in bold are significant (P<0.05). Full DistLM results, including marginal tests, are provided in SI Tables 9-12.

Response	Predictor	Adj R^2	Pseudo-*F*	Proportion	Cumulative Proportion	*P*	
Bony Fish & Elasmobranchs	**Habitat**	0.187	2.885	0.286	0.286	**0.000**	***
	Lagoon	0.210	1.531	0.059	0.345	**0.007**	**
	Longitude	0.233	1.995	0.037	0.382	**0.000**	***
	Latitude	0.247	1.640	0.030	0.412	**0.009**	**
	Depth	0.258	1.476	0.027	0.439	**0.047**	*
Crustaceans	**Habitat**	0.097	1.817	0.216	0.216	**0.000**	***
	Lagoon	0.119	1.407	0.065	0.281	**0.042**	*
	Longitude	0.121	1.073	0.024	0.306	0.358	
Molluscs	**Habitat**	0.076	1.661	0.192	0.192	**0.001**	***
	Latitude	0.111	2.408	0.053	0.245	**0.006**	**
	Longitude	0.139	2.050	0.044	0.289	**0.020**	*
	Lagoon	0.185	1.945	0.079	0.369	**0.008**	**
Echinoderms	**Lagoon**	0.129	3.453	0.182	0.182	**0.000**	***
	Latitude	0.176	2.754	0.069	0.251	**0.020**	*
	Depth	0.189	1.478	0.036	0.287	0.210	
	Longitude	0.200	1.418	0.034	0.322	0.232	
	Habitat	0.235	1.313	0.122	0.443	0.186	

Table 4. Overview of bony fish, elasmobranch and crustacean taxonomic surveys conducted at the CKI, ordered by decreasing year. The results from this survey are in bold. The percentage of taxa detected over total diversity is calculated based on the updated diversity count of 644 taxa for bony fish and elasmobranchs, and 236 taxa for crustaceans (221 Decapoda, 14 Stomatopoda and 1 Euphausiacea). New records refer to either new occurrence records at the CKI, or new species descriptions. Harvey et al. (unpublished) recorded species list for BRUV surveys in 2016 and 2017 are provided in SI Tables 30 and 31.

Year	Trips	Method	Taxa detected	Taxa detected/ total diversity	New records	Reference
Bony Fish & Elasmobranchs						
2017	**1**	**eDNA metabarcoding**	**240**	**37.3%**	**18 new occurrence**	**This study**
2017	1	Baited remote underwater video (BRUV)	245	38.0%	21 new occurrence	Harvey et al. (unpublished)
2016	1	Baited remote underwater video (BRUV)	267	41.5%	3 new occurrence	Harvey et al. (unpublished)
2014	11	Underwater visual observations	N/A	N/A	67 new occurrence	Hobbs,

–2001		and line-fishing				Newman, et al., 2014
2004	1	Underwater visual observations (large >30 TL cm fish only)	58	9.0%	0	Bennett et al., 2018
2001	1	Underwater visual observations (commercial and recreational target species only)	74	11.5%	0	Hender et al., 2001
1989	1	Rotenone, spear and dipnets	448	69.6%	N/A	Allen & Smith-Vaniz, 1994
1973	1	Rotenone, explosives, quinaldine, angling and spearing	425	65.9%	12 new species	Allen & Smith-Vaniz, 1994
Crustaceans						
2017	**1**	**eDNA metabarcoding**	**88**	**37.3%**	**25 new occurrence**	**This study**
1994	1	Above and underwater observations (Decapoda only)	198	89.6%	78 new occurrence	Morgan, 1994

1061