Online Journal of Biological Sciences

Original Research Paper

The Identification and Quantification of Marine Fish from Environmental DNA in Sulawesi Waters (Makassar Strait, Flores Sea and Bone Bay), Indonesia

[1]Nita Rukminasari, [1]Andi Aliah Hidayani, [1]Wilma Joanna Carolina Moka, [4]Nur Indah Sari Arbit, [2]Sapto Andriyono and [3]Andi Parenrengi

[1]Department of Fisheries, Faculty of Marine Science and Fisheries, Hasanuddin University, Indonesia
[2]Department of Fisheries, Faculty of Fisheries and Marine, Airlangga University, Indonesia
[3]Research Centre for Fisheries, National Research and Innovation Agency, Indonesia
[4]Faculty of Animal Husbandry, West Sulawesi University, Indonesia

Article history
Received: 19-03-2024
Revised: 16-04-2024
Accepted: 02-05-2024

Corresponding Author:
NIta Rukminasari
Department of Fisheries, Faculty of Marine Science and Fisheries, Hasanuddin University, Indonesia
Email: nita.r@unhas.ac.id

Abstract: The emergence of environmental DNA (eDNA) represents a recent methodological breakthrough for evaluating the presence of aquatic vertebrate species. This approach offers a relatively simple method with significant implications for conservation biology. Our study aim was to augment our understanding of marine fish biodiversity in Sulawesi waters. We employed eDNA metabarcoding to investigate fish biodiversity within Sulawesi waters, specifically focusing on the Makassar Strait, Bone Bay, and Flores Sea. The eDNA was extracted from 4-liter water samples obtained from the surface (0-1m depth) and the water column (15 m depth) at five distinct sites across the study area. Methodological reliability was evaluated using a primer set (MiFish-U) to estimate fish diversity in Sulawesi waters. Analysis of nine water samples collected from Sulawesi waters revealed the presence of 36 marine fish taxa identified to the species level, representing 18 families across 13 orders. The majority of these taxa were associated with reef habitats, indicating the prevalence of coral reef ecosystems in the region. Among the surveyed regions, Bone Bay exhibited the highest species richness with 27 taxa, followed by the Makassar Strait with 14 taxa and the Flores Sea with 12 taxa. This investigation facilitated the estimation of fish diversity utilizing eDNA metabarcoding, thereby furnishing valuable baseline data.

Keywords: Diversity, Environmental, Marine, Metabarcoding, Sulawesi

Introduction

Understanding the geographical range of a species is fundamental for comprehending ecological patterns and assessing extinction risk, thus playing a crucial role in population-level conservation efforts (Begon and Townsend, 2021). However, obtaining precise distribution estimates is often challenging due to complex microhabitat structures and vegetation, particularly in aquatic environments. Environmental DNA (eDNA) has become a valuable tool in recent years for mapping the geographic distributions of aquatic vertebrate species (Ficetola *et al*., 2008; Goldberg *et al*., 2011; Jerde *et al*., 2011; Minamoto *et al*., 2012; Valentini *et al*., 2009). The capability to identify short DNA fragments from water samples enhances survey accuracy while reducing costs, thereby facilitating the detection of both rare and invasive species (Valentini *et al*., 2009).

Indonesia, boasting over 17,000 islands, holds the esteemed title of the world's largest archipelagic nation. Situated within the tropical belt, it encompasses a diverse array of ecosystems and landscapes, ranging from deep seas to lowland and mountainous forests and even snowy peaks. Teeming with life, Indonesia is a haven for an astounding variety of living things, boasting nearly 17% of the world's richness of species. From 270 distinct kinds of mammals to 386 feathered creatures, 328 reptilian forms, 204 amphibious wonders, and a staggering 280 species of fish, Indonesia thrives as a hotspot for biodiversity. Additionally, it encompasses 10% of all flowering plants, 12% of mammals, 25% of reptiles, and vast unexplored reservoirs of microbial and genetic resources (Cleary and DeVantier, 2011). Furthermore, many living organisms exhibit endemism within specific regions of Indonesia (Hakim, 2017).

Situated at the heart of the global marine biodiversity triangle, Indonesia boasts some of the most diverse marine environments worldwide. The Wallace line demarcates distinct patterns of faunal diversity, morphology, and distribution. Sulawesi waters, including the Makassar Strait, Bone Bay, and Flores Sea, lie within the coral triangle, characterized by high biodiversity, and the Wallacea region, known for its high level of endemism (Ambo-Rappe and Moore, 2019). However, despite its ecological significance, Research investigating the efficiency of environmental DNA (eDNA) for identifying and quantifying marine fish species in this area is lacking.

Research indicates that eDNA metabarcoding surpasses traditional fish sampling methods in its effectiveness for studying fish distribution and diversity (Fujii *et al*., 2019; Sard *et al*., 2019). Moreover, it offers the advantage of surveying a larger number of locations in less time compared to conventional approaches, thus facilitating broader geographical coverage. Additionally, eDNA methods are non-invasive and obviate the need for euthanizing organisms (Kume *et al*., 2021). Detecting fish eDNA in water is indicative of the fish species present in that water body (Thomsen *et al*., 2012). While the relationship between fish presence and their DNA signal in water may change depending on location and time, eDNA offers promise for monitoring fish populations across different situations (Jerde *et al*., 2019). Research exploring the diversity of freshwater and marine fish in different settings has presented convincing proof endorsing eDNA metabarcoding as a strong method for monitoring aquatic ecosystems, aiding in their conservation and management (Andruszkiewicz *et al*., 2017; Cilleros *et al*., 2019; Evans *et al*., 2016; Shaw *et al*., 2016; Stat *et al*., 2019; Valentini *et al*., 2016).

The primary aim of our research was to augment our understanding of marine fish biodiversity in Sulawesi waters. We initiated this endeavor by conducting eDNA metabarcoding analysis on water samples collected from three distinct regions within Sulawesi waters, utilizing the MiFish-U primer set. Subsequently, we scrutinized the Operational Taxonomic Units (OTUs) generated through the MiFish-U pipeline to assess their efficacy in identifying marine fish species present in Sulawesi waters. Finally, leveraging the eDNA metabarcoding results, we computed diversity indices, including Shannon and Simpson species richness, to evaluate fish biodiversity across the designated areas within Sulawesi waters.

Materials and Methods

Study Sites

This research was carried out in three regions within Sulawesi waters (Makassar Strait, Flores Sea, and Bone Bay). We determined five representative sampling sites: Two each in the Makassar Strait and Bone Bay and one in the Flores Sea. The Makassar strait sampling sites were in Barru and Pangkep waters, the Flores Sea sampling site was in Bantaeng waters and the Bone Bay sampling sites were in Sinjai and Bone waters (Fig. 1).

eDNA Sampling

At designated stations, 4-liter water samples were gathered for eDNA metabarcoding. Samples were taken near the surface (0-1 mD) and at a depth of 15 m at each of the five stations, totaling 10 samples. Each 4-liter sample was placed in a fresh sterile polypropylene/HDPE container, properly labeled, and then cautiously placed in a coolbox due to the sensitivity of eDNA samples. Afterward, the samples were frozen at -20°C until eDNA extraction was performed.

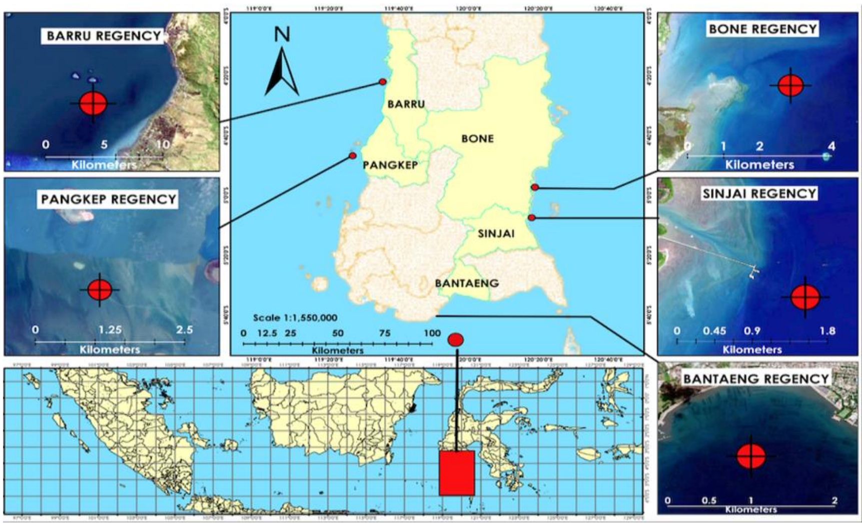

Fig. 1: Map showing the eDNA sample collection sites (red dots) in Sulawesi waters

DNA Extraction

Each 4 L water sample underwent filtration using 47 mm diameter filter paper. Employing a multi-filter technique, we aimed to trap and retain the DNA present in each water sample, replacing the filters after filtering approximately 2 L, thus resulting in two filters per sample. DNA extraction from each filter was performed following the CTAB method as detailed before. About 3 mL of CTAB buffer was poured onto each filter paper and the mixture was left to incubate in a water bath at 60°C for 3 h, with intermittent vortexing every half-hour. Phase separation was achieved by introducing 1 mL of chloroform, followed by vortexing for 30 sec and subsequent centrifugation at 12,000 rpm for 15 min. The resulting aqueous layer was carefully transferred to a fresh sterile tube and an equal volume of cold ethanol was added to precipitate the DNA, forming DNA pellets through centrifugation for 15 min. After two washes with 70% ethanol, the DNA pellet was dissolved in sterile Molecular Biology class water (Sigma Aldrich, USA) and stored at-20°C. The quality of DNA extracted from each station was evaluated using agarose gel electrophoresis and spectrophotometry (Thermo scientific™ NanoDrop™ one microvolume UV-Vis spectrophotometer) for preparing the DNA for sequencing.

Preparing Libraries and Conducting Next-Generation Sequencing

Library preparation involved two stages of PCR. Both the PCR products from the first and second stages were purified using AMPure XP beads before proceeding to the subsequent step. The first PCR was employed to amplify the target region of 12S rRNA mitochondrial DNA (mtDNA), a molecular marker recognized for its utility in identifying fish and other marine vertebrates (Suarez-Bregua *et al*., 2022), using the MiFish-U primer set (Miya *et al*., 2015). Each PCR reaction consisted of 12 Kapa HotStart HiFi 2× ReadyMix DNA polymerase, 1 μL of each 10 nM primer (F and R), 8 μL of dd H_2O, and 2 μL of DNA template. The DNA amplification PCR protocol involved: (1) Initial denaturation of the template DNA at 95°C for 5 min; 35 cycles of (2) Denaturation at 98°C for 30 sec, (3) Annealing at 65°C for 30 sec, (4) Primary extension at 72°C for 30 sec and (5) Final extension at 72°C for 5 min. Contamination was assessed using the 96 universal peqStAR PCR machine with negative controls (blank template). PCR product integrity was assessed via electrophoresis on a 2% agarose gel (100 mL TAE buffer and 2 g agarose). Each agarose well was loaded with a 3 μL aliquot of PCR product and a 100 bp DNA ladder. Electrophoresis was conducted at 50 volts for 60 min and the outcomes were observed using UV fluorescence in an alpha imager mini gel documentation system.

PCR products that successfully passed the electrophoresis quality assessment underwent a secondary PCR step for indexing purposes. Library markers, such as the IDT double index and Illumina sequencing adapters for Illumina-Nextera DNA unique dual index, set B, were introduced to the target amplicons during this second PCR phase. Each reaction comprised 12.5 μL of 2× ReadyMix and 2 μL of PCR product. The PCR cycle involved an initial denaturation at 95°C for 3 min, followed by 9 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, extension at 72°C for 30 sec, and a final extension at 72°C for 5 min. The purified indexed amplicon libraries were subsequently subjected to sequencing using an Illumina iSeq100 platform.

Bioinformatics and Data Analysis

We processed the raw sequence data output from the iSeq platform to generate FASTQ files, followed by data preprocessing using the fastp software developed by Chen *et al*. (2018). Upon applying error correction to the overlapping region between the paired-end reads, we assembled the paired-end dataset using FLASH 1.2.11, an assembly tool introduced by Magoč and Salzberg (2011). Sequences shorter than 100 base pairs or longer than 200 base pairs were filtered out from the assembled dataset. Subsequently, we classified the resulting sequences into taxonomic units with the tools of the CD-HIT with a similarity threshold of ≥97%, employing the mean relationship algorithm in search (Li *et al*., 2012). To assign representative OTU sequences to reference taxa, BLASTn analysis (version 2.9.0) was conducted to identify the closest matches in terms of sequence similarity (Zhang *et al*., 2000). To focus specifically on the eukaryotic community, bacterial OTUs were removed.

To quantify marine fish biodiversity based on the eDNA OTU dataset generated by the bioinformatics pipeline, we utilized assigned OTU taxa and their relative abundance (number of reads per OTU). Biodiversity indices, including Shannon and Simpson species richness, were estimated using primer-E version 5. The fish OTU datasets were consolidated and processed to create a pairwise distance matrix employing bray-Curtis dissimilarity. This facilitated the comparison of fish community composition across distinct regions. Non-Metric Multidimensional Scaling (NDMS) was employed to visually illustrate the dissimilarity between samples, while Analysis of Similarity (ANOSIM) was utilized to assess whether statistically significant differences in species composition existed among the different regions.

Results

eDNA Metabarcoding

The metabarcoding process produced sequence data from 9 of the 10 samples collected at the five sites in three

Sulawesi water bodies. We obtained 10,165 valid reads (Table 1) from the MiFish-U eDNA metabarcoding pipeline (Miya *et al*., 2015). These reads were assigned to 36 species-level marine fish OTUs with 98.22-100% identity with reference sequences.

The species assigned encompassed 36 genera from 18 families and 13 orders. The identity and relative abundance of species detected from surface water (0-1 m) and water column (15 mD) samples varied between sites (Fig. 1).

For both sites in Bone Bay, more marine fish species were assigned from surface eDNA samples (0-1 mD) than from water column (15 mD) samples. Only the water column sample provided valid data for the Flores Sea, while all Makassar Strait samples had similar and relatively low numbers of species (Fig. 1). At the family level, the *Pomacanthidae* had the highest percentage of reads at all sites, while 9 families (*Acanthuridae*, *Ambassidae*, *Apogonidae*, *Belonidae*, *Engraulidae*, *Gobiidae*, *Mugilidae*, *Sciaenidae* and *Zenarchopteridea*) had very few reads, each accounting for 2.8% of the total (Figs. 2-4).

Overall, at the family level, the *Pomacanthidae* family appeared to be dominant (13.9%), followed by the *Leiognathidae* and *Mullidae* (11.1%) and the *Balistidae* and *Dorosomatidae* (8.3%). However, there were variations in the proportions of species within these families between areas. In Bone Bay (Bone and Sinjai sites), the *Caesionidae* accounted for 7.4%, while in the Makassar Strait (Barru and Pangkep sites) the *Zenarchopteridae* accounted for 7.1%, and in the Flores Sea, the *Ambassidae* accounted for 8.3% of reads. Based on reads obtained from eDNA metabarcoding, the most abundant orders varied with depth and between sampling sites (Fig. 5). *Clupeiformes* was the most abundant fish order in surface water (0-1 mD) while Perciformes was the most abundant fish order in the water column (15 mD). In Bone Bay and the Flores Sea, Perciformes was the most abundant order, while *Clupeiformes* was the most abundant order in the Makassar Strait.

Species Assemblages and Composition as Revealed by the eDNA Dataset

The marine fish species composition varied substantially between locations (Fig. 6). The yellowtail fusilier *Caesio cuning* was a dominant species in Bone Bay (Bone and Sinjai sites), while the spoon-fin garfish *Zenarchopterus dispar* was dominant in the Makassar Strait (Pangkep and Barru sites) and *Ambassis urotaenia* was the dominant species in the Flores Sea (Bantaeng site). The highest number of reads was *Zenarchopterus dispar* accounting for 1432 and species and *Grammatorcynus bilineatus* (found at the Bantaeng site) had the least (14 reads). A phylogenetic tree at species and family level was constructed from the eDNA OTU sequences (Fig. 7).

The nMDS plot based on species assemblages in the eDNA OTU dataset showed significant between-site differences as well as segregation between the sea areas (Fig. 8). ANOSIM SIMPER results indicate that Bone Bay and Flores Sea are quite similar in terms of fish species composition, with a dissimilarity percentage of 52.09%.

This result is likely linked to the high abundance of *Caesio cuning* in Bone Bay and the Flores Sea. Conversely, species composition differed significantly between the Makassar Strait and Bone Bay, with a dissimilarity percentage of 87.56%. The three species contributing most to the differences between these sea areas were *Caesio cuning, Encrasicholina punctifer, and Grammatorcynus bilineautus.*

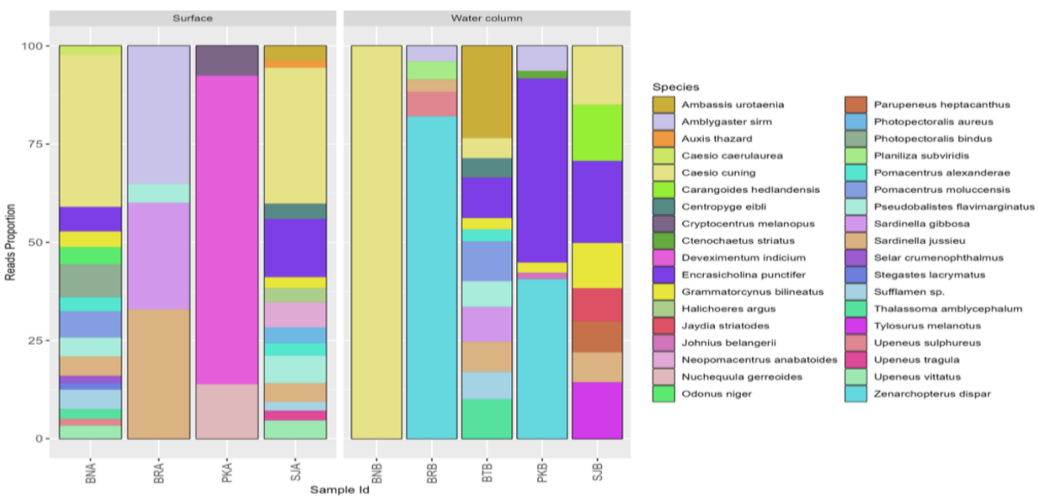

Fig. 2: Proportion of marine fish OTU reads assigned to each species-level taxon by the MiFish eDNA metabarcoding pipeline by site and depth. Site codes: BN = Bone, SJ = Sinjai (Bone Bay), BR = Barru, PK = Pangkep (Makassar Strait), BT = Bantaeng

Table 1: List of marine fish species identified through the eDNA metabarcoding method, along with the corresponding count of reads and their relative proportions

Order	Family	Species	Distribution	Identity (%)	Total reads	Read proportion (%)
Acanthuriformes	*Acanthuridae*	*Ctenochaetus striatus*	Indo-pacific	100.00	15	0.1
Acanthuriformes	*Leiognathidae*	*Deveximentum indicium*	Western pacific	100.00	312	3.1
Acanthuriformes	*Leiognathidae*	*Nuchequula gerreoides*	Indo-west pacific	98.82.00	55	0.5
Acanthuriformes	*Leiognathidae*	*Photopectoralis aureus*	Western PACIFIC	98.84.00	53	0.5
Acanthuriformes	*Leiognathidae*	*Photopectoralis bindus*	Indo-west pacific	100.00	145	1.4
Acanthuriformes	*Sciaenidae*	*Johnius belangerii*	Indo-west pacific	100.00	13	0.1
Beloniformes	*Belonidae*	*Tylosurus melanotus*	Indo-pasific	100.00	401	3.9
Beloniformes	*Zenarchopteridae*	*Zenarchopterus dispar*	Indo-pacific	100.00	1751	17.2
Carangiformes	*Carangidae*	*Carangoides hedlandensis*	Indo-west Pacific	98.82-99.41	397	3.9
Carangiformes	*Carangidae*	*Selar crumenophthalmus*	Pasific and Atlantic	99.41.00	30	0.3
Clupeiformes	*Dorosomatidae*	*Amblygaster sirm*	Indo-west pacific	100.00	280	2.8
Clupeiformes	*Dorosomatidae*	*Sardinella gibbosa*	Indo-west pacific	100.00	168	1.7
Clupeiformes	*Engraulidae*	*Encrasicholina punctifer*	Indo-PACIFIC	98.22-100	1306	12.8
Clupeiformes	*Dorosomatidae*	*Sardinella jussieu*	Western Indian ocean	100.00	605	6.0
Gobiiformes	*Gobiidae*	*Cryptocentrus melanopus*	Western PACIFIC	100.00	30	0.3
Kurtiformes	*Apogonidae*	*Jaydia striatodes*	Indo-west pacific	98.25.00	234	2.3
Labriformes	*Labridae*	*Halichoeres argus*	Indo-west pacific	100.00	48	0.5
Labriformes	*Labridae*	*Thalassoma amblycephalum*	Indo-Pacific	100.00	91	0.9
Mugiliformes	*Mugilidae*	*Planiliza subviridis*	Indo-pacific	100.00	81	0.8
Ovalentaria incertae sedis	*Ambassidae*	*Ambassis urotaenia*	Indo-pasific	98.21-99.4	164	1.6
Ovalentaria incertae sedis	*Pomacentridae*	*Stegastes lacrymatus*	Indo-pacific	100.00	31	0.3
Perciformes	*Caesionidae*	*Caesio caerulaurea*	Indo-west pacific	100.00	35	0.3
Perciformes	*Caesionidae*	*Caesio cuning*	Indo-west pacific	99.41-100	2019	19.9
Perciformes	*Pomacanthidae*	*Neopomacentrus anabatoides*	Western central pacific	98.81.00	85	0.8
Perciformes	*Pomacanthidae*	*Pomacentrus alexanderae*	Western pacific	100.00	120	1.2
Perciformes	*Pomacanthidae*	*Pomacentrus moluccensis*	Western pacific	100.00	165	1.6
Perciformes	*Pomacanthidae*	*Centropyge eibli*	Eastern Indian ocean	100.00	76	0.7
Scombriformes	*Scombridae*	*Grammatorcynus bilineatus*	Indian and pasific	100.00	462	4.5
Scombriformes	*Scombridae*	*Auxis thazard*	Atlantic, mediterranean, Indian and pacific	100.00	24	0.2
Syngnathiformes	*Mullidae*	*Parupeneus heptacanthus*	Indo-West pacific	100.00	221	2.2
Syngnathiformes	*Mullidae*	*Upeneus sulphureus*	Indo-West pacific	100.00	139	1.4
Syngnathiformes	*Mullidae*	*Upeneus tragula*	Eastern Indian ocean to western pacific	100.00	33	0.3
Syngnathiformes	*Mullidae*	*Upeneus vittatus*	Indo-pacific	98.25.00	120	1.2
Tetraodontiformes	*Balistidae*	*Odonus niger*	Indo-pacific	100.00	76	0.7
Tetraodontiformes	*Balistidae*	*Pseudobalistes flavimarginatus*	Indo-pacific	97.08-100	230	2.3
Tetraodontiformes	*Balistidae*	*Sufflamen* sp.	Western Indian Ocean	100.00	150	1.5

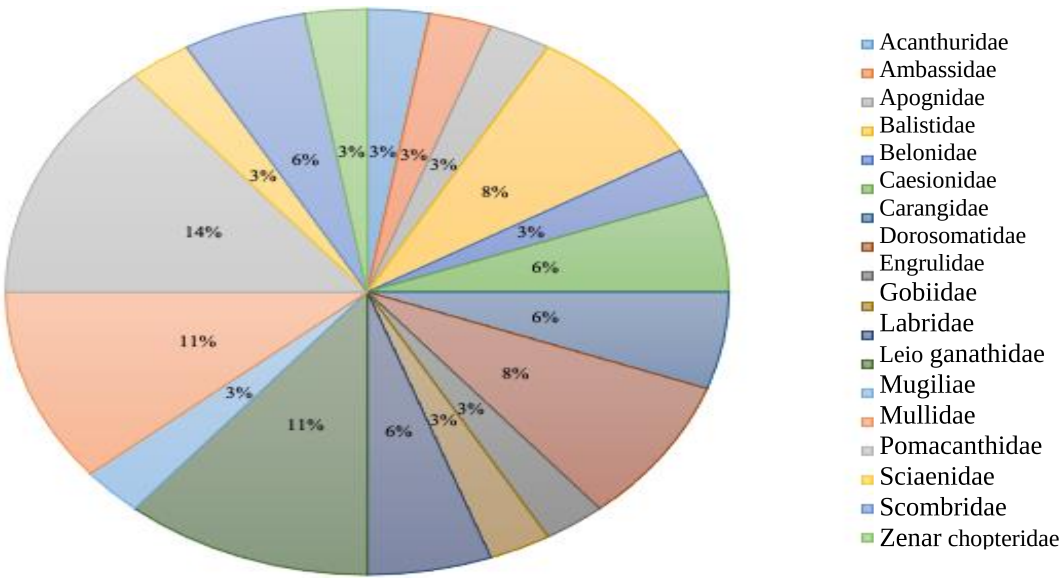

Fig. 3: The ratio of each fish family within the aggregated reads from eDNA samples gathered at the five research locations across three Sulawesi marine areas

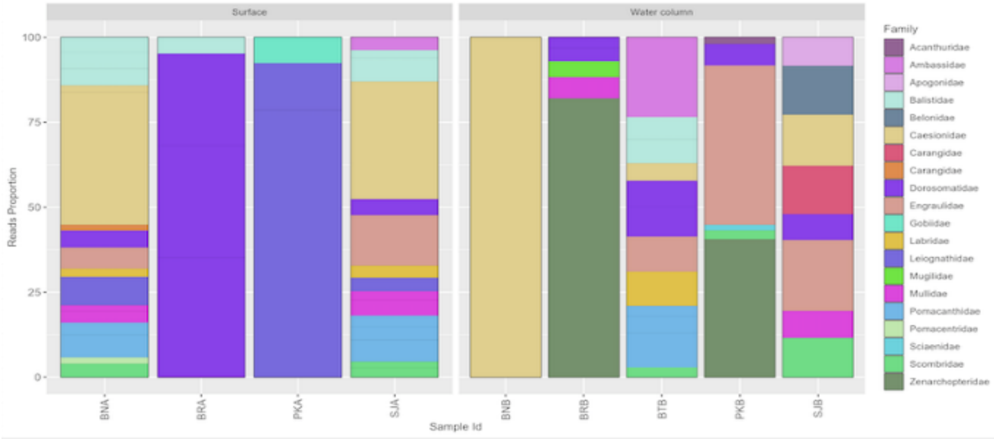

Fig. 4: Proportion of marine fish OTU reads assigned to each family-level taxon by the MiFish eDNA metabarcoding pipeline by site and depth. Site codes: BN = Bone, SJ = Sinjai (Bone Bay), BR = Barru, PK = Pangkep (Makassar Strait), BT = Bantaeng (Flores Sea). Final code letter: A = surface waters (0-1 m depth); B = water column (15 m depth)

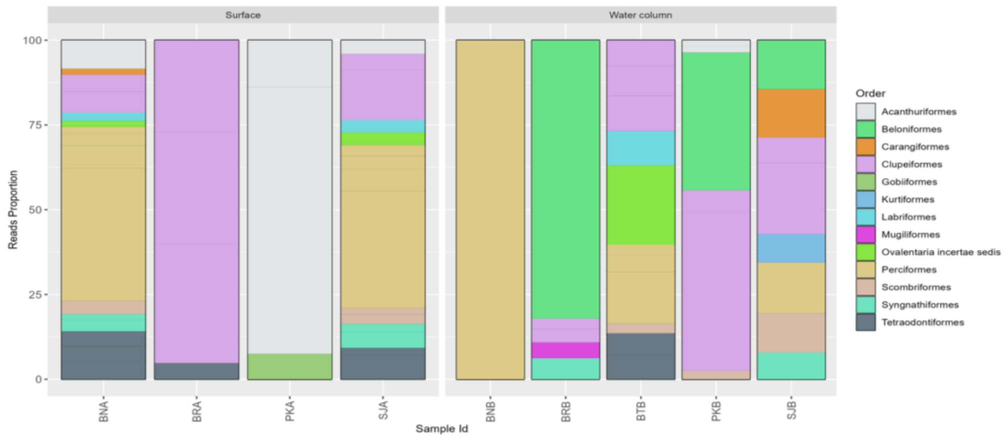

Fig. 5: Proportion of marine fish OTU reads assigned to each order-level taxon by the MiFish eDNA metabarcoding pipeline by site and depth. Site codes: BN = Bone, SJ = Sinjai (Bone Bay), BR = Barru, PK = Pangkep (Makassar Strait), BT = Bantaeng (Flores Sea). Final code letter: A = surface waters (0-1 m depth); B = water column (15 m depth)

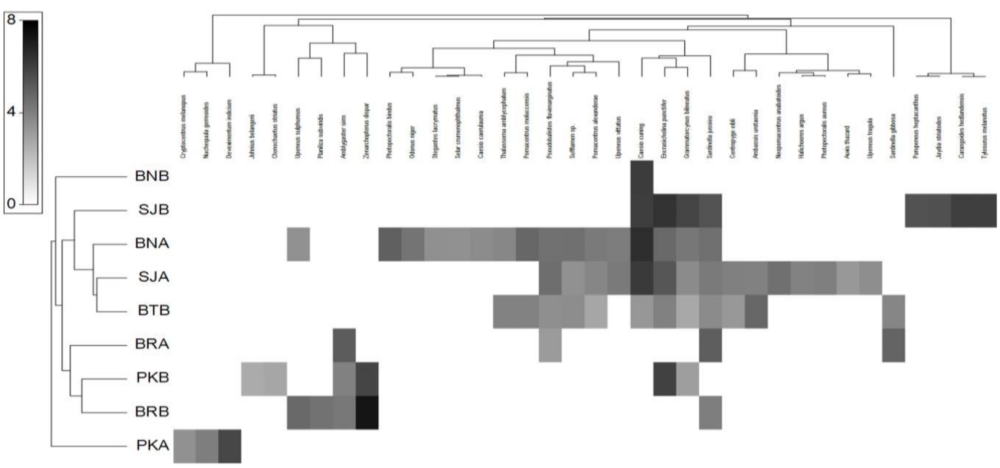

Fig. 6: Marine fish assigned OTU read abundance and site/depth cluster analysis. Site codes: BN = Bone Waters, SJ = Sinjai Waters (Bone Bay), BR = Barru Waters, PK = Pangkep Waters (Makassar Strait), BT = Bantaeng Waters (Flores Sea). Final code letter: A = surface waters (0-1 m depth); B = water column (15 m depth)

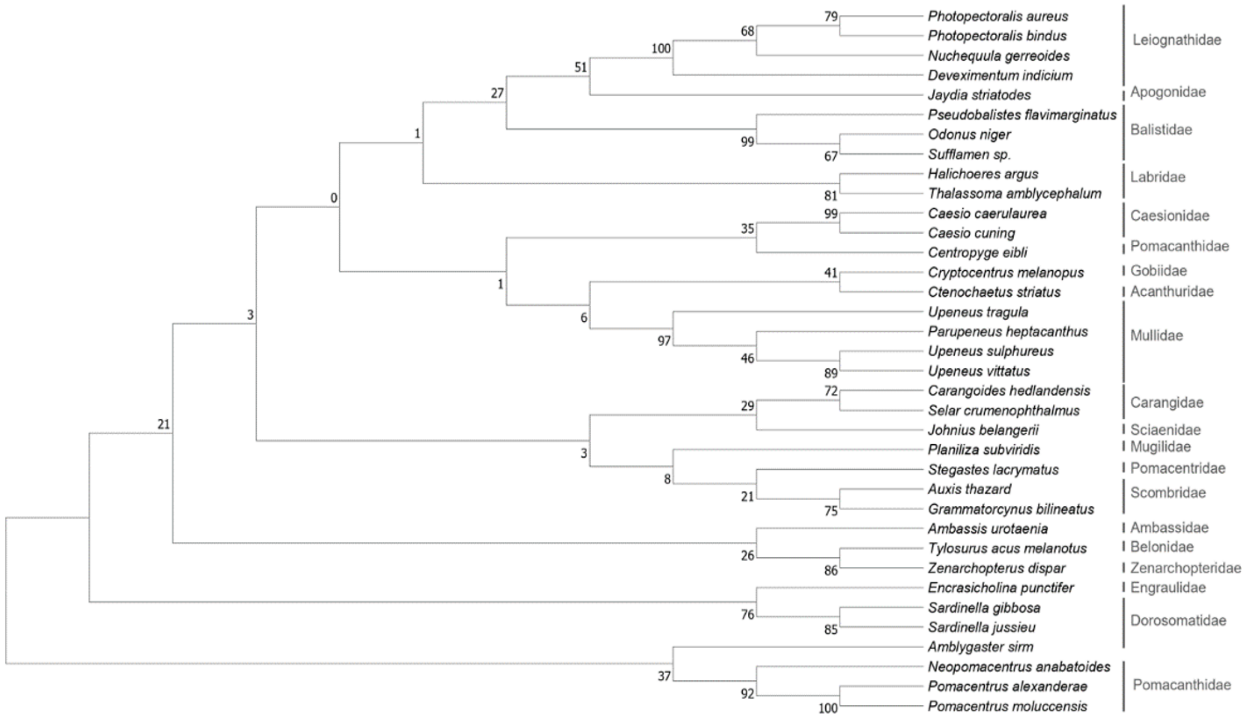

Fig. 7: Neighbor-joining phylogenetic tree based on OTU sequences produced through eDNA metabarcoding from 5 sites in Sulawesi waters

Table 2: Shannon Index (H') based on eDNA metabarcoding data from five locations within Sulawesi waters

Source of sample	Bone bay: Bone	Bone bay: Sinjai	Makassar strait: Barru	Makassar strait: Pangkep	Flores sea: Bantaeng	Average
Surface	0.978	0.981	0.535	0.286	No data	0.556
Water column	0.000	0.878	0.310	0.492	1.009	0.538

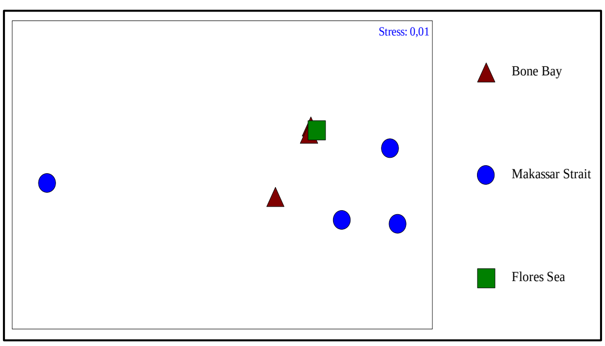

Fig. 8: An nMDS plot illustrates the dissimilarities in fish community composition among sites and sea areas based on the eDNA OTU dataset

Fish Biodiversity in Sulawesi Waters Based on eDNA Metabarcoding

The fish community biodiversity in the three Sulawesi waters (Makassar Strait, Bone Bay, and Flores Sea) examined in this study was different for each area, with variations in the identity and proportions of species within higher-level taxonomic groups between areas. The highest number of species was recorded in Bone Bay (27 species), followed by Makassar Strait (14 species) and Flores Sea (12 species). Three species were found in all three Sulawesi water areas, namely: *Encrasicholina punctifer*, *Grammatorcynus bilineatus*, and *Sardinella jussieu*.

The Shannon index (H') varied between sites and depths (Table 2). Bantaeng site (Flores Sea) had the highest Shannon index with 1.009 from the 15 m water column sample, followed by the Bone Bay (Sinjai and Bone) surface water samples with values approaching one, while the highest value in the Makassar strait was from the Barru surface site with just over 0.5. Species richness (Table 3) also varied between sites and depths. Species richness was highest in the Sinjai 15 m water column sample (0.973). For both indices, the lowest value (0.000) was from the Bone 15 m depth water column sample, where just one species (*Caesio cuning*) was detected. The average values across the five sites were slightly lower at the 15 m depth, despite the highest site-level values also occurring at this depth.

Table 3: Simpson Species richness Index (SI) calculated using eDNA metabarcoding in five locations within Sulawesi waters

	Bone bay		Makassar strait			
					Flores sea	
Source of sample	Bone	Sinjai	Barru	Pangkep	Bantaeng	Average
Surface	0.813	0.834	0.889	0.599	No data	0.627
Water column	0.000	0.973	0.443	0.633	0.935	0.597

Discussion

Molecular identification is a valuable tool for precise species recognition and enjoys widespread usage, although it can have limitations stemming from incomplete databases (Teletchea, 2009). At present, metabarcoding stands as a highly effective method for gauging the species present in a habitat, bypassing the necessity for costly and time-intensive surveys (Foote *et al*., 2012; Piggott, 2016; Rees *et al*., 2014; Roussel *et al*., 2015). Difficulties endure in implementing metabarcoding, encompassing worries regarding its vulnerability to contamination from non-target DNA, biases linked with the primers utilized, sequencing irregularities, possible misidentification of species, and sampling biases (Sato *et al*., 2017). Additionally, this method necessitates adequate equipment support and the processing of bioinformatic data. The benefits of eDNA metabarcoding are enhanced when combined with other approaches, especially when the majority of the species-level OTUs identified exhibit similarity values falling within the range of 95-100%, with a substantial portion sharing 100 or 99% identity with GenBank voucher sequences (Andriyono *et al*., 2019).

The *Perciformes* was the most speciose order based on the eDNA metabarcoding identification process (Table 1). Within this order, the *Pomacanthidae* family is commonly found among reef fish in Indonesia, as are other fish families associated with coral reefs identified in this study including the *Acanthuridae*, *Gobiidae*, *Carangidae*, and *Scombridae* (Wiadnya *et al*., 2023). We also detected economically valuable fish from the order *Clupeiformes*, as four distinct OTUs within the *Clupeidae* family were assigned to species level: *Sardinella Jussieu, Sardinella gibbose, Amblygaster sirm, and Encrasicholina punctifer*.

When it comes to identifying species, our study revealed a comparatively limited number of fish species (36) compared to a previous investigation conducted by Andriyono *et al*. (2019), which identified 53 marine fish species (with a sequence identity of 99-100%). These 36 species we found belong to 18 families, which is fewer than the 27 families documented in the study by Andriyono *et al*. (2019). The discrepancies in the tallies of identified fish species and families can be attributed to various factors, including the utilization of different genetic markers, variances in the geographical regions under scrutiny, and the array of species present at our research sites. It's worth noting that prior studies typically utilized water sample volumes of less than 1 L, whereas this study employed 4 L. For instance, Thomsen *et al*. (2012) collected 500 mL water samples during each sampling occasion, while Yamamoto *et al*. (2017); Andruszkiewicz *et al*. (2017) used 1L water samples. In our initial experiment, we demonstrated that collecting a 4 L water sample was sufficient for consistently achieving successful PCR amplification. This suggests that akin to other tropical environments, it may be essential to collect relatively larger water samples to ensure the success of high-throughput sequencing. Enhancing the detection capacity of eDNA metabarcoding at a specific location relies on the quantity of DNA present in a sample (Schultz and Lance, 2015). Generally, the greater the volume of water sampled, the more species can be identified (Miya *et al*., 2016). Replicates as well as total volume, may also be a factor, as the collection of 3×1 L replicate samples yielded different and additional taxa from each sample in an eDNA study in the Banggai Islands, Central Sulawesi (Moore *et al*., 2021).

The relatively low number of species identified is likely due to the incomplete DNA barcoding dataset for local fish species within the GenBank online database, a concern noted by several other studies (Madduppa *et al*., 2021; Marwayana *et al*., 2022; Moore *et al*., 2021). We anticipate that the eDNA technique holds the potential to reveal a greater number of documented fish species than currently feasible, given the limitations of the existing database. This issue can be better addressed with the establishment of a more comprehensive DNA barcoding database for local marine organisms. Nonetheless, the DNA sequences obtained from the eDNA samples in this study are valuable as they provide baseline data collected in the present timeframe. With the anticipation of additional DNA barcoding sequences being generated for local marine species in the future, we may be able to clarify the identities of previously uncertain or unidentified species that have been the subject of study thus far. The fluctuation in environmental conditions across seasons can also be expected to affect eDNA concentration. This is because the behavior of species, water stratification, temperature, and exposure to ultraviolet radiation undergo changes (Pilliod *et al*., 2014; Zhu, 2006).

The efficacy of eDNA metabarcoding has been successfully demonstrated and supported in studies involving various aquatic organisms that are challenging to collect, such as endangered species (Ikeda *et al*., 2016; Laramie *et al*., 2015; Thomsen *et al*., 2012), as well as those that are endemic (Jerde *et al*., 2011) or invasive (Dejean *et al*., 2012; Takahara *et al*., 2013). Moreover, eDNA offers the capability to provide an overview of biodiversity in a region, facilitating periodic assessments and comparisons with diversity in other areas (Thomsen *et al*., 2012). Notably, this approach is environmentally friendly, reduces survey expenses that can be substantial due to the need for extensive equipment, and is, in other words, highly cost-effective (Smart *et al*., 2016). Additional research employing eDNA metabarcoding may also have relevance in acquiring data beyond biodiversity, including the quantitative assessment of fish species (Alam *et al*., 2020).

To conduct a quantitative investigation, it is essential to establish standardized techniques for collecting and pre-treating samples meant for NGS sequencing analysis. One of the key strengths of eDNA metabarcoding in biodiversity assessment lies in its capacity to generate substantial amounts of information compared to traditional surveys, as extensive datasets are valuable for statistical analyses. However, research teams from diverse countries have amassed substantial data volumes employing different methods for water collection, eDNA preparation techniques, sequencing protocols, and bioinformatic analysis platforms.

In this study, we found that most identified marine fishes were reef fish which belong to several families namely: *Pomacanthidae*, *Acanthuridae*, *Gobiidae*, *Carangidae,* and *Scombridae*. In this study, the use of the eDNA approach demonstrates the efficiency of molecular methods in biodiversity research. This technique allows for the relatively rapid collection of data on species diversity in the Sulawesi waters region. A study conducted in the southern region of Java Island reported reef fish group composition similar to our findings obtained through eDNA metabarcoding in Sulawesi waters. In Prigi Bay, Trenggalek, nine marine fish families were identified: *Serranidae*, *Caesionidae*, *Acanthuridae*, *Lutjanidae*, *Mullidae*, *Nemipteridae*, *Scaridae*, *Haemulidae* and *Carangidae* (Wibowo and Adrim, 2014). Additionally, it was noted that Chaetodontidae fish serve as bioindicators for coral reefs′ health (Reese, 1981). In our study, we detected this particular group of fish using the eDNA metabarcoding method, suggesting that coral reefs prevail as the dominant ecosystem in this region. Moreover, the phylogenetic trees and species-based site clusters can assist in identifying sites that best represent the study area for representative surveys (Bessey *et al*., 2020; Sato *et al*., 2017; Sigsgaard *et al*., 2020).

One limitation of our current eDNA metabarcoding investigation is the uncertainty regarding the native or potentially invasive status of several common fish species. In recent times, there have been advancements in eDNA techniques that enhance detection precision for evaluating intraspecific genetic diversity (Tsuji *et al*., 2020; Uchii *et al*., 2016). Consequently, these approaches may prove valuable for appraising native invasive fish populations and their impact on native biodiversity.

Conclusion

The eDNA metabarcoding technique, utilizing the MiFish-U primer, facilitated the effective detection of tropical marine fish species in the waters surrounding Sulawesi. This eDNA approach offers enhanced insights into the fish species present in three areas within Sulawesi waters. Our study found 36 fish species from 13 orders and 18 families, with the majority falling within the categories of economically valuable fisheries resources, many of which are reef-dwelling fish. The eDNA metabarcoding method is poised to play a foundational role in providing the necessary data for understanding the diversity of marine fish in Sulawesi waters, complementing traditional survey and monitoring techniques. However, addressing the discrepancies and gaps in eDNA results will necessitate further investigation, potentially involving alternative sampling methods and considering water circulation in and out of Sulawesi waters, as well as efforts to enhance the reference sequence databases. Research focusing on seasonal variations in fish community structures through eDNA metabarcoding could enrich our understanding of the relationship between these communities and anthropogenic factors. Expanding the use of eDNA metabarcoding, increasing sampling frequency and site coverage, enables comprehensive analysis of eDNA from entire water bodies. This reveals patterns for specific species and groups, including their occurrence frequencies on a monthly, yearly, and location-specific basis.

Acknowledgment

We are grateful to the Ministry of Education, culture, research, and Technology, Republic of Indonesia who funded this research under the fundamental research for national collaboration scheme through contract number 02381/UN4.22/PT.01.03/2023 and 124/E5/PG.02.00.PL/2023. We also thank Juwiti and Ichsan Ashari Ahmad for assistance with field sampling and Abigail Moore for assistance in manuscript proofreading and editing.

Funding Information

This research was a part of the project titled ‘‘Mapping the distribution and genetic diversity of marine organisms in Sulawesi waters (Makassar Strait, Flores Sea, and Bone

Bay) using eDNA and metabarcoding", funded by the Indonesian government (ministry of education, culture, research and technology). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author's Contributions

Nita Rukminasari: Conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper and sample collection and approved the final draft.

Andi Aliah Hidayani, Sapto Andriyono and Nur Indah Sari Arbit: Analyzed the data, prepared figures and/or tables, sample collection and approved the final draft.

Wilma Joanna Carolina Moka: Performed the experiments, analyzed the data, prepared figures and/or tables and approved the final draft.

Andi Parenrengi: Conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper and approved the final draft.

Ethics

This research did not use human or animal as a subject of research.

Grant Disclosures

The following grant information was disclosed by the authors.

The ministry of education, culture, research and technology, Republic of Indonesia.

Competing Interests

The authors declare there are no competing interests.

References

Alam, M. J., Kim, N. K., Andriyono, S., Choi, H. K., Lee, J. H., & Kim, H. W. (2020). Assessment of fish biodiversity in four Korean rivers using environmental DNA metabarcoding. *PeerJ*, *8*, e9508. https://doi.org/10.7717/peerj.9508

Ambo-Rappe, R., & Moore, A. M. (2019). Sulawesi Seas, Indonesia. *World seas: An Environmental Evaluation*, 559-581. https://doi.org/10.1016/B978-0-08-100853-9.00032-4

Andriyono, S., Alam, M. J., & KIM, H. W. (2019). Environmental DNA (eDNA) metabarcoding: Diversity study around the Pondok Dadap fish landing station, Malang, Indonesia. *Biodiversitas Journal of Biological Diversity*, *20*(12). https://doi.org/10.13057/biodiv/d201241

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*(4), e0176343. https://doi.org/10.1371/journal. pone.0176343

Begon, M., & Townsend, C. R. (2021). *Ecology: from individuals to ecosystems*. John Wiley & Sons. ISBN-10: 9781119279358.

Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., ... & Keesing, J. (2020). Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, *2*(4), 493-504. https://doi.org/10.1002/edn3.74

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890. https://doi.org/10.1093/bioinformatics/bty560

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., ... & Brosse, S. (2019). Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources*, *19*(1), 27-46. https://doi.org/10.1111/1755-0998.12900

Cleary, D. F. R., & Devantier, L. (2011). Indonesia: Threats to the country's biodiversity. *Encyclopedia of Environmental Health*, *1*, 187-197. https://doi.org/10.1016/B978-0-444-52272-6.00504-3 https://doi.org/10.1111/2041-210X.12836

Dejean, T., Valentini, A., Miquel, C., Taberlet, P., Bellemain, E., & Miaud, C. (2012). Improved detection of an alien invasive species through environmental DNA barcoding: The example of the American bullfrog Lithobates catesbeianus. *Journal of Applied Ecology*, *49*(4), 953-959. https://doi.org/10.1111/j.1365-2664.2012.02171.x

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., ... & Lodge, D. M. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 29-41. https://doi.org/10.1111/1755-0998.12433

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, *4*(4), 423-425. https://doi.org/10.1098/rsbl.2008.0118

Foote, A. D., Thomsen, P. F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L. A., ... & Gilbert, M. T. P. (2012). Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. https://doi.org/10.1371/journal.pone.0041781

Fujii, K., Doi, H., Matsuoka, S., Nagano, M., Sato, H., & Yamanaka, H. (2019). Environmental DNA metabarcoding for fish community analysis in backwater lakes: A comparison of capture methods. *PLoS One*, *14*(1), e0210357. https://doi.org/10.1371/journal.pone.0210357

Goldberg, C. S., Pilliod, D. S., Arkle, R. S., & Waits, L. P. (2011). Molecular detection of vertebrates in stream water: A demonstration using Rocky Mountain tailed frogs and Idaho giant salamanders. *PLOS One*, *6*(7), e22746. https://doi.org/10.1371/journal.pone.0022746

Hakim, L. (2017, November). Managing biodiversity for a competitive ecotourism industry in tropical developing countries: New opportunities in biological fields. In *AIP Conference Proceedings* (Vol. *1908*, No. 1). AIP Publishing. https://doi.org/10.1063/1.5012708

Ikeda, K., Doi, H., Tanaka, K., Kawai, T., & Negishi, J. N. (2016). Using environmental DNA to detect an endangered crayfish Cambaroides japonicus in streams. *Conservation Genetics Resources*, *8*, 231-234. https://doi.org/10.1007/s12686-016-0541-z

Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). "Sight‐unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, *4*(2), 150-157. https://doi.org/10.1111/j.1755-263X.2010.00158.x

Jerde, C. L., Wilson, E. A., & Dressler, T. L. (2019). Measuring global fish species richness with eDNA metabarcoding. https://doi.org/10.1111/1755-0998.12929

Kume, M., Lavergne, E., Ahn, H., Terashima, Y., Kadowaki, K., Ye, F., ... & Kasai, A. (2021). Factors structuring estuarine and coastal fish communities across Japan using environmental DNA metabarcoding. *Ecological Indicators*, *121*, 107216. https://doi.org/10.1016/j.ecolind.2020.107216

Laramie, M. B., Pilliod, D. S., & Goldberg, C. S. (2015). Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation*, *183*, 29-37. https://doi.org/10.1016/j.biocon.2014.11.025

Li, W., Fu, L., Niu, B., Wu, S., & Wooley, J. (2012). Ultrafast clustering algorithms for metagenomic sequence analysis. *Briefings in Bioinformatics*, *13*(6), 656-668. https://doi.org/10.1093/bib/bbs035

Madduppa, H., Cahyani, N. K. D., Anggoro, A. W., Subhan, B., Jefri, E., Sani, L. M. I., ... & Bengen, D. G. (2021). eDNA metabarcoding illuminates species diversity and composition of three phyla (chordata, mollusca and echinodermata) across Indonesian coral reefs. *Biodiversity and Conservation*, *30*(11), 3087-3114. https://doi.org/10.1007/s10531-021-02237-0

Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, *27*(21), 2957-2963. https://doi.org/10.1093/bioinformatics/btr507

Marwayana, O. N., Gold, Z., Meyer, C. P., & Barber, P. H. (2022). Environmental DNA in a global biodiversity hotspot: Lessons from coral reef fish diversity across the Indonesian archipelago. *Environmental DNA*, *4*(1), 222-238. https://doi.org/10.1002/edn3.257

Minamoto, T., Yamanaka, H., Takahara, T., Honjo, M. N., & Kawabata, Z. I. (2012). Surveillance of fish species composition using environmental DNA. *Limnology*, *13*, 193-197. https://doi.org/10.1007/s10201-011-0362-4

Miya, M., Minamoto, T., Yamanaka, H., Oka, S. I., Sato, K., Yamamoto, S., ... & Doi, H. (2016). Use of a filter cartridge for filtration of water samples and extraction of environmental DNA. *JoVE Journal of Visualized Experiments*, (117), e54741. https://doi.org/10.3791/54741

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., ... & Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*(7), 150088. https://doi.org/10.1098/rsos.150088

Moore, A. M., Jompa, J., Tassakka, A. C. M. A., Yasir, I., Ndobe, S., Umar, W., ... & Barber, P. H. (2021). Sharks and rays (Chondrichthyes) around Banggai Island, Banggai MPA, Indonesia: Biodiversity data from an environmental DNA pilot study. *Aquaculture, Aquarium, Conservation and Legislation*, *14*(2), 725-745.

Piggott, M. P. (2016). Evaluating the effects of laboratory protocols on eDNA detection probability for an endangered freshwater fish. *Ecology and Evolution*, *6*(9), 2739-2750. https://doi.org/10.1002/ece3.2083

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2014). Factors influencing detection of eDNA from a stream‐dwelling amphibian. *Molecular Ecology Resources*, *14*(1), 109-116. https://doi.org/10.1111/1755-0998.12159

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA–a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*(5), 1450-1459. https://doi.org/10.1111/1365-2664.12306

Reese, E. S. (1981). Predation on corals by fishes of the family Chaetodontidae: Implications for conservation and management of coral reef ecosystems. *Bulletin of Marine Science*, *31*(3), 594-604. https://www.ingentaconnect.com/content/umrsmas/bullmar/1981/00000031/00000003/art00011

Roussel, J. M., Paillisson, J. M., Treguier, A., & Petit, E. (2015). The downside of eDNA as a survey tool in water bodies. *Journal of Applied Ecology*, 823-826. https://doi.org/10.1111/1365-2664.12428

Sard, N. M., Herbst, S. J., Nathan, L., Uhrig, G., Kanefsky, J., Robinson, J. D., & Scribner, K. T. (2019). Comparison of fish detections, community diversity, and relative abundance using environmental DNA metabarcoding and traditional gears. *Environmental DNA*, *1*(4), 368-384. https://doi.org/10.1002/edn3.38

Sato, H., Sogo, Y., Doi, H., & Yamanaka, H. (2017). Usefulness and limitations of sample pooling for environmental DNA metabarcoding of freshwater fish communities. *Scientific reports*, *7*(1), 14860. https://doi.org/10.1038/s41598-017-14978-6

Schultz, M. T., & Lance, R. F. (2015). Modeling the sensitivity of field surveys for detection of environmental DNA (eDNA). *PloS One*, *10*(10), e0141503. https://doi.org/10.1371/journal.pone.0141503

Shaw, J. L., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation*, *197*, 131-138. https://doi.org/10.1016/j.biocon.2016.03.010

Smart, A. S., Weeks, A. R., van Rooyen, A. R., Moore, A., McCarthy, M. A., & Tingley, R. (2016). Assessing the cost-efficiency of environmental DNA sampling. *Methods in Ecology and Evolution*, *7*(11), 1291-1298. https://doi.org/10.1111/2041-210X.12598 https://doi.org/10.1038/s41598-017-12501-5

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*, *33*(1), 196-205. https://doi.org/10.1111/cobi.13183 https://doi.org/10.1371/journal.pone.0175186

Suarez-Bregua, P., Alvarez-Gonzalez, M., Parsons, K. M., Rotllant, J., Pierce, G. J., & Saavedra, C. (2022). Environmental DNA (eDNA) for monitoring marine mammals: Challenges and opportunities. *Frontiers in Marine Science*, *9*, 987774. https://doi.org/10.3389/fmars.2022.987774

Sigsgaard, E. E., Torquato, F., Frøslev, T. G., Moore, A. B., Sørensen, J. M., Range, P., ... & Thomsen, P. F. (2020). Using vertebrate environmental DNA from seawater in biomonitoring of marine habitats. *Conservation Biology*, *34*(3), 697-710. https://doi.org/10.1111/cobi.13437

Takahara, T., Minamoto, T., & Doi, H. (2013). Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PloS One*, *8*(2), e56584. https://doi.org/10.1371/journal.pone.0056584

Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries*, *19*, 265-293. https://doi.org/10.1007/s11160-009-9107-4

Thomsen, P. F., Kielgast, J. O. S., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., ... & Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, *21*(11), 2565-2573. https://doi.org/10.1111/j.1365-294X.2011.05418.x

Tsuji, S., Maruyama, A., Miya, M., Ushio, M., Sato, H., Minamoto, T., & Yamanaka, H. (2020). Environmental DNA analysis shows high potential as a tool for estimating intraspecific genetic diversity in a wild fish population. *Molecular Ecology Resources*, *20*(5), 1248-1258. https://doi.org/10.1111/1755-0998.13165

Uchii, K., Doi, H., & Minamoto, T. (2016). A novel environmental DNA approach to quantify the cryptic invasion of non-native genotypes. *Molecular Ecology Resources*, *16*(2), 415-422. https://doi.org/10.1111/1755-0998.12460

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology and Evolution*, *24*(2), 110-117. https://doi.org/10.1016/j.tree.2008.09.011

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, Nicolas., Copp, G, H., Geniez, Philippe., Pont, Didier., Argillier, Christine., Baudoin, Jean-Marc., Peroux, Crivelli, A, J., Olivier, A., Acqueberge, M., Le, B, M., Møller, P., R., Willerslev, E., & Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, *25*(4), 929-942. https://doi.org/10.1111/mec.13428

Wiadnya, D. G. R., Kurniawan, N., Hariati, A. M., Astuti, S. S., Paricahya, A. F., Dailami, M., & Kusuma, W. E. (2023). DNA barcoding of the most common marine ornamental fish species spilled over from a small-sized marine protected area, Bali Barat National Park, Indonesia. *Biodiversitas Journal of Biological Diversity*, *24*(1). https://doi.org/10.13057/biodiv/d240107

Wibowo, K., & Adrim, M. (2014). Komunitas ikan-ikan karang Teluk Prigi, Trenggalek, Jawa Timur. *Zoo Indonesia*, *22*(2). https://doi.org/10.52508/zi.v22i2.320

Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., ... & Miya, M. (2017). Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, *7*(1), 40368. https://doi.org/10.1038/srep40368 https://doi.org/10.1016/j.ecolind.2015.11.022

Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, *7*(1-2), 203-214. https://doi.org/10.1089/10665270050081478

Zhu, B. (2006). Degradation of plasmid and plant DNA in water microcosms monitored by natural transformation and real-time Polymerase Chain Reaction (PCR). *Water Research*, *40*(17), 3231-3238. https://doi.org/10.1016/j.watres.2006.06.040