Splash, splash. Who's there? Advantages and limitations of the environmental DNA (eDNA) metabarcoding in assessing megadiverse but poorly known tropical community of fishes

Danial Hariz Zainal Abidin
Universiti Sains Malaysia
Siti Azizah Mohd. Nor
Universiti Malaysia Terengganu
Sébastien Lavoué
Universiti Sains Malaysia
Masazurah A. Rahim
Fisheries Research Institute
Noor Adelyna Mohammed Akib (adelyna@usm.my)
Universiti Sains Malaysia

Research Article

Keywords:

Posted Date: February 23rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1350797/v1

Abstract

Biodiversity surveys are crucial to monitor the health of threatened aquatic ecosystems, such as tropical estuaries and mangroves. Conventional monitoring methods are intrusive, time-consuming, substantially expensive and frequently only provide rough assessments in complex habitats. Recent advanced molecular methods such as environmental DNA (eDNA) using high-throughput sequencing technology are promising although only few applications in tropical estuarine ecosystems have been reported. In this study, we explore the advantages and limitations of an eDNA metabarcoding survey on the fish community of the Merbok Estuary (Peninsular Malaysia). COI and 12S eDNA metabarcoding assays collectively detected 178 species belonging to 127 genera, 68 families, and 25 orders. This approach captured significantly more species than in any previous traditional surveys, including a few species of conservation importance. However, we highlight three limitations: (1) in the absence of a comprehensive reference database the identities of several species are unresolved; (2) a fraction of previously documented specimen-based diversity was not captured by the current method, may be as a consequence of PCR primer specificity, and (3) the detection of non-resident species; stenohaline freshwater taxa (e.g. cyprinids, channids, osphronemids) and marine coral reef taxa (e.g. some sharks, holocentrids and syngnathids), not known to frequent estuaries, leading to the supposition that their DNA have drifted into the estuary through water movement. The community analysis revealed that fish diversity along the estuary is not homogenous, with the upstream more diverse than further downstream. This may be the consequence of the salinity or pollution gradients. In summary, we demonstrated the practicality of eDNA metabarcoding in assessing fish community and structure within a complex and rich tropical environment within a short sampling period. However, some limitations need to be considered and addressed to fully exploit the efficacy of this approach, in particular the development of a comprehensive reference genetic database.

Introduction

Estuaries associated with mangrove forests are at the interface of freshwater and marine environments, serving essential ecological functions including protective, feeding, spawning, and rearing habitats for a diverse array of aquatic organisms[1]. Mangroves provide critical natural services such as coastline protection, nutrient synthesis, and fishery resources, which are essential for sustaining local communities' socioeconomic livelihoods[2]. Regrettably, such vital human-nature relationships are jeopardized by habitat degradation, pollution, and overfishing[3]. Well-protected tropical estuaries, particularly those comprising mangrove habitats, support diverse and complex biological communities that include salinity-tolerant resident species as well as frequenters, including numerous marine species, that rely on this ecosystem for food, shelter, breeding, or nursing[4]. Assessing the ecosystem's diversity and community structure is critical for evaluating its health and key to successfully protect its biodiversity[5,6]. Thus, inventories and long-term biodiversity surveillance of native communities in estuarine mangrove ecosystems are required to ensure stability and resilience against anthropogenic disturbance, monitoring alien species invasion, and to prevent the loss of native, sometimes endemic, species[1,6,7]. Long-term monitoring efforts aid in identifying which species are most vulnerable to environmental stressors, determining new threats, and revealing the interconnectedness among species in an ecosystem[8]. Conventional methods are, however, time-consuming, often cost-prohibitive and requires expertise in taxonomy[9]. Consequently, biodiversity in many mangrove habitats, notably those in Southeast Asia, is still not adequately monitored, impeding their management and conservation.

Located within the Sundaland biodiversity hotspot[10], Malaysia is home to astonishing levels of diversity and endemism. In particular, [11] reported the presence of more than 1,400 marine and brackish species in coastal Malaysian waters, a significant proportion either residing or frequenting mangrove ecosystems. The Merbok Estuary, located in northwest Peninsular Malaysia, is not only one of the region's largest remaining patches of mangrove forest (~ 40,000 ha)[12], it is also recognised as the World's most diverse ecosystem for mangrove species in terms of species richness per unit area[13]. In 1951, the Merbok Estuary was gazetted as a permanent forest reserve known as the Sungai Merbok Mangrove Forest Reserve. In spite of its value and critical role in providing vital ecosystem services and livelihoods to the local community, few biodiversity inventories have been conducted in Merbok Estuary [14–16]. Recent ichthyodiversity surveys in Merbok Estuary using three different approaches (morphological and DNA barcoding on adults, and metabarcoding of larvaes) revealed the presence of about 180 species[7,12,17]. Although these studies discovered a large number of species, their respective list of species only partially overlapped to each other, indicating limitations in each approach and incomplete taxonomic coverage. Thus, there is a need to explore alternative monitoring approaches for capturing an exhaustive species diversity list[18,19].

Environmental DNA (eDNA), also referred to as the metabarcoding approach is recognised as a revolutionary method to effectively conduct multi-taxa inventory surveys using bulk DNA extracted from environmental samples (e.g. water, soil)[20,21]. It overcomes the shortcomings of conventional survey methods to efficiently characterise fish assemblages in aquatic habitats as have been reported in estuarine[22–24], marine[25,26], and freshwater ecosystems[27,28]. The use of eDNA as a powerful monitoring instrument is increasingly acknowledged as it reveals greater diversity at a lower cost compared to conventional surveying methods[29–32]. The eDNA metabarcoding method is said to outperform conventional surveys in taxa detection (i.e. identification of novel species and conservation targets) and determination of the local community composition[29,32]. This approach has also been employed to address other objectives, including detection of invasive species[33,34], identification of cryptic species[35,36], understanding spawning ecology[37,38], and in assessing ecosystem health and dynamics[39].

In this study, we utilised the eDNA metabarcoding method to estimate fish diversity in the Merbok Estuary. Through this, we aim to augment the Merbok Estuary previous checklists assembled through traditional surveys towards a complete species database. We also evaluated the complementarity of two eDNA metabarcoding assays (COI and 12S) to characterise the diversity patterns and composition of fish communities in this ecologically and biologically diverse landscapes of Merbok Estuary. Our study demonstrated the potential of eDNA metabarcoding as an efficient tool in inventorying the fish diversity of this globally important mangrove hotspot.

Results

eDNA-based fish detection in Merbok Estuary

A total of 12,958,643 and 10,757,026 raw amplicon reads were generated from the COI and 12S assays, respectively (Supplementary Table 3). The mean number of filtered reads (after post-quality processing and chimera removal) was 184,017 and 140,118 per sample for COI and 12S assays, respectively. Of these, the COI metabarcoding assay identified 8,332 MOTUs whereas the 12S assay identified 859 MOTUs. The COI primer pair amplified taxa from multiple animal phyla (e.g. Insecta, Gastropoda, Aves, Mammalia, Chondrichthyes, and Actinoperygii), whereas the 12S primers were more specific in amplifying vertebrates primarily (Actinoperygii, Amphibia, Aves, and Mammalia) and a small percentage of diatoms. Even though COI assay produced a higher number of reads leading to the detection of almost ten times more MOTUs when compared to the 12S assay, many of the COI MOTUs were assigned to non-fish taxa. 98.51% of the 12S MOTUs were assigned to fish taxa whereas only a mere 3.90% of the COI MOTUs were assigned to fish taxa. The final taxonomic assignment after multiple layers of filtering resulted in a total of 71 Actinopterygii and seven Chondrichthyes MOTUs identified from COI metabarcoding assay and 235 Actinopterygii MOTUs from 12S metabarcoding assay.

Collectively, both COI and 12S metabarcoding assays detected 178 fish species belonging to 127 genera, 68 families, and 25 orders (Table 3). Overlapped species detections and cross-amplification of taxonomic groups between both assays are visualised in the *circlize* plot[40] (Figure 2). Of these, 174 were bony fish species (class: Actinopterygii) and the other four were elasmobranch taxa (class: Chondrichthyes). The most speciose orders observed in the bony fish taxa were Perciformes (with 40 species representing 22.5% of the total number of species detected), followed by Carangiformes (26 species, 14.6%), and Gobiiformes (25 species, 14%). Among these, the predominant families include Gobiidae (gobies; 18 species), Carangidae (jacks and pompanos; 13 species), Serranidae (groupers and sea basses; 9 species), and Sciaenidae (croakers; 8 species). The four elasmobranch species belong to three families, the Carcharhinidae (requiem sharks; 2 species), Dasyatidae (whiptail stingrays; 1 species), and Gymnuridae (butterfly rays; 1 species). We detected seven fish species that are listed as either 'Endangered' (EN) or 'Near Threatened' (NT) in the IUCN Red List of Threatened Species (IUCN, 2020): blacktail reef shark (*Carcharhinus amblyrhynchos*, EN), blacktip reef shark (*Carcharhinus melanopterus*, NT), scaly whipray (*Brevitrygon walga*m, NT), long-tailed butterfly ray (*Gymnura poecilura*, NT), Indonesian shortfin eel (*Anguilla bicolor*, NT), Bombay-duck (*Harpadon nehereus*, NT), and narrow-barred Spanish mackerel (*Scomberomorus commerson*, NT). Two of the detected species were classified as invasive species: Nile tilapia (*Oreochromis niloticus*) and the guppy (*Poecilia vivipara*).

Comparison of eDNA metabarcoding detections and previous specimen-capture records

Sampling campaigns conducted in Merbok Estuary during the last decade (2010-2019) recorded a total of 165 species from 102 genera, 49 families, and 19 orders[7,12,16] whereas our eDNA metabarcoding assessment conducted in the same area recorded a total of 178 species (Figure 3; Supplementary Table 2; Table 3). Overlapping taxa detected from both methods at different taxonomic ranks is shown in Figure 3. At the ordinal, familial, generic, and specific levels, approximately 90% (17 orders), 82% (40 families), 60% (61 genera) and 41% (67 species) are detected both by previous capture surveys and this eDNA metabarcoding assays, respectively.

The most speciose orders detected by both methods were Perciformes, Carangiformes, and Gobiiformes. It is highly apparent that eDNA metabarcoding succeeded in capturing higher diversity with ~40% more species than previous conventional surveys. The eDNA metabarcoding detected an additional eight orders: Carcharhiniformes (ground sharks), Anabantiformes (gouramies), Blenniiformes (blennies), Cypriniformes (carps and minnows), Cyprinodontiformes (killifish), Gonorynchiformes (milkfish), Holocentriformes (squirrelfishes and soldierfishes), and Syngnathiformes (trumpetfishes). In addition, the metabarcoding assays unexpectedly detected stenohaline freshwater species belonging to the families Cyprinidae (e.g. *Barbodes binotatus*, *Mystacoleucus obtusirostris*, *Tor tambra*, *Rasbora pauciperforata*), Channidae (*Channa limbata* and *C. striata*), and Osphronemidae (*Trichogaster lalius* and *Trichopodus trichopterus*) (Table 3). Nine families recorded by conventional surveys were not detected by eDNA metabarcoding, namely the families Hemiscylliidae (longtailed carpet sharks), Dussumieriidae (round herrings), Hemiramphidae (halfbeaks), Phallostethidae (priapium fish), Paralichthyidae (sand flounders), Bothidae (lefteye flounders), Lethrinidae (emperor breams), Drepanidae (sicklefishes), and Stromatidae (butterfish).

Diversity patterns and composition

The MOTU richness recovered from the two metabarcoding assays appeared to vary across samples and designated zones (Figure 4). The COI assay detected many brackish and marine fishes belonging to the orders Carangiformes, Gobiiformes, Centrarchiformes, and Perciformes. The zones closer to the river mouth (Zones A and B) were dominated by marine families such as Carangidae, Cirrhitidae, Eleotridae, Sciaenidae, and Plotosidae, which are associated with pelagic habitats or sandy substrates (Supplementary Figure 1). Interestingly, we detected the presence of the blacktail reef shark, *Carcharhinus amblyrhynchos* (identity= 100%, E-value= 0.021), a marine shark species, with the highest read abundance recorded from site A5.2 from the COI assay. The 12S metabarcoding assay however, revealed richer species diversity, in particular from the families Scombridae, Mugilidae, Scatophagidae, Sciaenidae, and Plotosidae (Supplementary Figure 1). Notwithstanding, the MOTU analysis of both assays on the five pooled samples (samples TA, TB, TC, T1, and T2) did not detect any additional species.

The α-diversity analysis based on the indices of richness (observed MOTUs and Chao1) and diversity (Shannon) showed varied patterns among the designated zones for both COI and 12S assays after rarefaction (Figure 5). The coastal area of the Merbok Estuary (Zone A) harboured the highest MOTU richness based on COI whereas the 12S estimated higher richness in the midstream and upstream of the Merbok Estuary (Zones B and C), as revealed by the Chao1 and Shannon indices. The sample-based MOTU accumulation curve displayed direct relationship of diversity with number of samples (Figure 6), with higher detections observed in the 12S metabarcoding assay. NMDS ordination plots for both metabarcoding assays, based on zones and pooled samples, exhibit partial to total overlap among fish communities (Figure 7). Comparing MOTUs on a pair-wise basis revealed considerable disparities in MOTU diversity among the three zonations (PERMANOVA, Bray-Curtis, p<0.001) (Table 4). In the ordination of the COI metabarcode, Zone A and Zone C community clusters were separated, with the inertia ellipse of Zone B partially overlapping both clusters. Within the 12S assay community analysis, Zone A, on the other hand, has

a substantially broader community ellipse that encompasses both Zone B and Zone C clusters. The PERMANOVA analysis returned significantly different variances among the three zones (Zones A, B, and C) but minimal differentiation based on salinity measurement at all sampling sites (Table 1).

Discussion

We investigated the efficiency of the eDNA metabarcoding approach in estimating fish diversity of a tropical brackish environment and compared it to previous conventional surveys. The feasibility and vast potential of eDNA metabarcoding as a rapid tool in assessing fish diversity in Merbok Estuary was well demonstrated. Metabarcoding overcomes the limitations of conventional survey methods (e.g. the use of invasive/ lethal sampling gears that may disrupt habitats), resulting in a significant reduction of direct impact on the natural ecosystem[41]. In this regard, eDNA-based metabarcoding methods provide a non-invasive and cost-effective alternative for surveying aquatic ecosystems, and species with conservation interest. But its efficiency is contingent on the availability of a comprehensive and reliable reference taxonomic database. Furthermore, we also observed several conflicting findings; the detection of taxa unlikely to be present in the surveyed areas and the non-detection of previously recorded taxa (albeit a moderate proportion). We discuss all these points below.

Fish diversity detected via conventional surveys vs. eDNA metabarcoding

Our results lend further support to the use of an eDNA metabarcoding approach on water samples collected from the Merbok Estuary at capturing the diversity of its fish community over time-consuming conventional surveys; a finding noted in other similar studies[29,31,42]. As illustrated, the eDNA metabarcoding assays detected ~82% of the fish families previously recorded during the last decade in Merbok Estuary by traditional surveys and detected additional <100 species (i.e. residents, migrants, or frequenters) inhabiting Merbok Estuary, within just a two-day sampling period. Several were new records. Some of these species are regarded to be in larval stage and thus uncapturable using the sampling gears employed in previous surveys. Furthermore, taxonomic identification of larva is very challenging even when sampled[17]. With the capacity to differentiate species at any stage of development (i.e. larva, juvenile, or adults), the eDNA metabarcoding approach improves the accuracy of community composition estimation and provides additional information on species interactions within a community[42]. Thus, the eDNA assays provide a quick and powerful estimation of an aquatic ecosystem diversity through a more powerful detection efficacy[28].

The species detected by eDNA metabarcoding reflect the estuarine-mangrove community in Merbok Estuary, including estuarine residents (e.g. *Butis butis*, *Lates calcarifer*, *Epinephelus coioides*), marine-estuarine-dependent species that include amphidromous species (e.g. *Batrachomoeus trispinosus*, *Acentrogobius caninus*, *Siganus guttaus*), marine species and marine migrants that use the estuary as a nursery and/or feeding grounds (e.g. *Lutjanus johnii*, *Caranx ignobilis*, *Eleutheroma tetradactylum*), and estuarine-freshwater-dependent species (e.g. *Scatophagus argus*)[7,12,43].

Limitation one: absence of a comprehensive reference database

Despite its potential to uncover higher diversity in this estuary, we highlight here three limitations of the eDNA method from water samples. Firstly, the major challenge is the absence of a comprehensive reference fish database of the two genes in this area and in general, of tropical fish community. This is particularly true in megadiverse taxonomic and cryptic groups in tropical hotspot regions, such as Merbok. Based on historical literature review of formal and informal documentations (e.g.[44-47]), more than 500 species of fish could occur in Merbok region (coastal and brackish) with many of them being small and cryptic and not yet genetically examined. Furthermore, the significant number of misidentified species sequence entries in GenBank[48] results in possible erroneous taxonomic assignment in the eDNA samples[49].

In recent years, ecosystem managers are turning to eDNA metabarcoding to glean information about community composition and diversity. The target loci (i.e. COI and 12S) were purposely selected because they are frequently employed in DNA barcoding and phylogenetic studies[7,50,51]. However, there remain significant gaps in the reference database of several taxon groups which hampers the optimal utility of eDNA metabarcoding[52]. Thus, future monitoring surveys must prioritize the generation of reference sequences to include more annotated sequences in global databases (e.g. BOLD[53] and NCBI GenBank[48]). Establishing high-quality reference databases of local biodiversity should be a prerequisite in DNA-based biodiversity monitoring. We have established a COI-based barcode of 134 fish species occurring in Merbok Estuary in a previous study[7]. More barcoding efforts on the region's ichthyodiversity are underway, including sequences generated from other mitochondrial markers (e.g. 12S and cytochrome *b*). This will elevate the efficacy of eDNA metabarcoding and offers the opportunity to re-analyse sequences generated in this study against updated databases. The eDNA dataset generated in this study will be valuable for estimating the Merbok fish community composition changes across time.

Limitation two: detection of metabarcodes from surrounding areas due to drifting tissue materials (not associated with the actual presence of whole fish)

Water movements could potentially transmit eDNA from adjacent waterways that flow into (freshwater streams) or out of (marine ecosystem) this estuary, extending the spatial coverage and hence taxonomic detection[36,42]. This factor needs to be considered when assessing the diversity of fish communities along a heterogeneous spatial gradient, such as estuaries and mangrove ecosystems[54].

A few species from strictly freshwater families of Cyprinidae, Anabantidae, Channidae, Osphronemidae, Aplocheilidae or Poeciliidae were unexpectedly encountered through the eDNA assays. None of these have been recorded in previous surveys within the estuary but are known to occur upstream, in the freshwater part of the Merbok and Muda basins. For instance, [55] recorded the presence of the blue panchax, *Aplocheilus armatus* (Aplocheilidae), and the dwarf snakehead, *Channa limbata* (Channidae), in the streams of Gunung Jerai Forest Reserve, located ~15 km north of Merbok Estuary. We opine that the presence was due to downstream movement of DNA traces into the estuary and not the actual presence of the fish individual. A more recent diversity survey conducted in the streams of Ulu Muda Forest Reserve (located within the Muda basin and connected to the Merbok Estuary), located ~80 km northeast of our

study area, recorded several taxa that correspond to our eDNA metabarcodes such as *Barbodes binotatus*, *Mystacoleucus obtusirostris*, *Tor tambra*, *Anabas testudineus*, and *Channa striata*[56]. Similarly, the detection of coral reef endemic species is likely due to eDNA transportation into the estuary by oceanic and tidal currents. This is supported by the timing of our sampling during high tide, when the ocean level reaches its maximum and seawater freely enters the estuary. [22] demonstrated that species from the surrounding marine and freshwater habitats are more easily detected during this period.

Our metabarcoding assays also detected a few taxa of conservation importance including the blacktail reef shark *Carcharhinus amblyrhynchos*, the blacktip reef shark *Carcharhinus melanopterus*, and the Indonesian shortfin eel, *Anguilla bicolor*, none of which had been previously recorded, thus inferring their occurrence in the study area or in nearby habitats. The discovery of the endangered *Carcharhinus amblyrhynchos* was unexpected. However, this species was as one of the common catches in Peninsular Malaysia a few decades ago based on a report on Malaysian shark fisheries[57]. Apparently, it still exists within the Merbok region and in all likelihood in several other places. Moving forward, more intensive studies to identify these areas are needed. The discovery of these threatened species highlights the value of eDNA metabarcoding as a tool for assessing biodiversity and represents a promising step towards a holistic conservation in estuaries, mangrove and coastal ecosystems[20,23,41].

Limitation three: eDNA metabarcoding methods may not successfully capture whole species diversity

Despite its enormous potential, some taxa may still go undetected due to several inherent limitations of the metabarcoding assay[58]. The current study missed nine families previously identified by conventional surveys. This may be attributed to the following reasons: small number of individuals present in the environment, hence insufficient eDNA in the water sample[59,60], rapid eDNA degradation[61] , or due to primer specificity[36]. Often eDNA metabarcoding methods exhibit some degree of taxonomic selectivity due to primer specificity. To reduce the impact of selectivity, it is advised to use several DNA loci or primer sets[62,63].

Two genetic markers were used in our study to compensate the drawbacks of using a single primer in the metabarcoding assays, such as weak taxonomic discriminatory power and primer binding biases[62]. We utilized COI and 12S markers in our eDNA assays, in reference to previous metabarcoding studies[29,64,65]. Our study showed that the COI assay produced higher reads with broader detections across taxa when compared to the 12S assay, which is similar to the findings of[29,64]. In comparison to the COI assay, 12S assay retrieved a higher number of species. The discrepancies in taxonomic detection between the COI and 12S assays are most likely attributable to the combination of few factors, including primer bias, differences in the completeness of the reference database, and the different taxonomic resolution of the primers[21,66].

The 12S MiFish primers[67] are reputed to be teleost specific and have been successfully employed in other eDNA surveys over a wide range of aquatic environments, including estuaries[22,24], marine[42,68], and freshwater ecosystems[60,69]. Despite being the standard barcode gene[70], the COI marker appeared to be a less common choice as a fish metabarcoding primer. Notwithstanding the wider taxonomic coverage of COI sequences than other loci, its high variability within this standard barcode region makes it difficult to design universal fish-specific primers for short amplicons, hence reducing the COI optimal capacity in eDNA biodiversity assessment[71].

Diversity patterns and composition

Biodiversity is inextricably linked to environment, especially in estuaries and mangrove ecosystems[72]. Based on the ecological characteristics of ecosystem landscapes in Merbok Estuary, there are three pre-defined zones (i.e. Zone A, Zone B, and Zone C). These three zones have different degrees of anthropogenic impact. Comparing the three zones: Zone B, the area with the lowest anthropogenic disturbances, harbours the highest diversity estimates based on the Chao1 and Shannon's indices on the 12S assay (Figure 5b). Our findings agree with[16] who likewise found the greatest abundance of fish species in the midstream of the Merbok Estuary. This is an indicator that optimal conditions to support high biodiversity warrants a pristine environment[22]. The NMDS ordination allied with PERMANOVA based on read counts of the two metabarcoding assays, confirming clustering results that three distinct fish communities were distinguishable by the designated zones. However, the results of the PERMANOVA showed that salinity measures at all sampling sites did not affect the community structuring of detected fish taxa, which appeared to minimally vary across sites in Merbok Estuary (Table 1).

Fish communities respond to many environmental variables such as salinity, water temperature, food availability, or sediment type to determine their dispersion throughout the aquatic habitats, especially in estuaries and mangrove ecosystems[73]. Apart from the variations observed in the zones, no other notable spatial trends in the MOTUs read abundance across the Merbok Estuary were observed. However, a detailed spatial and geographical comparison of fish communities requires a more extensive sampling coverage, a higher sample density, and comprehensive physicochemical data collection[22,26], which was not the primary goal of our study. Even so, a recent study demonstrated that species assemblage composition differed noticeably across different environments on a small geographic scale, indicating the specificity of eDNA signals despite vast oceanic water flow[26]. Even though our study has shown that eDNA can provide valuable information on the community composition of different habitats within a specific region, we affirm that complementary conventional approaches with better spatial fidelity will allow us to assign species and attribute them to specific habitats with greater certainty supporting the eDNA method. Thus, we advocate incorporating eDNA metabarcoding sampling into traditional surveys as a best practice for performing whole biodiversity surveys in a complex ecosystem like the estuaries and the mangroves. [52] demonstrated that combining both eDNA and conventional approaches uncovers a broader taxa diversity and provides a more holistic perspective of species compositions in an aquatic environment.

Conclusion

This is one of the first studies aimed to evaluate the practicability and value of an eDNA metabarcoding approach to assess fish diversity in a tropical mangrove estuarine environment in this region. Our findings demonstrated strong advantages of using such approach, albeit with some limitations. We

showed that eDNA metabarcoding is capable of rapidly capturing a large proportion of the fish diversity of a complex ecosystem within a short sampling period, including taxa of conservation importance. However, the absence of a comprehensive and reliable genetic database for comparison, the transportation of exogen eDNA by water movements from surrounding areas, and the specificity of primers need to be recognised as limitations and addressed. Overall, this study confirms that the fish diversity of Merbok Estuary is extremely rich requiring more intensive and extensive studies to fully document.

Methods

Ethics statement

The study was conducted following the relevant national and international guidelines. Since the experiments were performed only with water samples and did not involve any endangered/ protected fish species or experiments on live specimens, an ethics statement is not required for this study. This study was carried out following the recommendations and approval by the Universiti Sains Malaysia Animal Ethics Committee. All methods were performed in accordance with the relevant guidelines and regulations.

Merbok Estuary capture records

A compiled checklist of fish species in Merbok Estuary was constructed from our previous and other recent surveys[7,12,16]. We also included their conservation status as according to the International Union for Conservation of Nature (IUCN) Red List (https://www.iucnredlist.org/), habitat, pelagic zone inhabited, and migration type (where available).

Study sites and water sample collections

Field samplings were conducted at Merbok Estuary, Kedah, northwest Peninsular Malaysia (Figure 1). The Merbok Estuary spans approximately 30 km with a catchment area of about 550 km2 and supports approximately 40 km2 of mangrove forest12. The sampling activity was conducted on 21st and 22nd February 2018, which coincides with the dry season in this region with estimated monthly rainfalls between 110 mm and 170 mm74. The sampling sites along the eastuary were categorised into three zones: Zone A (Merbok Estuary; KM 1.0 – KM 10.0 from the open ocean); Zone B (Midstream Merbok River (KM 10.0 – KM 20.0); and Zone C (KM 20.0 – KM 30.0). Zone A covers the area of the estuary and coastal beaches. Two sand beaches – Pantai Tanjung Dawai and Pantai Merdeka are located within this area. These areas are impacted by tourism activities, commercial fisheries, and local housing. Located at the river midstream, the vegetation in Zone B is the least impacted ecosystem with minimal anthropogenic disturbance (although impacted by polluted waters from the upstream), with most intertidal zones covered by pristine mangroves. Zone C is the most disturbed area with apparent pollution, evident by the observation of floating debris and oily substance from the probable source of the nearby residential and agricultural areas. One of the main tourist attraction sites, the Semeling Jetty Tourist Complex, is located in this area.

A total of 54 sampling sites, 18 from each zone, were surveyed (Table 1). Water samples were collected during high tide from three positions along a horizontal transect of the river: one from each bank of the estuary, and the third from the middle of the estuary. Such design is postulated to maximise the spatial coverage of the collection area. Using sterile disposable plastic bottles, one litre of water was collected at a depth of 1.0 m (from the water surface) at each site. All sampling equipment, including the sample bottles were cleaned using 10% of commercial bleach (~7.4% sodium hypochlorite), rinsed with distilled water, and wiped clean with 70% ethanol to minimise cross-site and exogenous DNA contamination. Water temperature (data not shown) and salinity (Table 1) were measured with a multiparameter instrument (YSI Pro 1030). All water samples were packed in individual plastic bags and immediately stored on ice for transport.

Laboratory contamination control

Filtration, DNA extraction, pre-PCR, and post-PCR procedures were conducted in four separate facilities, where no other environmental or organismal samples were processed. Working surfaces were decontaminated before use with a 10% bleach solution. Only sterilised consumables (e.g. filter capsules, syringes, gloves, tubes, and tips) were used throughout the experiment, and pipettes were decontaminated using 10% bleach solution and 70% ethanol before every use.

Filtration and DNA extraction

All water samples were filtered using Sterivex™-GV Sterile Vented Filter Unit, 0.22 µm polyethersulfone membrane, within 12 hours of sampling. For each of the three designated zones, one negative control (1 L of pure water) was included and filtered to monitor contamination during the filtration and subsequent extraction steps. After filtration, all membranes were individually stored in a sterile sample bag with silica beads and preserved at -20°C until extraction. DNA was extracted from the filter membrane using a DNeasy Blood and Tissue Kit (Qiagen) with slight modifications from the manufacturer's protocol: 540 µL of ATL lysis buffer and 60 µL of Proteinase K were added to each sample, and the incubation at 56°C was extended to 3 hours. The final elution was done in 80.0 µL volume. Then, all three filtered negative controls were extracted in the same way as the field samples. The extracted DNA was then transferred into a new labelled tube, secured with parafilm, and stored at -20°C until further use. The DNA extract concentration was quantified using Qubit dsDNA HS Assay (Invitrogen, USA).

Library preparation and amplicon sequencing

PCR metabarcoding assays were employed using two genetic markers, the COI and 12S rRNA genes. The COI gene was amplified using the metazoan universal primers mlCOIintF[74] and jgHCO2198[75], which rendered amplicons of 313 base pairs (bp) in size (Table 2). Meanwhile, the 12S gene was amplified using the primer pair MiFish-U[67], targeting teleost fish which rendered amplicons of approximately 219 bp in size (Table 2). A two-step PCR was conducted to prepare the libraries for Illumina MiSeq sequencing. Contamination was monitored using negative controls for all experiments.

For preparing the COI and 12S library, we adapted the "16S Metagenomic Sequencing Library Preparation Protocol"[76]. The first PCR was performed to amplify both regions with an overhanging linker sequence (Table 2) for each Nextera XT index (Illumina, USA). PCR amplifications were done in triplicates for each sample. The PCR mixture for COI (20 µL) contained 10.0 µL of KAPA HiFi HotStart ReadyMix (2X), 1.0 µL of each primer (5.0 µM), 0.3 µL of BSA (20 µg/µL), 5.7 µL nuclease-free water, and 2.0 µL of DNA template. The PCR reaction started with denaturation at 95°C for 10 minutes followed by 35 cycles at 94°C for 1 minute, 46°C for 1 minute, and 72°C for 1 minute, and a final extension at 72°C for 5 minutes. Meanwhile, the 12S PCR mixture (20 µL) contained 10.0 µL of KAPA HiFi HotStart ReadyMix (2X), 1.2 µL of each primer (5.0 µM), 4.1 µL nuclease-free water, and 3.5 µL of DNA template. The PCR cycle profile was set with an initial 3 minutes denaturation at 95°C, followed by 35 cycles of 20 seconds denaturation at 98°C, 25 seconds annealing at 66°C, and 17 seconds extension at 72°C, and a final 5 minutes extension at 72°C. Successful PCR replicates were pooled and purified using the AMPure XP kit (Beckman Coulter, USA).

All samples were quantified using Qubit dsDNA HS Assay (Invitrogen, USA). In addition to the 54 individual samples, equimolar amounts of PCR products from each sample deriving from each site were then pooled before the second index PCR. Five extra pooled samples were added, resulting in a total of 59 samples. Pooling strategy of the additional five samples is described in Supplementary Table 1. These five additional samples were included to increase the MOTU discovery from the collected individual samples. All samples were then amplified in the second index PCR following the Nextera XT index kit (Illumina, USA) protocol. All samples were also purified using the AMPure XP kit (Beckman Coulter, USA) and validated using Agilent TapeStation 4200. The library concentration and quality were measured by staining procedure (Quanti-iT™ Pico Green dsDNA Assay, Invitrogen USA) and spectrophotometry (Nanodrop, Thermo Scientific, USA). Both COI and 12S amplicon libraries were separately sequenced on an Illumina MiSeq platform (Illumina, Singapore) with COI amplicons running at 2 x 300 bp paired-end format and the 12S amplicons at 2 x 150 bp paired-end format, following the manufacturer's instructions.

Bioinformatic and data analyses

MiSeq read quality for all samples was assessed using FastQC[77] with summary statistics for each run visualised in MultiQC[78]. Individual sequencing reads were demultiplexed and trimmed with quality filtering using BBDuk within the BBTools package (https://sourceforge.net/projects/bbmap). Sequences were further processed in USEARCH v11.0.667 (https://www.drive5.com/usearch/). Pre-processed reads were dereplicated, and all low abundance sequence clusters (<1% of the total number of unique sequences) were removed from the subsequent analyses. Finally, the dereplicated sequences were clustered into molecular operational taxonomic units (MOTUs) at 97% threshold using UPARSE[79] as implemented in USEARCH. Within the same analysis tool, chimeric reads were screened and removed. MOTUs represented by a single or double sequence(s) (i.e. singletons and doubletons) were discarded. BLASTn was used in querying the MOTUs against the NCBI GenBank database[48]. Each MOTU was assigned to a unique species based on sequence similarity greater than 97%, a cut-off threshold that minimises erroneous taxonomic assignments[67].

The BLAST top hits (those with the highest identity with query sequence) were applied to each representative MOTU for species assignment. Stringent sequence and taxon filtering were employed to generate a dataset with high certainty. In order to reduce the possibility of introducing false-positive results, MOTUs represented by less than 0.02% of the generated reads were identified and discarded because they could be the consequence of contamination[36,80,81]. After removing low-frequency noise, spurious species assignments were inspected against the negative control annotation. The sequence abundance of species detected in the negative control was subtracted from the field samples[36,82]. All the remaining MOTUs assigned to species were then manually evaluated. All MOTUs with detection hits of taxa outside of our targeted taxonomic groups of bony fish and elasmobranchs were omitted. Species identities were then cross-checked to the species distribution reported by FishBase (https://www.fishbase.se/), IUCN (https://www.iucnredlist.org/), Eschmeyer's Catalog of Fishes[83], and Sharks of the World Illustrated Guide[84]. Marine or freshwater species that are unlikely to inhabit the study areas or not previously recorded in the region (taxon exotic to the study area) were identified as false positive[36]. Bioinformatic analyses including BBTools, USEARCH, UPARSE, and QIIME were run using Unix shell script in a high-performance computing (HPC) workstation.

Statistical analyses

Statistical analyses were performed in RStudio version 1.4.1106[85] using several packages. Alpha (α) diversity analyses were carried out using the phyloseq package[86]. The MOTU bar plots and heatmaps based on the relative abundances (read counts) were computed with normalised datasets for both assays (COI = 2730; 12S = 113,809). MOTU richness (α-diversity) patterns among the three defined zones (Zones A, B, and C) and the pooled samples were based on non-parametric Chao1 and Shannon indices. Sample-based MOTU accumulation curve was plotted with the specaccum function implemented in the vegan package[87] to compare species richness detected from both assays. Beta (β) diversity was calculated using Bray-Curtis metric (function vegdist) within the vegan package. In each sample, read numbers of MOTUs were transformed into fourth square root values and translated into relative abundances for the Bray-Curtis dissimilarity calculations. Biodiversity patterns across the datasets were estimated with the non-metric multidimensional (NMDS) ordination using the vegan function metaMDS. Permutational multivariate analysis of variance (PERMANOVA) was calculated with 999 free permutations with function adonis in vegan to assess community composition divergence with zones and salinity level defined as factors. Plots from all performed analyses were visualised using ggplot2 package[88].

Declarations

Data availability

All raw sequences used in this study will be uploaded to Dryad repository upon manuscript acceptance.

Acknowledgements

The Ministry of Higher Education Malaysia supported this study through Fundamental Research Grant Scheme (FRGS) with project code: FRGS/1/2016/STG05/USM/02/2. We sincerely appreciate the assistance from Firdaus Ridzuan, Siti Amalia Aisyah, Siti Zuliana, and Norli Fauzani in the sample collection. We thank the local fishermen of Merbok Estuary and the staffs of Penang Fisheries Research Institute for their logistic support in the sampling activities. The authors would also like to thank Dr. Norfatimah Mohamed Yunus for her helpful guidance in preparing the eDNA library.

Author contributions statement

D.H.Z.A., N.A.M.A., S.A.M.N., and S.L. design and conception of experiments. D.H.Z.A., N.A.M.A., M.A.R. data collection and samples processing. D.H.Z.A. bioinformatics, statistical analysis, and data visualisation. S.A.M.N. and N.A.M.A. resources and funding acquisition. D.H.Z.A. wrote the first version with the collaboration of all authors. All authors reviewed and approved the final version.

Additional information

Supplementary information accompanies this paper at (doi address)

Competing Interest: The authors declare no competing interest.

References

1. Levin, L. A. *et al.* The function of marine critical transition zones and the importance of sediment biodiversity. *Ecosystems* **4**, 430-451 (2001).
2. Wagner, G. M. & Sallema-Mtui, R. in *Estuaries: A Lifeline of Ecosystem Services in the Western Indian Ocean Estuaries of the World* (eds S. Diop, P. Scheren, & J. Machiwa) 183-207 (2016).
3. Brown, C. J. *et al.* The assessment of fishery status depends on fish habitats. *Fish and Fisheries* **20**, 1-14 (2019).
4. De La Morinière, E. C., Pollux, B., Nagelkerken, I. & Van der Velde, G. Post-settlement life cycle migration patterns and habitat preference of coral reef fish that use seagrass and mangrove habitats as nurseries. *Estuarine, Coastal and Shelf Science* **55**, 309-321 (2002).
5. Branton, M. & Richardson, J. S. Assessing the value of the umbrella-species concept for conservation planning with meta-analysis. *Conservation Biology* **25**, 9-20 (2011).
6. Dudgeon, D. *et al.* Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews* **81**, 163-182 (2006).
7. Zainal Abidin, D. H. *et al.* DNA-based taxonomy of a mangrove-associated community of fishes in Southeast Asia. *Scientific Reports* **11**, 1-15, doi:10.1038/s41598-021-97324-1 (2021).
8. Gauthier, G. *et al.* Long-term monitoring at multiple trophic levels suggests heterogeneity in responses to climate change in the Canadian Arctic tundra. *Philosophical Transactions of the Royal Society B: Biological Sciences* **368**, 20120482 (2013).
9. Valentini, A. *et al.* Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology* **25**, 929-942 (2016).
10. Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853 (2000).
11. Chong, V. C., Lee, P. K. & Lau, C. M. Diversity, extinction risk and conservation of Malaysian fishes. *Journal of fish biology* **76**, 2009-2066, doi:10.1111/j.1095-8649.2010.02685.x (2010).
12. Zainal Abidin, D. H. *et al.* Ichthyofauna of Sungai Merbok Mangrove Forest Reserve, northwest Peninsular Malaysia, and its adjacent marine waters. *Check List* **17**, 601 - 631, doi:10.15560/17.2.601 (2021).
13. Ong, J. *et al.* in *Hutan paya laut Merbok, Kedah: Pengurusan hutan, persekitaran fizikal dan kepelbagaian flora.* Vol. 23 *Siri kepelbagaian biologi hutan* (ed Ku Aman KA Abd Rahim AR, Abu Hassan MN, Abdullah M, Nor Hazliza MB, Latiff A) 21–33 (Jabatan Perhutanan Semenanjung Malaysia, 2015).
14. Hookham, B., Shau-Hwai, A. T., Dayrat, B. & Hintz, W. A baseline measure of tree and gastropod biodiversity in replanted and natural mangrove stands in Malaysia: Langkawi Island and Sungai Merbok. *Tropical Life Sciences Research* **25**, 1 (2014).
15. Jamaluddin, J. A. F. *et al.* DNA barcoding of shrimps from a mangrove biodiversity hotspot. *Mitochondrial DNA Part A* **30**, 618-625, doi:10.1080/24701394.2019.1597073 (2019).
16. Mansor, M., Mohammad-Zafrizal, M., Nur-Fadhilah, M., Khairun, Y. & Wan-Maznah, W. Temporal and spatial variations in fish assemblage structures in relation to the physicochemical parameters of the Merbok estuary, Kedah. *Journal of Natural Sciences Research* **2**, 110-127 (2012).
17. Alshari, N. F. M. A. H. *et al.* Metabarcoding of Fish Larvae in the Merbok River Reveals Species Diversity and Distribution Along its Mangrove Environment. *Zoological Studies* **60**, 60-76, doi:10.6620/ZS.2021 (2021).
18. Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J.-C. & Altermatt, F. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. *Nature Communications* **7**, 1-9 (2016).
19. Hupało, K. *et al.* An urban Blitz with a twist: rapid biodiversity assessment using aquatic environmental DNA. *Environmental DNA* **3**, 200-213 (2020).
20. Bohmann, K. *et al.* Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution* **29**, 358-367 (2014).
21. Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L. H. Environmental DNA. *Molecular Ecology* **21**, 1789-1793 (2012).
22. Ahn, H. *et al.* Evaluation of fish biodiversity in estuaries using environmental DNA metabarcoding. *PloS One* **15**, e0231127 (2020).
23. Polanco F, A. *et al.* Detecting aquatic and terrestrial biodiversity in a tropical estuary using environmental DNA. *Biotropica* **53**, 1606-1619 (2021).
24. Zhang, H., Yoshizawa, S., Iwasaki, W. & Xian, W. Seasonal fish assemblage structure using environmental DNA in the Yangtze Estuary and its adjacent waters. *Frontiers in Marine Science* **6**, 515, doi:10.3389/fmars.2019.00515 (2019).

25. Stat, M. *et al.* Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Scientific Reports* **7**, 1-11 (2017).
26. West, K. *et al.* Large-scale eDNA metabarcoding survey reveals marine biogeographic break and transitions over tropical north-western Australia. *Diversity and Distributions* **27**, 1942-1957 (2021).
27. Hallam, J., Clare, E. L., Jones, J. I. & Day, J. J. Biodiversity assessment across a dynamic riverine system: A comparison of eDNA metabarcoding versus traditional fish surveying methods. *Environmental DNA* **3**, 1247-66 (2021).
28. Seymour, M. *et al.* Environmental DNA provides higher resolution assessment of riverine biodiversity and ecosystem function via spatio-temporal nestedness and turnover partitioning. *Communications Biology* **4**, 1-12 (2021).
29. Aglieri, G. *et al.* Environmental DNA effectively captures functional diversity of coastal fish communities. *Molecular Ecology* **30**, 3127-3139 (2021).
30. Fujii, K. *et al.* Environmental DNA metabarcoding for fish community analysis in backwater lakes: A comparison of capture methods. *PLoS One* **14**, e0210357 (2019).
31. Lecaudey, L. A., Schletterer, M., Kuzovlev, V. V., Hahn, C. & Weiss, S. J. Fish diversity assessment in the headwaters of the Volga River using environmental DNA metabarcoding. *Aquatic Conservation: Marine and Freshwater Ecosystems* **29**, 1785-1800 (2019).
32. Zou, K. *et al.* eDNA metabarcoding as a promising conservation tool for monitoring fish diversity in a coastal wetland of the Pearl River Estuary compared to bottom trawling. *Science of The Total Environment* **702**, 134704 (2020).
33. Klymus, K. E., Marshall, N. T. & Stepien, C. A. Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *Plos One* **12**, 24, doi:10.1371/journal.pone.0177643 (2017).
34. Wilson, C. *et al.* Tracking ghosts: combined electrofishing and environmental DNA surveillance efforts for Asian carps in Ontario waters of Lake Erie. *Manag. Biol. Invasion* **5**, 225-231, doi:10.3391/mbi.2014.5.3.05 (2014).
35. Alexander, J. B. *et al.* Development of a multi-assay approach for monitoring coral diversity using eDNA metabarcoding. *Coral Reefs* **39**, 159-171, doi:10.1007/s00338-019-01875-9 (2020).
36. Port, J. A. *et al.* Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology* **25**, 527-541, doi:10.1111/mec.13481 (2016).
37. Fritts, A. K. *et al.* Development of a quantitative PCR method for screening ichthyoplankton samples for bigheaded carps. *Biological Invasions* **21**, 1143-1153 (2019).
38. Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M. & Minamoto, T. The release rate of environmental DNA from juvenile and adult fish. *PLoS One* **9**, e114639 (2014).
39. Amberg, J. J., Merkes, C. M., Stott, W., Rees, C. B. & Erickson, R. A. Environmental DNA as a tool to help inform zebra mussel, Dreissena polymorpha, management in inland lakes. *Manag. Biol. Invasion* **10**, 96 (2019).
40. Gu, Z., Gu, L., Eils, R., Schlesner, M. & Brors, B. circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811-2812, doi:10.1093/bioinformatics/btu393 (2014).
41. Zainal Abidin, D. H. & Noor Adelyna, M. A. Environmental DNA (eDNA) Metabarcoding as a Sustainable Tool of Coastal Biodiversity Assessment in *Universities as Living Labs for Sustainable Development* 211-225 (Springer, 2020).
42. Yamamoto, S. *et al.* Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports* **7**, 1-12 (2017).
43. Whitfield, A. K. Fish species in estuaries–from partial association to complete dependency. *Journal of Fish Biology* **97**, 1262-1264 (2020).
44. Carpenter, K. & Niem, V. *The living marine resources of the Western Central Pacific. Volume 5. Bony fishes part 3 (Menidae to Pomacentridae).* Vol. 5, 2791–3380 (Food and Agriculture Organization of the United Nations, 2001).
45. Carpenter, K. E. & Niem, V. *FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 6. Bony fishes part 4 (Labridae to Latimeriidae), estuarine crocodiles, sea turtles, sea snakes and marine mammals*. Vol. 6, 3381–4218 ((Food and Agriculture Organization of the United Nations, 2001).
46. Carpenter, K. E. & Niem, V. H. *The living marine resources of the Western Central Pacific: Batoid fishes, chimaera and bony fishes part 1 (Elopidae to Linophrynidae)*. Vol. 3, 1397–2068 (Food and Agriculture Organization of the United Nations, 1999).
47. Carpenter, K. E. & Niem, V. H. *The living marine resources of the Western Central Pacific. Volume 4. Bony fishes part 2 (Mugilidae to Carangidae).* Vol. 4, 2069–2790 (Food and Agriculture Organization of the United Nations, 1999).
48. Benson, D. A. *et al.* GenBank. *Nucleic Acids Research* **46**, D41-D47 (2018).
49. Pentinsaari, M., Ratnasingham, S., Miller, S. E. & Hebert, P. D. N. BOLD and GenBank revisited - Do identification errors arise in the lab or in the sequence libraries? *PloS One* **15**, e0231814-e0231814, doi:10.1371/journal.pone.0231814 (2020).
50. Ardura, A., Planes, S. & Garcia-Vazquez, E. Applications of DNA barcoding to fish landings: authentication and diversity assessment. *Zookeys*, **365**, 49-65, doi:10.3897/zookeys.365.6409 (2013).
51. Zainal Abidin, D. H. *et al.* Population genetics of the black scar oyster, Crassostrea iredalei: repercussion of anthropogenic interference. *Mitochondrial DNA Part A* **27**, 647-658 (2016).
52. Kelly, R. P. *et al.* Genetic and manual survey methods yield different and complementary views of an ecosystem. *Frontiers in Marine Science* **3**, 283 (2017).
53. Ratnasingham, S. & Hebert, P. D. BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular Ecology Notes* **7**, 355-364 (2007).
54. Vasconcelos, R. P. *et al.* Global patterns and predictors of fish species richness in estuaries. *Journal of Animal Ecology* **84**, 1331-1341 (2015).

55. Shah, A. S. R. M., Hashim, Z. H. & Sah, S. A. M. Freshwater fishes of Gunung Jerai, Kedah Darul Aman: A preliminary study. *Tropical Life Sciences Research* **20**, 59 (2009).
56. Md. Zain, K. *et al.* Fish diversity along streams in Ulu Muda Forest Reserve, Kedah, Peninsular Malaysia. *Malayan Nature Journal* **73**, 349-361 (2021).
57. Southeast Asian Fisheries Development Centre (SEAFDEC). Status and trends of sharks fisheries in South East Asia in *Malaysia Shark Fisheries* (Fisheries and Resources Monitoring System (FIRMS), Rome, 2004).
58. Zhang, S., Zhao, J. & Yao, M. A comprehensive and comparative evaluation of primers for metabarcoding eDNA from fish. *Methods in Ecology and Evolution* **11**, 1609-1625 (2020).
59. Doi, H. *et al.* Environmental DNA analysis for estimating the abundance and biomass of stream fish. *Freshwater Biology* **62**, 30-39 (2017).
60. Hayami, K. *et al.* Effects of sampling seasons and locations on fish environmental DNA metabarcoding in dam reservoirs. *Ecology and Evolution* **10**, 5354-5367 (2020).
61. Collins, R. A. *et al.* Persistence of environmental DNA in marine systems. *Communications Biology* **1**, doi:10.1038/s42003-018-0192-6 (2018).
62. Morey, K. C., Bartley, T. J. & Hanner, R. H. Validating environmental DNA metabarcoding for marine fishes in diverse ecosystems using a public aquarium. *Environmental DNA* **2**, 330-342 (2020).
63. Shaw, J. L. *et al.* Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* **197**, 131-138 (2016).
64. Siegenthaler, A. *et al.* Metabarcoding of shrimp stomach content: Harnessing a natural sampler for fish biodiversity monitoring. *Molecular Ecology Resources* **19**, 206-220, doi:https://doi.org/10.1111/1755-0998.12956 (2019).
65. Stoeckle, M. Y., Das Mishu, M. & Charlop-Powers, Z. Improved environmental DNA reference library detects overlooked marine fishes in New Jersey, United States. *Frontiers in Marine Science* **7**, 226 (2020).
66. Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* **9**, 134-147 (2018).
67. Miya, M. *et al.* 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**, 150088 (2015).
68. Mariani, S., Baillie, C., Colosimo, G. & Riesgo, A. Sponges as natural environmental DNA samplers. *Current Biology* **29**, R401-R402 (2019).
69. Bylemans, J., Gleeson, D. M., Duncan, R. P., Hardy, C. M. & Furlan, E. M. A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. *Environmental DNA* **1**, 402-414 (2019).
70. Hebert, P. D., Ratnasingham, S. & De Waard, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, S96-S99 (2003).
71. Collins, R. A. *et al.* Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution* **10**, 1985-2001 (2019).
72. Chin, A. T. *et al.* Beta diversity changes in estuarine fish communities due to environmental change. *Marine Ecology Progress Series* **603**, 161-173 (2018).
73. Sloterdijk, H. *et al.* Composition and structure of the larval fish community related to environmental parameters in a tropical estuary impacted by climate change. *Estuarine, Coastal and Shelf Science* **197**, 10-26 (2017).
74. Leray, M. *et al.* A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* **10**, 34 (2013).
75. Geller, J., Meyer, C., Parker, M. & Hawk, H. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular ecology resources* **13**, 851-861 (2013).
76. Illumina. 16S Metagenomic Sequencing Library Preparation. https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf 2013, 1–28 (2013).
77. Andrews, S. FastQC: a quality control tool for high throughput sequence data. (Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom, 2010).
78. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016).
79. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996-998 (2013).
80. Andruszkiewicz, E. A. *et al.* Biomonitoring of marine vertebrates in Monterey Bay using eDNA metabarcoding. *PLoS One* **12**, e0176343 (2017).
81. Goldberg, C. S. *et al.* Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* **7**, 1299-1307 (2016).
82. Nguyen, N. H., Smith, D., Peay, K. & Kennedy, P. Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytologist* **205**, 1389-1393 (2015).
83. Fricke, R., Eschmeyer, W. N. & Van der Laan, R. Eschmeyer's Catalog of Fishes: Genera, species, references. http://www.calacademy.org/scientists/catalog-of-fishes-family-group-names/ (2021).
84. Ebert, D. A. & Fowler, S. *Sharks of the World*. (Princeton University Press, 2013).
85. R Core Team. RStudio: integrated development for R. *RStudio, Inc., Boston, MA URL http://www. rstudio. com* **42**, 14 (2015).

86. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS One* **8**, e61217 (2013).
87. Oksanen, J. *et al.* Package 'vegan'. *Community Ecology Package* **2**, 1-295 (2013).
88. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York* (2016).
89. Chakrabarty, P., Amarasinghe, T. & Sparks, J. Rediscription of ponyfishes (Teleostei: Leiognathidae) of Sri Lanka and the status of Aurigequula Fowler 1918. *Ceylon Journal of Science (Biological Sciences)* **37** (2009).
90. Rüber, L., Tan, H. H. & Britz, R. Snakehead (Teleostei: Channidae) diversity and the Eastern Himalaya biodiversity hotspot. *Journal of Zoological Systematics and Evolutionary Research* **58**, 356-386 (2020).
91. McCosker, J. E. & Psomadakis, P. N. Snake eels of the genus Ophichthus (Anguilliformes: Ophichthidae) from Myanmar (Indian Ocean) with the description of two new species. *Zootaxa* **4526**, 71-83 (2018).
92. Meisner, A. D. Phylogenetic systematics of the viviparous halfbeak genera Dermogenys and Nomorhamphus (Teleostei: Hemiramphidae: Zenarchopterinae). *Zoological Journal of the Linnean Society* **133**, 199-283 (2001).
93. Hata, H., Lavoué, S. & Motomura, H. Taxonomic status of seven nominal species of the anchovy genus *Stolephorus* described by Delsman (1931), Hardenberg (1933), and Dutt and Babu Rao (1959), with redescriptions of *Stolephorus tri* (Bleeker 1852) and *Stolephorus waitei* Jordan and Seale 1926 (Clupeiformes: Engraulidae). *Ichthyological Research* **67**, 7-38, doi:https://doi.org/10.1007/s10228-019-00697-7 (2020).
94. Hata, H. & Motomura, H. Validity of *Encrasicholina pseudoheteroloba* (Hardenberg 1933) and redescription of *Encrasicholina heteroloba* (Rüppell 1837), a senior synonym of *Encrasicholina devisi* (Whitley 1940)(Clupeiformes: Engraulidae). *Ichthyological Research* **64**, 18-28 (2017).
95. Kottelat, M. The fishes of the inland waters of Southeast Asia: a catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. *Raffles Bulletin of Zoology* **27**, 1-663 (2013).
96. Katwate, U., Kumkar, P., Britz, R., Raghavan, R. & Dahanukar, N. The identity of *Aplocheilus andamanicus* (Köhler, 1906)(Teleostei: Cyprinodontiformes), an endemic Killifish from the Andaman Islands, with notes on *Odontopsis armata* van Hasselt. *Zootaxa* **4382**, 159-174 (2018).
97. Larson, H., Hadiaty, R. & Hubert, N. A new species of the gobiid fish genus *Pseudogobiopsis* (Teleostei, Gobiidae, Gobionellinae) from Indonesia. *The Raffles Bulletin of Zoology* **65**, 175-180 (2017).
98. Larson, H. K. & Hammer, M. P. A revision of the gobiid fish genus *Pseudogobius* (Teleostei, Gobiidae, Tridentigerinae), with description of seven new species from Australia and South-east Asia. *Zootaxa* **4961**, 1–85-81–85 (2021).

Tables

Table 1. List of sampling sites within the sampling transects at the three designated zones with collection metadata.

Zone	Sampling transects	Sampling sites	GPS coordinates	Salinity (ppt)	Collection date
A	A1	A1.1	05° 40′ 26.0″N, 100° 20′ 37.8″E	30.0	21/02/2018
		A1.2	05° 40′ 44.2″N, 100° 20′ 31.7″E	29.0	21/02/2018
		A1.3	05° 41′ 12.1″N, 100° 20′ 45.2″E	34.0	21/02/2018
	A2	A2.1	05° 40′ 55.9″N, 100° 21′ 29.0″E	36.0	21/02/2018
		A2.2	05° 40′ 27.3″N, 100° 21′ 28.4″E	35.0	21/02/2018
		A2.3	05° 40′ 05.3″N, 100° 21′ 15.7″E	29.0	21/02/2018
	A3	A3.1	05° 40′ 09.2″N, 100° 22′ 00.9″E	33.0	21/02/2018
		A3.2	05° 40′ 26.9″N, 100° 22′ 17.6″E	34.5	21/02/2018
		A3.3	05° 40′ 41.5″N, 100° 22′ 27.6″E	34.0	21/02/2018
	A4	A4.1	05° 39′ 59.1″N, 100° 23′ 01.7″E	33.0	21/02/2018
		A4.2	05° 39′ 52.4″N, 100° 22′ 54.3″E	36.0	21/02/2018
		A4.3	05° 39′ 40.3″N, 100° 22′ 43.5″E	36.0	21/02/2018
	A5	A5.1	05° 39′ 25.1″N, 100° 23′ 03.0″E	36.0	21/02/2018
		A5.2	05° 39′ 37.6″N, 100° 23′ 14.0″E	36.0	21/02/2018
		A5.3	05° 39′ 51.6″N, 100° 23′ 17.0″E	36.0	21/02/2018
	A6	A6.1	05° 39′ 38.1″N, 100° 23′ 48.4″E	36.0	21/02/2018
		A6.2	05° 39′ 28.7″N, 100° 23′ 51.2″E	37.0	21/02/2018
		A6.3	05° 39′ 19.6″N, 100° 23′ 50.5″E	34.0	21/02/2018
B	B1	B1.1	05° 38′ 35.6″N, 100° 23′ 59.8″E	28.5	21/02/2018
		B1.2	05° 38′ 35.4″N, 100° 24′ 07.7″E	35.0	21/02/2018
		B1.3	05° 38′ 38.0″N, 100° 24′ 23.6″E	30.0	21/02/2018
	B2	B2.1	05° 38′ 10.0″N, 100° 24′ 20.7″E	36.0	21/02/2018
		B2.2	05° 38′ 01.9″N, 100° 24′ 17.8″E	36.0	21/02/2018
		B2.3	05° 37′ 54.2″N, 100° 24′ 21.0″E	37.0	21/02/2018
	B3	B3.1	05° 38′ 11.3″N, 100° 24′ 56.5″E	37.0	21/02/2018
		B3.2	05° 38′ 19.0″N, 100° 24′ 54.6″E	36.0	21/02/2018
		B3.3	05° 38′ 29.4″N, 100° 24′ 51.9″E	36.0	21/02/2018
	B4	B4.1	05° 38′ 35.9″N, 100° 25′ 34.2″E	35.0	21/02/2018
		B4.2	05° 38′ 27.9″N, 100° 25′ 35.1″E	35.0	21/02/2018
		B4.3	05° 38′ 21.1″N, 100° 25′ 33.0″E	34.0	21/02/2018
	B5	B5.1	05° 38′ 01.6″N, 100° 26′ 00.7″E	31.0	22/02/2018
		B5.2	05° 38′ 06.4″N, 100° 26′ 04.2″E	30.0	22/02/2018
		B5.3	05° 38′ 14.2″N, 100° 26′ 02.9″E	29.0	22/02/2018
	B6	B6.1	05° 38′ 42.9″N, 100° 26′ 30.1″E	31.0	22/02/2018
		B6.2	05° 38′ 37.2″N, 100° 26′ 33.2″E	32.0	22/02/2018
		B6.3	05° 38′ 36.5″N, 100° 26′ 39.9″E	32.0	22/02/2018
C	C1	C1.1	05° 39′ 28.5″N, 100° 27′ 00.2″E	32.0	22/02/2018
		C1.2	05° 39′ 27.1″N, 100° 26′ 53.9″E	31.5	22/02/2018
		C1.3	05° 39′ 26.9″N, 100° 26′ 49.2″E	32.0	22/02/2018
	C2	C2.1	05° 40′ 12.5″N, 100° 27′ 03.2″E	32.0	22/02/2018
		C2.2	05° 40′ 08.0″N, 100° 27′ 06.0″E	33.0	22/02/2018
		C2.3	05° 40′ 03.1″N, 100° 27′ 07.8″E	31.5	22/02/2018
	C3	C3.1	05° 40′ 48.6″N, 100° 27′ 26.8″E	31.0	22/02/2018

	C3.2	05° 40′ 50.3″N, 100° 27′ 23.3″E	31.0	22/02/2018
	C3.3	05° 40′ 53.9″N, 100° 27′ 21.4″E	30.0	22/02/2018
C4	C4.1	05° 41′ 01.8″N, 100° 28′ 08.2″E	30.0	22/02/2018
	C4.2	05° 40′ 59.3″N, 100° 28′ 07.7″E	30.0	22/02/2018
	C4.3	05° 40′ 57.5″N, 100° 28′ 08.3″E	28.0	22/02/2018
C5	C5.1	05° 41′ 13.9″N, 100° 28′ 50.6″E	30.0	22/02/2018
	C5.2	05° 41′ 14.9″N, 100° 28′ 51.1″E	30.0	22/02/2018
	C5.3	05° 41′ 18.8″N, 100° 28′ 51.8″E	30.0	22/02/2018
C6	C6.1	05° 41′ 20.0″N, 100° 29′ 15.6″E	28.5	22/02/2018
	C6.2	05° 41′ 20.6″N, 100° 29′ 17.6″E	29.0	22/02/2018
	C6.3	05° 41′ 21.3″N, 100° 29′ 19.3″E	29.0	22/02/2018

Table 2. Illumina adapters and primers used in the metabarcoding assay.

	Oligonucleotide sequence (5′– 3′)	Reference
Adapters		
Forward overhang	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus-specific sequence]	76
Reverse overhang	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[locus-specific sequence]	
Primers		
COI		
mlCOIintF	GGWACWGGWTGAACWGTWTAYCCYCC	74
jgHCO2198	CATAGTGGGGTATCTAATCCCAGTTTG	75
12S		
MiFish-U-F	GTCGGTAAAACTCGTGCCAGC	67
MiFish-U-R	CATAGTGGGGTATCTAATCCCAGTTTG	

Table 3. Fish taxa (at the order, family and species levels) detected from eDNA by COI and 12S metabarcoding assays.

Order	Family	Species	Common name	IUCN status[a]	Habitat[b]	Pelagic zone[c]	Migration[d]	Notes	Me CR[e]
Carcharhiniformes	Carcharhinidae	*Carcharhinus amblyrhynchos*	Blacktail reef shark	EN	M	R	O	Endangered	
		Carcharhinus melanopterus	Blacktip reef shark	NT	B, M	R	A	Near threatened	
Myliobatiformes	Dasyatidae	*Brevitrygon walga*	Scaly whipray	NT	M	D		Near threatened	X
	Gymnuridae	*Gymnura poecilura*	Long-tailed butterfly ray	NT	M	D		Near threatened	X
Acanthuriformes	Siganidae	*Siganus fuscescens*	Mottled spinefoot	LC	M	R	O		X
		Siganus guttatus	Orange-spotted spinefoot	LC	M, B	R			X
		Siganus sutor	Shoemaker spinefoot	LC	B, M	R			
	Leiognathidae	*Photopectoralis bindus*	Shortnose ponyfish	NE	B, M	D	AM		X
		Leiognathus equula	Common ponyfish	LC	F, B	D	AM		X
		Deveximentum ruconius	Deep pugnose ponyfish	NE	F, B, M	D	AM		X
		***Deveximentum* sp.**	Ponyfish					BLAST detection as *Leiognathus* cf. *ruconius*. Refer to Note[1].	
		Photopectoralis bindus	Orangefin ponyfish	NE	B, M	D	A		
	Lobotidae	*Lobotes surinamensis*	Tripletail	LC	B, M	B	O		
	Scatophagidae	*Scatophagus argus*	Spotted scat	LC	F, B, M	R	AM		X
Anabantiformes	Anabantidae	*Anabas testudineus*	Climbing perch	LC	F, B	D	P		
	Channidae	**Channa limbata*	Dwarf snakehead	LC	F	B		BLAST detection as *Channa gachua*. Refer to Note[2].	
		Channa striata	Striped snakehead	LC	F, B	B			
	Osphronemidae	*Trichogaster lalius*	Dwarf gourami	LC	F	B			
		Trichopodus trichopterus	Three spot gourami	LC	F	B			
Anguilliformes	Anguillidae	*Anguilla bicolor*	Indonesian shortfin eel	NT	F, B, M	D	C	Near threatened	
	Muraenidae	*Gymnothorax reticularis*	Moray eel	NE	M	D			
	Ophichthidae	*Neenchelys buitendijki*	Fintail serpent eel	NE	M	D			
		**Ophichthus lithinus*	Evermann's snake eel	NE	M	B		BLAST detection as *Ophichthus evermanni*. Refer to Note[3].	
Aulopiformes	Synodontidae	*Harpadon nehereus*	Bombay-duck	NT	B, M	B	O	Near threatened	
		Saurida undosquamis	Brushtooth lizardfish	LC	M	R	A		
Batrachoidiformes	Batrachoididae	*Batrachomoeus trispinosus*	Three-spined frogfish	NE	B, M	R			X

Beloniformes	Adrianichthyidae	*Oryzias melastigma*	Marine medaka	LC	F, B	B			
		Oryzias javanicus	Javanese ricefish	LC	F, B	BP			X
	Belonidae	*Ablennes hians*	Flat needlefish	LC	B, M	R	O		
		Tylosurus crocodilus	Hound needlefish	LC	M	R	O		
	Exocoetidae	*Parexocoetus brachypterus*	Sailfin flyingfish	NE	M	N	O		
	Zenarchopteridae	*Zenarchopterus buffonis*	Buffon's river-garfish	NE	B, M	R			X
		**Dermogenys collettei*	Halfbeak	LC	F, B, M	N		BLAST detection as *Dermogenys pusilla*. Refer to Note[4].	
		Zenarchopterus dunckeri	Duncker's river garfish	NE	B	P			
Blenniiformes	Blenniidae	*Omobranchus punctatus*	Muzzled blenny	LC	B, M	B			
Carangiformes	Carangidae	*Atule mate*	Yellowtail scad	LC	B, M	R			X
		Scomberoides tol	Needlescaled queenfish	LC	B, M	R			X
		Alepes vari	Herring scad	LC	B, M	N			
		Carangoides praeustus	Brownback trevally	LC	M	D	O		
		Decapterus macrosoma	Shortfin scad	LC	M	R			
		Selar crumenophthalmus	Bigeye scad	LC	M	R			
		Caranx ignobilis	Giant trevally	LC	B, M	R			X
		Megalaspis cordyla	Torpedo scad	LC	M	R			X
		Scomberoides commersonnianus	Talang queenfish	LC	B, M	R	AM		X
		Carangoides malabaricus	Malabar trevally		M	R	A		
		Caranx melampygus	Bluefin trevally	LC	B, M	R			
		Decapterus maruadsi	Japanese scad	LC	M	R			
		Scomberoides lysan	Doublespotted queenfish	LC	B, M	R			
	Coryphaenidae	*Coryphaena hippurus*	Common dolphinfish	LC	B, M	N	O		
	Polynemidae	*Eleutheronema tetradactylum*	Fourfinger threadfin	NE	F, B, M	N	AM		X
		Polydactylus sextarius	Blackspot threadfin	NE	B, M	D	A		
	Latidae	*Lates calcarifer*	Barramundi	LC	F, B, M	D	C		X
	Toxotidae	*Toxotes chatareus*	Spotted archerfish	LC	F, B	P	A		
	Cynoglossidae	*Cynoglossus bilineatus*	Fourlined tonguesole	LC	B, M	D			X
		Cynoglossus lingua	Long tongue sole	LC	F, B, M	D	AM		X
		Cynoglossus	Speckled	LC	F, B, M	D			X

		puncticeps	tonguesole						
	Psettodidae	*Psettodes erumei*	Indian halibut	DD	M	D			
	Soleidae	*Solea ovata*	Ovate sole	LC	M	D			
		Zebrias quagga	Fringefin zebra sole	LC	M	D	A		
Centrarchiformes	Cirrhitidae	*Oxycirrhites typus*	Longnose hawkfish	LC	M	R	NM		
	Terapontidae	*Terapon jarbua*	Jarbua terapon	LC	F, B, M	D	C		X
		Terapon theraps	Largescaled terapon	LC	F, B, M	R			X
Cichliformes	Cichlidae	*Oreochromis niloticus*	Nile tilapia	LC	F, B	B	P	Invasive	
		Oreochromis sp. 1	Tilapia						
		Oreochromis sp. 2	Tilapia						
Clupeiformes	Chirocentridae	*Chirocentrus dorab*	Dorab wolf-herring	LC	B, M	R	A		
	Clupeidae	*Anodontostoma chacunda*	Chacunda gizzard shad	LC	F, B, M	N	AN		X
		Escualosa thoracata	White sardine	LC	F, B, M	N	AM		X
		Amblygaster clupeoides	Bleeker smoothbelly sardinella	LC	M	R			
	Engraulidae	**Stolephorus* sp.	Anchovy					BLAST detection as *Stolephorus dubiosus*. Refer to Note[5].	
		***Encrasicholina heteroloba*	Shorthead anchovy	LC	M	R	O	BLAST detection as *Encrasicholina devisi*. Refer to Note[6].	
		***Encrasicholina pseudoheteroloba*	Shorthead anchovy	LC	M	R	O	BLAST detection as *Encrasicholina heteroloba*. Refer to Note[6].	
		Encrasicholina punctifer	Buccaneer anchovy	LC	M	R	O		
		Stolephorus baganensis	Bagan anchovy	LC	B, M	N	A		X
		Thryssa baelama	Baelama anchovy	LC	M	P			
		Thryssa kammalensis	Kammal thryssa	DD	B, M	N	O		X
	Pristigasteridae	*Ilisha melastoma*	Indian ilisha	LC	B, M	N	AM		X
Cypriniformes	Cyprinidae	*Barbodes binotatus*	Spotted barb	LC	F	B			
		Esomus metallicus	Flying barb	LC	F, B	B			
		***Mystacoleucus obtusirostris*	Masai barb	LC	F	B		BLAST detection as *Mystacoleucus marginatus*. Refer to Note[7].	
		Tor tambra	Malayan mahseer	DD	F	B			
		Rasbora pauciperforata	Redstripe rasbora	LC	F	B			
		Tor sp.	Mahseer	DD	F	B	P		

Cyprinodontiformes	Aplocheilidae	***Aplocheilus armatus*	Blue panchax	LC	F, B	B		BLAST detection as *Aplocheilus panchax*. Refer to Note[8].	
	Poeciliidae	*Poecilia vivipara*	Guppy	NE	F, B	B		Invasive	
Elopiformes	Elopidae	*Elops machnata*	Tenpounder	LC	B, M	N	O		
Gobiiformes	Eleotridae	*Butis butis*	Duckbill sleeper	LC	F, B, M	D			X
		Butis koilomatodon	Mud sleeper	NE	B, M	D	AM		X
		Butis melanostigma	Black-spotted gudgeon	NE	F, B, M	D	A		
		Butis sp.	Gudgeon						
		***Giuris margaritaceus*	Snakehead gudgeon	LC	F, B, M	D	A	BLAST detection as *Giuris margaritacea*. Refer to Note[9].	
		Oxyeleotris marmorata	Marble goby	LC	F, B	D			
		Ophiocara porocephala	Northern mud gudgeon	LC	F, B, M	A			
	Gobiidae	*Acentrogobius caninus*	Tropical sand goby	LC	B, M	D	AM		X
		Boleophthalmus boddarti	Boddart's goggle-eyed goby	LC	F, B, M	D	AM		X
		Exyrias puntang	Puntang goby	LC	B, M	R			X
		Pseudogobius fulvicaudus	Oxudercid	NE	B	D			X
		Pseudogobius olorum	Bluespot goby	NE	F, B, M	D			X
		Trypauchen vagina	Mudburrowing goby	LC	B, M	D	AM		X
		Acentrogobius janthinopterus	Robust mangrove goby	NE	F, B, M	R	A		
		Asterropteryx semipunctata	Starry goby	LC	M	R			
		Boleophthalmus pectinirostris	Great blue spotted mudskipper	NE	F, B, M	D			
		Drombus triangularis	Brown drombus	LC	F, B, M	D	A		
		***Pseudogobiopsis oligactis*	Goby	LC	F, B	D		BLAST detection as *Eugnathogobius oligactis*. Refer to Note[10].	
		Eugnathogobius variegatus	Gudgeon	LC	B, M	D			
		Mugilogobius sp.	Goby						
		Periophthalmodon schlosseri	Giant mudskipper	LC	F, B, M	D	A		
		***Pseudogobius poicilosoma*	Northern fatnose goby	LC	F, B, M	B		BLAST detection as *Pseudogobius javanicus*. Refer to Note[11].	
		Scartelaos histophorus	Walking goby	LC	B, M	D			
		Hemigobius hoevenii	Banded mulletgoby	NE	F, B, M	D	AM		X

		Stigmatogobius pleurostigma	Gudgeon	NE	F, B	B		
Gonorynchiformes	Chanidae	*Chanos chanos*	Milkfish	LC	F, B, M	B	A	
Holocentriformes	Holocentridae	*Sargocentron punctatissimum*	Speckled squirrelfish	LC	M	R		
Mugiliformes	Mugilidae	*Crenimugil crenilabis*	Fringelip mullet	LC	F, B, M	R	NM	X
		Ellochelon vaigiensis	Squaretail mullet	LC	F, B, M	D	C	X
		Planiliza subviridis	Greenback mullet	NE	F, B, M	D	C	X
		Planiliza macrolepis	Largescale mullet	LC	F, B, M	D	C	
		Mugil cephalus	Flathead grey mullet	LC	F, B, M	B	C	
		Paramugil parmatus	Broad-mouthed mullet	NE	F, B, M	D	C	
Perciformes	Gerreidae	*Gerres filamentosus*	Whipfin silver-biddy	LC	F, B, M	D	AM	X
	Apogonidae	*Ostorhinchus semilineatus*	Half-lined cardinal	DD	B, M	R		
	Labridae	*Calotomus spinidens*	Spinytooth parrotfish	LC	M	R		
		Pseudocheilinus hexataenia	Sixline wrasse	LC	M	R		
	Haemulidae	*Pomadasys kaakan*	Javelin grunter	NE	B, M	R		X
	Lutjanidae	*Lutjanus argentimaculatus*	Mangrove red snapper	LC	B, M	R	O	X
		Lutjanus decussatus	Checkered snapper	LC	M	R		
		Lutjanus rivulatus	Blubberlip snapper	LC	M	R		
		Lutjanus johnii	John's snapper	LC	B, M	R	O	X
		Lutjanus indicus	Snapper	LC	F, B, M	R		
		Lutjanus malabaricus	Malabar blood snapper	LC	B, M	R		
	Ambassidae	*Ambassis vachellii*	Vachelli's glass perchlet	NE	F, B, M	D	O	X
		Ambassis nalua	Scalloped perchlet	LC	F, B, M	D	A	
	Platycephalidae	*Platycephalus indicus*	Bartail flathead	DD	B, M	R	O	X
	Pomacentridae	*Abudefduf sordidus*	Blackspot sergeant	LC	B, M	R	NM	
		Neoglyphidodon melas	Bowtie damselfish	NE	M	R	NM	
	Sciaenidae	*Dendrophysa russelii*	Goatee croaker	LC	F, B, M	D	AM	X
		Johnius amblycephalus	Bearded croaker	LC	F, B, M	D		X
		Johnius belangerii	Belanger's croaker	LC	B, M	D	AM	X
		Pennahia anea	Donkey croaker	LC	B, M	D		X
		Chrysochir aureus	Reeve's	LC	B, M	B		

			croaker						
		Otolithes ruber	Tigertooth croaker	LC	B, M	BP	AM		X
		Pennahia ovata	Croaker	DD	M	BP			X
		Nibea soldado	Soldier croaker	LC	F, B, M	D	A		X
	Scorpaenidae	*Pterois volitans*	Red lionfish	LC	M	R			
	Serranidae	*Epinephelus bleekeri*	Duskytail grouper	DD	M	D			X
		Epinephelus coioides	Orange-spotted grouper	LC	B, M	R			X
		Epinephelus sexfasciatus	Sixbar grouper	LC	M	R			X
		Cephalopholis sonnerati	Tomato hind	LC	M	R	NM		
		Epinephelus chlorostigma	Brownspotted grouper	LC	M	R	NM		
		Epinephelus lanceolatus	Giant grouper	DD	B, M	R			
		Variola albimarginata	White-edged lyretail	LC	M	R			
		Epinephelus areolatus	Areolate grouper	LC	M	R			
		Plectropomus laevis	Blacksaddled coralgrouper	LC	M	R			
	Sillaginidae	*Sillago sihama*	Silver sillago	LC	B, M	R	AM		X
	Sphyraenidae	*Sphyraena barracuda*	Great barracuda	LC	B, M	R			X
		Sphyraena qenie	Blackfin barracuda	NE	M	R			X
	Tetrarogidae	*Trichosomus trachinoides*	Waspfish	NE	M	D			X
	Nemipteridae	*Nemipterus bipunctatus*	Delagoa threadfin bream	LC	M	D	NM		
		Nemipterus marginatus	Red filament threadfin bream	LC	M	D	NM		
	Mullidae	*Parupeneus rubescens*	Rosy goatfish	LC	M	R			
		Upeneus tragula	Freckled goatfish	LC	B, M	R	O		
Scombriformes	Gempylidae	*Rexea prometheoides*	Royal escolar	NE	M		B		
	Scombridae	*Scomberomorus commerson*	Narrow-barred Spanish mackerel	NT	M	N	O	Near threatened	
		Rastrelliger brachysoma	Short mackerel	DD	B, M	N	O		
		Rastrelliger kanagurta	Indian mackerel	DD	M	N	O		
	Trichiuridae	*Lepturacanthus savala*	Savalai hairtail	NE	M, B	BP			X
		Trichiurus japonicus	Largehead hairtail	LC	B, M	B	A		
Siluriformes	Ariidae	*Hexanematichthys sagor*	Sagor catfish	NE	B, M	D	AM		X
		Plicofollis	Flatmouth sea	LC	B, M	D	AM		X

		platystomus	catfish					
		Plicofollis argyropleuron	Longsnouted catfish	NE	B, M	D		X
		Arius jella	Blackfin sea catfish	NE	B, M	D	A	
		Nemapteryx caelata	Engraved catfish	NE	B, M	D	A	
	Plotosidae	*Plotosus canius*	Gray eel-catfish	NE	F, B, M	D	AM	X
		Plotosus lineatus	Striped eel catfish	NE	B, M	R	A	
Syngnathiformes	Syngnathidae	*Hippichthys spicifer*	Bellybarred pipefish	NE	F, B, M	D		
Tetraodontiformes	Tetraodontidae	*Dichotomyctere fluviatilis*	Green pufferfish	LC	F, B	D	P	X
		Dichotomyctere nigroviridis	Spotted green pufferfish	NE	F, B	D		X
	Triacanthidae	*Triacanthus biaculeatus*	Short-nosed tripodfish	NE	B, M	D		

Number of species: **178**

Number of genera: **127**

Number of families: **68**

Number of orders: **25**

Number of classes: **2**

[a]LC: Least Concern; VU: Vulnerable; NT: Near Threatened; EN: Endangered; NE: Not Evaluated; DD: Data Deficient

[b]M: Marine; B: Brackish; F: Freshwater

[c]D: Demersal; R: Reef-associated; BP: Benthopelagic; N: Neritic

[d]AM: Amphidromous; AN: Anadromous; C: Catadromous; O: Oceanodromous; P: Potamodromous; NM: Non-migratory

[e]Species overlapped with previous capture records[7,12,16]

Note[1] Changed *Leiognathus* cf. *ruconius* to *Devexiментum* sp.

Reason: The species "*ruconius*" is transferred to the genus "*Devexiментum*" (following[89]; although this taxonomic change is still disputed because there is no type specimen for this species). In addition, because *Devexiментum ruconius* is already listed and this list includes only distinct species, this species of *Devexiментum* must be an unidentified species.

Note[2] Changed *Channa gachua* to *Channa limbata*

Reason: Following the conclusions of species groupings and distribution by[90].

Note[3] Changed *Ophichthus evermanni* to *Ophichthus lithinus*

Reason: *Ophichthus evermanni* is currently considered a junior synonym of *Ophichthus lithinus* (see[91]).

Note[4] Changed *Dermogenys pusilla* to *Dermogenys collettei*

Reason: All populations of *Dermogenys pusilla* of Peninsular Malaysia were described as a new species, *Dermogenys collettei*, by[92].

Note[5] Changed *Stolephorus dubiosus* to *Stolephorus* sp.

Reason: *Stolephorus dubiosus* is currently considered a junior synonym of *Stolephorus baganensis* (see[93]) but *S. baganensis* is already listed; so, the reference sequence of "*Stolephorus dubiosus"* in GenBank must be from an "unidentified" species.

Note[6] Changed *Encrasicholina devisi* to *Encrasicholina heteroloba* and *Encrasicholina heteroloba* to *Encrasicholina pseudoheteroloba*

Reason: *Encrasicholina devisi* is currently considered a junior synonym of *Encrasicholina heteroloba* and *Encrasicholina pseudoheteroloba* was (until recently) misidentified as *Encrasicholina pseudoheteroloba* (see[94]).

Note[7] Changed *Mystacoleucus marginatus* to *Mystacoleucus obtusirostris*

Reason: *Mystacoleucus marginatus* is currently considered a junior synonym of *Mystacoleucus obtusirostris* (see[95]).

Note[8] Changed *Aplocheilus panchax* to *Aplocheilus armatus*

Reason: Following the taxonomic revision of [96] who recognized all population of *Aplocheilus* in Sundaland as *Aplocheilus armatus.*

Note[9] Changed *Giuris margaritacea* to *Giuris margaritaceus*

Reason: The gender of the genus name "Giuris" is masculine, therefore (according to [95]) the species name must agree in gender with the genus name and be spelled "*margaritaceus*".

Note[10] Changed *Eugnathogobius oligactis* to *Pseudogobiopsis oligactis*

Reason: following the reclassification of the species "*oligactis*" into the genus *Pseudogobiopsis*; see[97].

Note[11] Changed *Pseudogobius javanicus* to *Pseudogobius poicilosoma*

Reason: *Pseudogobius javanicus* is currently considered a junior synonym of *Pseudogobius poicilosoma*[98].

Table 4. PERMANOVA results for the zones division and site salinity levels in fish community composition. Analysis was calculated using Bray-Curtis distances for both COI and 12S metabarcoding assays.

COI	*df*	Sum of Sqs	Mean Sqs	*F*	R^2	*p*-value	
Zone	2	1.82060	0.91032	8.70450	0.25061	0.001	***
Salinity	1	0.21530	0.21526	2.05830	0.02963	0.053	
Residual	50	5.22900	0.10458		0.71976		
Total	53	7.26490			1.00000		
12S	*df*	Sum of Sqs	Mean Sqs	*F*	R^2	*p*-value	
Zone	2	0.57180	0.28590	6.52800	0.20158	0.001	***
Salinity	1	0.07502	0.07502	1.71300	0.02645	0.063	
Residual	50	2.18980	0.04380		0.77197		
Total	53	2.83662			1.00000		

***p<0.001

Figures

Figure 1

Location of sampling sites across the Merbok Estuary. Inset map shows the location of the study area within Peninsular Malaysia. Photographs of the representative sites within the three zones: (a) site A1.1, (b) site B3.2, (c) site C6.3. Map is generated using ArcMap 10.8 and edited in Adobe Photoshop CC 2019.

Figure 2

Actinopterygian and elasmobranch taxa detected from eDNA by the COI and 12S metabarcoding assays. *Circlize* plot showing assays being mapped to the 25 orders detected by eDNA metabarcoding. The purple ribbons represent overlapped detections from both COI and 12S metabarcoding assays.

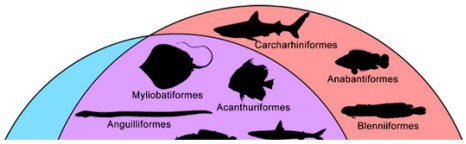

Figure 3

Venn diagram showing the number of taxa recovered from the local capture records (blue), eDNA metabarcoding assays (red), and both methods (purple) at different taxonomic ranks: (a) order, (b) family, (c) genus, and (d) species. Blown-up diagram at the top showing the orders detected by both methods.

Figure 4

Barplots showing relative read abundance in all samples per fish family; (a) COI assay and (b) 12S assay.

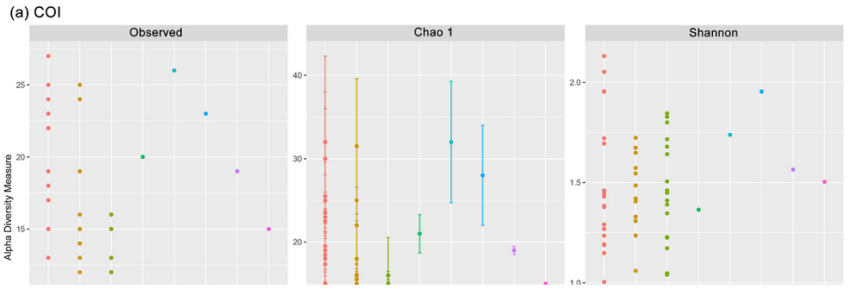

Figure 5

Alpha (α) diversity plots based Observed, Chao1, and Shannon estimator grouped by zones and sample pools: (a) COI assay and (b) 12S assay.

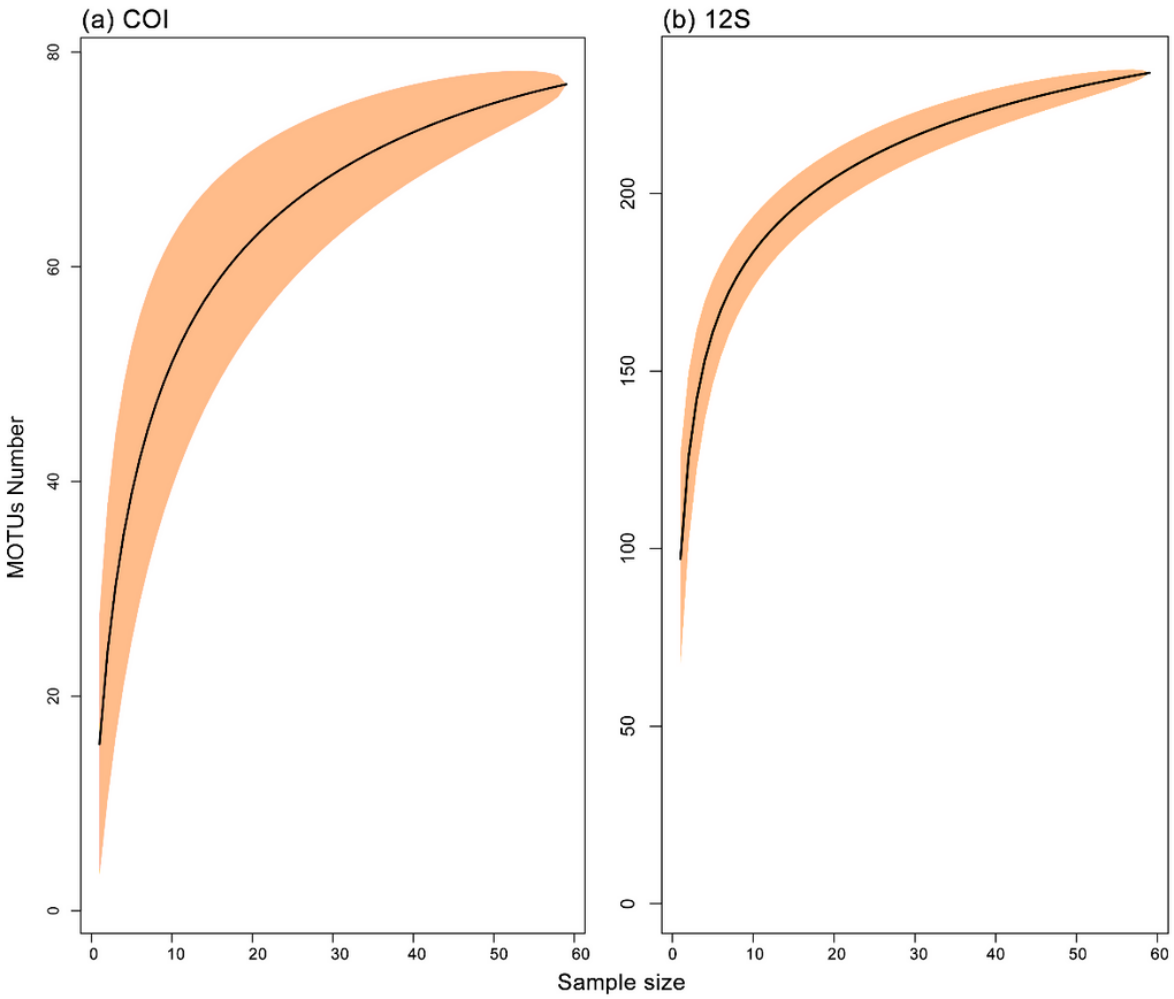

Figure 6

MOTU accumulation curves representing the number of MOTU identified in all samples analysed by eDNA metabarcoding assays: (a) COI assay and (b) 12S assay. The light-shaded area equates to the 95% confidence interval.

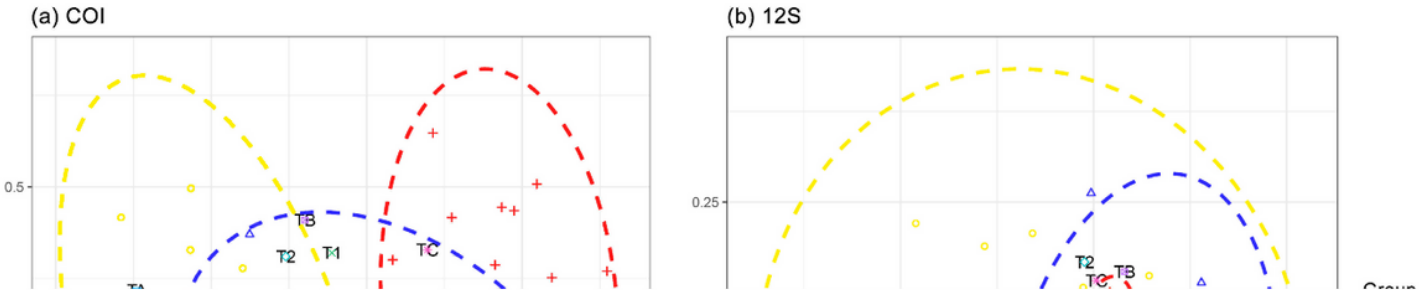

Figure 7

Nonmetric multidimensional scaling (NMDS) ordination of fish community in Merbok Estuary using the Bray-Curtis coefficient for both metabarcoding assays: (a) COI assay and (b) 12S assay. Different symbols denote individual samples analysed and the community clusterings are colour-coded (see legend). The ordination stress value is indicated at the bottom of each plot.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ScientificReportseDNASupplementaryInfoFeb2022.pdf