

# Assessing fish diversity of Merbok Estuary and adjacent marine waters by DNA taxonomy: towards building comprehensive DNA barcode reference library

Danial Hariz Zainal Abidin (✉ [geneprodigy@gmail.com](mailto:geneprodigy@gmail.com))

Centre for Global Sustainability Studies (CGSS), Level 5, Hamzah Sendut Library, Universiti Sains Malaysia, 11800, Penang

Siti Azizah Mohd. Nor

Institute of Marine Biotechnology, Universiti Malaysia Terengganu, 21030, Terengganu,

Sébastien Lavoué

School of Biological Sciences, Universiti Sains Malaysia, 11800, Penang

Masazurah A. Rahim

Fisheries Research Institute, Batu Maung, 11960, Penang

Noorul Azliana Jamaludin

Fisheries Research Institute, Kampung Acheh, 32000, Perak

Noor Adelyna Mohammed Akib

Centre for Global Sustainability Studies (CGSS), Level 5, Hamzah Sendut Library, Universiti Sains Malaysia, 11800, Penang

---

## Research Article

**Keywords:** Mangrove Forests, Biodiversity, Morphological Examination, Molecular Approach, Fish Taxa

**Posted Date:** March 6th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-273253/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The Merbok Estuary comprises one of the largest remaining mangrove forests in Peninsular Malaysia. Its value is significant as it provides both direct and indirect services to local and global communities. It also offers a unique opportunity to study the structure and functioning of mangrove ecosystems. However, its biodiversity is still partially described, limiting its research value. Recent inventories based on morphological examination, documented 138 fish species residing, frequenting or subject to entering the Merbok Estuary. Using a molecular approach, we assessed the fish diversity of the Merbok Estuary and its adjacent waters in DNA barcoding 350 specimens assignable to 135 species initially identified based on morphology. Our results revealed the presence of 140 MOTUs, 130 of them are congruent with morphology-based species delimitation. In two cases, barcodes did not permit to differentiate between two morphotypes while they unveiled cryptic diversity within six other species, calling for further taxonomic investigations. This study provides a comprehensive core-list of fish taxa in Merbok Estuary, demonstrating the advantages of combining morphological and molecular evidence to describe diverse but still poorly studied tropical fish communities. It also delivers a large DNA reference collection for brackish fishes occurring in this region which will facilitate further biodiversity-oriented research studies and management activities.

## Introduction

Estuaries and coastal wetlands which feature mangrove ecosystem are transition zones that link terrestrial and freshwater habitats with the sea<sup>1</sup>. Mangrove ecosystem delivers essential ecosystem services, including shoreline protection, nutrient production and fisheries resources. In consequence, mangrove ecosystem plays a vital role in supporting local communities' socio-economic pursuits<sup>2</sup>. Unfortunately, such crucial human-nature relationship is threatened by habitat pollution, destruction, and overfishing<sup>3</sup>. It is also impacted by other factors such as species invasion, and climate change<sup>1</sup>.

The less disturbed tropical estuaries, especially their mangrove area, generally harbour rich, diversified and complex faunal communities<sup>4</sup>, combining the presence of salinity-tolerant resident species along with regularly or occasionally frequenters. Frequenters include mainly marine species that use this ecosystem either to feed, shelter, breed, or nurse their young<sup>5</sup>. Describing and monitoring the biodiversity in these ecosystems is primordial for long term sustainability because it is considered that biodiversity ensures stability and resistance towards any disturbance or potential invasion through complex species-species interactions<sup>1</sup>. However, biodiversity is still poorly documented in many mangrove ecosystems that hampers further research on their functioning and management.

Malaysia is recognised as one of the world's biodiversity hotspots with astounding levels of diversity and endemism<sup>6,7</sup>. Considering only fishes, <sup>8</sup> reported the presence of a total of 1418 brackish and marine species in Malaysian waters, occupying various coastal habitats, including mangroves and estuaries. The remaining coastal mangroves in Malaysia (Peninsular Malaysia and Malaysian Borneo) occupy less than

6000 km<sup>2</sup> <sup>9</sup>. Along the west coast of Peninsular Malaysia, facing the Strait of Malacca, mangrove-forests often occupy estuarine banks in saline tidal areas. One of the largest remaining intact patches of mangrove forests is located within the Merbok Estuary (north-west of Peninsular Malaysia, Figure 1). The estuary was gazetted as a permanent forest reserve, the Sungai Merbok Mangrove Forest Reserve in 1951, and is the second largest mangrove forest in Peninsular Malaysia after the Larut Matang Forest Reserve. The Merbok Estuary and its surroundings are a dynamic and productive ecosystem that supports the world's highest mangrove species diversity per unit area within a contiguous habitat, with 39 of the estimated 70 true mangroves species described globally<sup>10</sup>. This area also represents important resource grounds for local populations <sup>11,12</sup>.

Due to its biological, ecological, and socio-economic importance, the Merbok Estuary has been the focus of research in the 21<sup>st</sup> century, including some biodiversity inventories (trees and gastropods<sup>13</sup>; shrimps<sup>11</sup>; fishes<sup>12,14</sup>; mangrove trees<sup>10</sup>). The latest ichthyological survey has inventoried up to 138 fish species (from 47 families) in the estuary, demonstrating a rich fish fauna<sup>15</sup>. However, because of taxonomical uncertainties using only morphological characteristics, the identifications of some species are challenging, especially for some speciose families such as Mugilidae, Gobiidae or Eleotridae<sup>15</sup>. Furthermore, cryptic diversity is frequently encountered in tropical highly biodiverse regions<sup>16,17</sup>, and it is likely that some morphology-based species hide more than one species. In Merbok as elsewhere, a precise account of species diversity is a necessary requirement for further researches and numerous studies have highlighted the complementarity between morphological and molecular approaches to reveal biodiversity<sup>18-20</sup>. To date, there is no attempt to compare morphology-based results on fish diversity with molecule-based approach in this mangrove species-rich Merbok Estuary.

Since its introduction in the past decades, DNA barcoding has emerged as the global molecular taxonomic method across fishes based on a defined molecular marker, a ~650 base pairs long fragment of the mitochondrial cytochrome oxidase I gene (COI)<sup>21</sup>. Several regional DNA barcoding studies have demonstrated its efficacy to delimitate fish species, for instance, in Australia<sup>22</sup>, South China Sea<sup>23</sup>, Indian Ocean<sup>24</sup>, and Indo-Pacific coral fishes<sup>25</sup>. DNA barcoding has proven to be a reliable method in identifying cryptic and potentially new fish species<sup>26-29</sup>, marine larval fishes<sup>30-33</sup>, and food origins<sup>34</sup>.

In this study, we assemble a large reference library of DNA barcodes of 350 fish individuals from Merbok Estuary and its adjacent waters for the purpose to describe the fish diversity in this region in providing a critical look at previous morphology-based results. Such species inventory approach has found wide applicability in a plethora of research that hinges on reliable species identification, from ecosystem health management, biodiversity conservation, and aquaculture to fishery management<sup>35,36</sup>. All of these uses pertain to the Merbok Estuary.

## Results

### Fish Diversity

A total of 350 specimens (out of 441 collected) were successfully sequenced for the COI gene, representing 135 morphological species, 94 genera, 47 families, 17 orders, and two classes, Chondrichthyes and Actinopterygii (species list shown Table 1). Two of these species (i.e. *Cryptocentrus* sp. and *Johnius* sp.) were only identified to the generic level using morphology whereas for three other species, *Dichotomyctere* cf. *fluviatilis*, *Brachygobius* cf. *kabiliensis*, and *Lagocephalus* aff. *Lunaris*, we used open nomenclature.

The most diverse orders were Perciformes (42 species representing 31.1% of the total number of species), followed by Carangiformes (21 species, 15.5%), Gobiiformes (18 species, 13.3%), and Clupeiformes (16 species, 11.9%) (Figure 2a). At the family level, Gobiidae has the highest species richness with 15 species (11.1%), followed by Carangidae (12 species, 8.9%), Engraulidae (9 species, 6.7%), and Ariidae (8 species, 5.9%) (Figure 2b). The three most diverse genera were the anchovy genus *Stolephorus*, the flatfish genus *Cynoglossus* with five species each, followed by the grouper genus *Epinephelus* with four species. According to the International Union for Conservation of Nature (IUCN) Red List, five species are “Near Threatened” (four Chondrichthyes: *Telatrygon zugei*, *Brevitrygon walga*, *Gymnura poecilura*, *Chiloscyllium indicum* and one actinopterygian: *Arius gagora*) whereas others are listed as “least concern” or “data deficient”. One recorded species is an alien invasive species (AIS), *Oreochromis mossambicus* (the Mozambique tilapia) from the African region<sup>37</sup>.

## DNA-based identification

Sequence length for all 350 generated barcodes was longer than 600 bp with no indels or stop codon detected. The nucleotide composition showed a mean percentage of 18.32% (G), 27.97% (C), 24.07% (A), and 30.7% (T). More than half of the species (56%, 76 species) were represented by multiple specimens while 59 species were represented by a single specimen (Table 1). Mean number of specimens per species was 2.59. Increment in the K2P genetic divergence was directly related to the hierarchical taxonomic relationship: within species mean divergence = 0.85% (SE=0.01), within congeners mean divergence = 16.7%, (SE=0.01) and within families mean divergence = 18.17% (SE=0) (Table 2).

Deep intraspecific K2P divergences, exceeded the standard threshold distance of 2%<sup>21,38</sup>, were observed in six species: *Eleutheronema tetradactylum* (16.66%), *Osteomugil perusii* (14.24%), *Planiliza subviridis* (13.44%), *Deveximentum indicium* (9.05%), *Gerres oyena* (4.29%) and *Lutjanus russellii* (4.12%) (Table 3). Barcoding gap analysis demonstrated that almost all species represented by multiple sequences is supported by a barcode gap (Figure 3). Notably, only one species, *D. indicium*, had its maximum intraspecific distance (9.05%) similar to its nearest neighbour distance (9.04%).

Both Bayesian Inference (BI) (Figure 4) and Maximum Likelihood (ML) (Figure S1) trees exhibited well-defined clusters at the level of orders and families, with minimal differences in topologies. The BI tree was better resolved and used to represent the species delimitation (Figure 4). The Molecular Operational Taxonomic Units (MOTU) delimitation analyses yielded variable, but relatively high number of MOTUs for each method. The RESL analysis revealed 139 MOTUs assigned to dedicated BINs. Similarly, the ABGD

analysis also identified 139 MOTUs ( $P=0.0010-0.0599$ ) within the initial partition for all substitution models (Table S2). The single-threshold GMYC analysis recognised 140 MOTUs that were taxonomically concordant with the identification by the other two analyses except for one species, *Hyporhamphus quoyi*, that partitioned into two MOTUs. Several incongruences between MOTUs and morphology-based species delimitation are highlighted in Figure 4 (red bars) involving multiple species, as presented in Table 3.

Two MOTUs (= two BINs) were shared by more than one species: *Alepes melanoptera* and *Caranx sexfasciatus* (BIN "BOLD:AAB5775") and *Dichotomyctere nigroviridis* and *Dichotomyctere cf. fluviatilis* (BIN "BOLD:AAF2344") (Table 3) as these species pairs displayed no intraspecific genetic divergence and were very closely related to their nearest neighbour (<2%).

## Discussion

### Species delimitation methods

One of the premises of DNA barcoding is the detection of the so-called "barcode gap", which can be observed when the intraspecific divergence is smaller than the interspecific mean distances<sup>39</sup>. Absence of gap between two morphological species is indicative that these species represent only different forms within one species whereas the presence of a gap within a morphological species is evidence for species-level cryptic diversity<sup>40</sup>.

Employing multiple methods and schemes in clustering (using barcode gaps) the generated DNA barcodes provide an efficient approach in identifying putative species. Even though these methods may have individual pitfalls, especially in analysing singletons, they can yield a robust outcome when combined<sup>41</sup>. Our study utilised three different methods in clustering the sequences into MOTUs, namely Refined Single Linkage (RESL), Automatic Barcode Gap Discovery (ABGD), and Generalized Mixed Yule Coalescent (GMYC). Despite different analytical assumptions supporting each method, all three methods yielded similar results: RESL and ABGD analyses identified each 139 MOTUs whereas the GMYC analysis identified 140 MOTUs within our dataset. These results demonstrate a robust pattern of MOTUs in our dataset; even the GMYC method which is known to overestimate MOTUs counts compared to other methods<sup>42</sup>, delimited only one additional MOTU. Because both RESL and ABGD analyses had closer correspondence to the number of species defined by morphological identification, we based our discussion on species account on these two methods.

Our results show that DNA barcoding (using COI gene) and morphology-based approach accurately converge on the delimitation of 131 species (about 97% of the examined species) in Merbok Estuary region. DNA barcoding approach further revealed possible cryptic diversity within six species whereas it did not detect any difference between two pairs of species, calling for taxonomic revisionary studies.

The mean conspecific K2P divergence (0.85%) was 20-fold lower than the mean congeneric divergence (16.7%). This increase in genetic divergence with increment in taxonomic levels is logical and expected<sup>35</sup>. However, both mean genetic estimates are higher than those previously recorded in other regions. Most molecular assessment of marine fishes displayed conspecific divergence within the range of 0.25-0.39% whereas congeneric divergence were within the range of 4.56-9.93%<sup>22-24,36,43</sup>, but <sup>25</sup> found similar pattern of high average conspecific and congeneric divergence within the Indo-Pacific coral reef fishes (1.06% and 15.34%, respectively).

### Taxonomic conundrum

We observed high intraspecific COI variability in six species, namely *Eleutheronema tetradactylum* (16.66%), *Osteomugil perusii* (14.24%), *Planiliza subviridis* (13.44%), *Deveximentum indicium* (9.05%), *Gerres oyena* (4.29%) and *Lutjanus russellii* (4.12%). Such high intra-specific genetic divergence suggests either misidentification or the presence of more than one species<sup>25,44</sup>. We dismiss the first possibility because the morphological examination of incriminated specimens, based on existing keys, seems consistent. Therefore, such morphology-molecules incongruence may more likely be the signal of hidden diversity. Large genetic differentiation has been reported in *E. tetradactylum* among allopatric populations within the Indian Ocean<sup>45</sup>. Our results are consistent with this study, further indicating that differentiation in this species is not only allopatrically but, also, sympatrically distributed. Recent molecular taxonomic studies on the family Mugilidae in which are included *O. perusii* and *P. subviridis*, evidenced a very high level of cryptic diversity in the Indo-West Pacific region<sup>46,47</sup>. Several valid mullet species are actually complexes of several morphologically similar species for which extensive taxonomic revisions are needed. The taxonomy of *D. indicium* (family Leiognathidae) is still in flux with continual descriptions of new species<sup>48</sup>. Our results indicate the presence of two sympatric species under *D. indicium* in Merbok Estuary. *Gerres oyena* and *L. russellii* exhibit intraspecific differentiation of lower magnitude than those observed for the first four species discussed above, although still well above the threshold of 2%. *Lutjanus russellii* natively occurs in this region<sup>49</sup> but it is also farmed in Merbok estuary. Aquaculture activities regularly import non-native seeds from various sources, with no or poor records of origins, making it difficult to determine their origins. The divergence observed within this species (4.12%) could be the consequence of the presence of both native and alien (escaped from aquaculture farms) individuals in Merbok estuary<sup>15</sup>.

Two cases of shared MOTUs between species were detected involving the pairs *Alepes melanoptera* and *Caranx sexfasciatus* (BOLD:AAB5775), and *Dichotomyctere nigroviridis* and *Dichotomyctere cf. fluviatilis* (BOLD:AAF2344). Hybridisation among closely related taxa, incomplete lineage classification of a recent, on-going speciation event, or intraspecific morphological variability could account for this observation<sup>50</sup>. Similarly, <sup>51</sup> who studied the fish diversity in the Parnaíba Delta, opined that a lack of taxonomy consensus may be observed when the rate of molecular variation does not accompany recent sympatric speciation event that lead to morphological differentiation.

## Establishment of DNA barcoding library

Precise identification of organisms is a prerequisite for assessing the biological and ecological status of an aquatic ecosystem. The current study illustrates yet another example of the complementarity of the morphological and molecular techniques to achieve this goal. DNA barcoding offers a non-invasive approach in aquatic diversity assessment and requires minimal expertise in conventional taxonomy<sup>52,53</sup>. Comprehensive DNA barcode reference library is crucial in any biodiversity assessment for providing selective autecological and biogeographic information for comparative analysis with previous assessment. Even though DNA databases like BOLD<sup>54</sup> and GenBank<sup>55</sup> are publicly available, a localised taxon-specific reference library is synoptically important as it is easier to curate and is a more practical reference for a focused site. The DNA barcodes reference library associated with voucher collections established in our study broadens the research spectrum in biological evaluation and biomonitoring effort. Future research endeavours will utilise this database to assess this habitat's ecological and health status through eDNA metabarcoding. The barcode data generated in this study will definitely contribute to the local as well as regional conservation efforts of fish diversity. Notwithstanding, to improve the resolution of the taxonomic coverage, the number of DNA barcodes for the singleton specimens should be increased through more intensive sampling and increased number of sites within the estuary.

Of the 135 species examined in this study, 61 species (~45%) were identified with high commercial value<sup>56</sup> thus, protection planning, and proper fishery management of these species are vital. Furthermore, we manage to barcode an invasive species - the Mozambique tilapia, *Oreochromis mossambicus* that were sampled within our studied site.

We DNA barcoded a rich and diverse mangrove-related fish assemblage. Of the 135 species initially identified based on morphology, barcodes of 131 species support their validity. We found hidden diversity within six species whereas the divergences between two pairs of valid species are below the interspecific threshold standard raising questions on their validity and calling for further taxonomic studies. Although certainly not complete, the establishment of a local DNA barcodes reference library is an essential step for future studies of fisheries, conservation and ecological management of this important site.

## Methods

### Ethics statement

This project was conducted according to the relevant national and international guidelines and did not involve any endangered or protected fish species. All fish specimens were either collected from the local fishermen, caught using non-invasive fishing gear by the authors, or bought from the local market. This study was carried out following the recommendations and approval by the Universiti Sains Malaysia Animal Ethics Committee.

### Sample collection

A total of 441 specimens were sampled between December 2018 to October 2019 at multiple locations along the Merbok Estuary and its vicinity (Figure 1). Specimens were collected either from local fishermen (who use the barrier-net method locally called 'pompong'), direct sampling by dip-net or bought from the major fish landing site (Kuala Muda Whispering Market). All specimens were caught within Merbok River and its adjacent waters. Samples collected from the fish landing site were retrieved from fishing vessels that operate within Zone A (from the shoreline up to 5 nautical miles) and Zone B (from 5 to 12 nautical miles)<sup>57</sup>. Information on the sampling localities (geographical coordinates) is shown in Table S1. Other collection data – dates, taxonomy and details of voucher specimens can be retrieved from the online project datasheet implemented in BOLD with project code – DBMR.

### **Sample processing and morphological identification**

A fin clip from each fresh specimen was taken and stored in 90% ethanol. Voucher specimens were fixed in 10% formalin for at least one week and then transferred into 70% ethanol for long term storage. All specimens were catalogued and deposited at the Museum of Biodiversity, Universiti Sains Malaysia.

Morphology-based species identifications and nomenclature follow<sup>15</sup>. We were unable to unequivocally assigned few specimens to a valid described species using available keys. In these cases, we used either "sp.", "cf.", or "aff."

### **Laboratory analyses**

Genomic DNA was extracted using DNeasy Blood & Tissue kit (Qiagen, Germany) following the given protocol of animal tissue DNA extraction. The purity and concentration of the isolated DNA were measured using a microvolume UV spectrophotometer (Quawell Q300, Quawell, CA) and stored at -20°C until further use. An approximately 650 bp fragment of the mitochondrial COI gene region was amplified using the combinations of the following primers previously designed by<sup>22</sup>:

FishF1-5'-TCAACCAACCACAAAGACATTGGCAC-3',

FishF2-5'-TCGACTAATCATAAAGATATCGGCAC-3',

FishR1-5'-TAGACTTCTGGGTGGCCAAAGAATCA-3' and

FishR2-5'-ACTTCAGGGTGACCGAAGAATCAGAA-3'.

Each sample was amplified in a final volume of 25 µL, containing 5.5 µL of 5x MyTaq™ Reaction Buffer Red (Bioline GmbH, Germany), 0.5 µL of each primer (100 ng/µL), 0.25 µL 5U Taq polymerase (iNtRON Biotechnology Inc., Korea), 2.5 µL of genomic DNA (50 ng/µL) and adequate nuclease-free water to complete the final reaction volume. Each amplification set was performed with the inclusion of a negative control (no template DNA) with thermal cycling conditions as follows: initial denaturation at 94°C for 4 minutes; followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 48°C for 50 seconds, and extension at 72°C for 1 minute; then a final extension at 72°C for 10 minutes. The PCR products were

then fractionated by 2% gel electrophoresis to check for successful amplification. All positive amplifications were then sent for purification and sequencing to Apical Scientific Sdn. Bhd. (Selangor, Malaysia) operating the ABI PRISM 3730XL automated sequencer and the ABI PRISM BigDye terminator cycle sequencing kit v3.1 (Applied Biosystems, Foster City, CA). Bidirectional sequencing was employed to decrease the probability of sequencing errors.

## Data analyses

Each generated chromatogram was manually screened prior to DNA alignment in MEGA X<sup>58</sup>. The sequences were proofread and independently aligned and then inspected for deletions, insertions and stop codons using the same software. All sequences have been uploaded in BOLD<sup>54</sup> and deposited in GenBank<sup>55</sup> (Accession nos. MW498499 - MW498843).

A total of 350 COI sequences were determined in this study. To assess the taxon discrimination between all specimens, pairwise genetic distances were calculated within and between species, genera, and families based on the Kimura 2-parameter (K2P) distance model<sup>59</sup> using the analytical tools available in the BOLD system platform. To depict a graphical representation of the genetic relationships of the sequences, Bayesian Inference (BI) and Maximum Likelihood (ML) analyses were run in BEAST 2<sup>60</sup> and raxmlGUI 2.0<sup>61</sup> program, respectively. The GTR+I+G substitution model was determined as the best one in PartitionFinder 2<sup>62</sup>, as implemented in the CIPRES portal<sup>63</sup>. The BI tree was constructed with the GTR+I+G substitution model, empirical base frequencies with four gamma categories, employing a relaxed lognormal clock and the birth-death model. Two Markov Chain Monte Carlo (MCMC) chains of 40 million were run independently, sampled every 1000 generations and the first 20% were discarded as burn-in. Both run performances were then assessed for convergence (ESS > 200) using Tracer 1.7.1 and combined using LogCombiner 2.4.8 before the final tree was constructed using TreeAnnotator 2.4.7, within the BEAST 2 package<sup>60</sup>. The ML tree was also built based on the GTR+I+G model with 1000 nonparametric bootstrap replicates. Both constructed trees were then viewed and edited in FigTree 1.4.4<sup>64</sup>.

Three different sequence-based methods were used to delimit the Molecular Operational Taxonomic Units (MOTUs) from the analysed sequences - (1) Refined Single Linkage (RESL), (2) Automatic Barcode Gap Discovery (ABGD), and (3) Generalized Mixed Yule Coalescent (GMYC). The first analysis was done within the BOLD platform using the RESL algorithm<sup>65</sup> to assign sequences to a dedicated Barcode Index Numbers (BIN). Next, the ABGD<sup>39</sup> analysis was run at the webserver (<https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html>) to census divergence within the analysed dataset for species delimitation. The ABGD analysis was run with the following settings: relative gap width  $X=1.0$ , intraspecific divergence ( $P$ ) values range from 0.001 to 0.0059 for all the distance metrics, while all other parameter values were kept as default. Finally, the GMYC method<sup>66</sup> was employed with the fully resolved, BI ultrametric COI tree (see above for the reconstruction method). A single-threshold GMYC analysis was run in RStudio<sup>67</sup> with the 'splits' package.

## Data availability

All the COI sequences analysed in this study are available in the BOLD system under the DBMR project. The sequences can also be retrieved in GenBank (Accession nos. MW498499 - MW498843).

## 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. Sarathchandra, C. et al. Significance of Mangrove Biodiversity Conservation in Fishery Production and Living Conditions of Coastal Communities in Sri Lanka. *Diversity* **10**, 20 (2018).
3. Brown, C. J. et al. The assessment of fishery status depends on fish habitats. *Fish and Fisheries* **20**, 1-14 (2019).
4. Kathiresan, K. & Bingham, B. L. Biology of mangroves and mangrove ecosystems. *Advances in Marine Biology* **40**, 84-254 (2001).
5. 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).
6. Asaad, I., Lundquist, C. J., Erdmann, M. V. & Costello, M. J. Delineating priority areas for marine biodiversity conservation in the Coral Triangle. *Biological Conservation* **222**, 198-211 (2018).
7. Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A. & Kent, J. Biodiversity hotspots for conservation priorities. *Nature* **403**, 853 (2000).
8. 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).
9. Wong, S. L. in *Matang Mangroves: A Century of Sustainable Management*. (Petaling Jaya, 2004).
10. 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).
11. 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).
12. 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).
13. 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).

14. Mansor, M., Najamuddin, A., Mohammad-Zafrizal, M., Khairun, Y. & Siti-Azizah, M. Length-weight relationships of some important estuarine fish species from Merbok estuary, Kedah. *Journal of Natural Sciences Research* **2**, 8-19 (2012).
15. Zainal Abidin, D. H. et al. Ichthyofauna of Sungai Merbok Mangrove Forest Reserve, northwest Peninsular Malaysia, and its adjacent marine waters. *Check List* (in press).
16. Lim, H. C., Zainal Abidin, M., Pulungan, C. P., de Bruyn, M. & Mohd Nor, S. A. DNA barcoding reveals high cryptic diversity of the freshwater halfbeak genus *Hemirhamphodon* from Sundaland. *PloS one* **11**, e0163596 (2016).
17. Mennesson, M. I., Bonillo, C., Feunteun, E. & Keith, P. Phylogeography of *Eleotris fusca* (Teleostei: Gobioidi: Eleotridae) in the Indo-Pacific area reveals a cryptic species in the Indian Ocean. *Conservation Genetics* **19**, 1025-1038 (2018).
18. Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S. & Carvalho, D. C. Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. *Genetica* **143**, 581-588 (2015).
19. Iyiola, O. A. et al. DNA barcoding of economically important freshwater fish species from north-central Nigeria uncovers cryptic diversity. *Ecology and Evolution* **8**, 6932-6951 (2018).
20. Stern, N., Rinkevich, B. & Goren, M. Integrative approach revises the frequently misidentified species of *Sardinella* (Clupeidae) of the Indo-West Pacific Ocean. *Journal of Fish Biology* **89**, 2282-2305 (2016).
21. 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).
22. Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 1847-1857 (2005).
23. Xu, L. et al. Assessment of fish diversity in the South China Sea using DNA taxonomy. *Fisheries Research* **233**, 105771 (2020).
24. Lakra, W. et al. DNA barcoding Indian marine fishes. *Molecular Ecology Resources* **11**, 60-71 (2011).
25. Hubert, N. et al. Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLoS one* **7**, e28987 (2012).
26. Adibah, A. & Darlina, M. Is there a cryptic species of the golden snapper (*Lutjanus johnii*)? *Genetics and Molecular Research* **13**, 8094-8104 (2014).
27. Bakar, A. A. et al. DNA barcoding of Malaysian commercial snapper reveals an unrecognized species of the yellow-lined *Lutjanus* (Pisces: Lutjanidae). *PloS one* **13**, e0202945 (2018).
28. Farhana, S. N. et al. Exploring hidden diversity in Southeast Asia's *Dermogenys* spp. (Beloniformes: Zenarchopteridae) through DNA barcoding. *Scientific Reports* **8**, 1-11 (2018).
29. Jaafar, T. N. A. M., Taylor, M. I., Nor, S. A. M., de Bruyn, M. & Carvalho, G. R. DNA barcoding reveals cryptic diversity within commercially exploited Indo-Malay Carangidae (Teleostei: Perciformes). *PLoS one* **7**, e49623 (2012).

30. Azmir, I., Esa, Y., Amin, S., Salwany, M. & Zuraina, M. DNA barcoding analysis of larval fishes in Peninsular Malaysia. *Journal of Environmental Biology* **41**, 1295-1308 (2020).
31. Chu, C. et al. Using DNA barcodes to aid the identification of larval fishes in tropical estuarine waters (Malacca Straits, Malaysia). *Zoological Studies* **58** (2019).
32. Hubert, N., Delrieu-Trottin, E., Irisson, J.-O., Meyer, C. & Planes, S. Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. *Molecular Phylogenetics and Evolution* **55**, 1195-1203 (2010).
33. Ko, H.-L. et al. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS one* **8**, e53451 (2013).
34. Chin, T. C., Adibah, A., Hariz, Z. D. & Azizah, M. S. Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. *Food Control* **64**, 247-256 (2016).
35. Hubert, N. et al. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS one* **3**, e2490 (2008).
36. Landi, M. et al. DNA barcoding for species assignment: the case of Mediterranean marine fishes. *PLoS one* **9**, e106135 (2014).
37. Russell, D., Thuesen, P. & Thomson, F. A review of the biology, ecology, distribution and control of Mozambique tilapia, *Oreochromis mossambicus* (Peters 1852)(Pisces: Cichlidae) with particular emphasis on invasive Australian populations. *Reviews in Fish Biology and Fisheries* **22**, 533-554 (2012).
38. Hebert, P. D., Cywinska, A. & Ball, S. L. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* **270**, 313-321 (2003).
39. Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* **21**, 1864-1877 (2012).
40. Meier, R., Zhang, G. & Ali, F. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. *Systematic Biology* **57**, 809-813 (2008).
41. Ortiz, D. & Francke, O. F. Two DNA barcodes and morphology for multi-method species delimitation in Bonnetina tarantulas (Araneae: Theraphosidae). *Molecular Phylogenetics and Evolution* **101**, 176-193 (2016).
42. Hajibabaei, M., Singer, G. A., Hebert, P. D. & Hickey, D. A. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends in Genetics* **23**, 167-172 (2007).
43. Mecklenburg, C. W., Møller, P. R. & Steinke, D. Biodiversity of arctic marine fishes: taxonomy and zoogeography. *Marine Biodiversity* **41**, 109-140 (2011).
44. Puckridge, M., Andreakis, N., Appleyard, S. A. & Ward, R. D. Cryptic diversity in flathead fishes (Scorpaeniformes: Platycephalidae) across the Indo-West Pacific uncovered by DNA barcoding. *Molecular Ecology Resources* **13**, 32-42 (2013).

45. Thirumaraiselvi, R. & Thangaraj, M. Genetic Diversity Analysis of Indian Salmon, *Eleutheronema tetradactylum* from South Asian Countries Based on Mitochondrial COI Gene Sequences. *Notulae Scientia Biologicae* **7**, 417-422 (2015).
46. Delrieu-Trottin, E. et al. Biodiversity inventory of the grey mullets (Actinopterygii: Mugilidae) of the Indo-Australian Archipelago through the iterative use of DNA-based species delimitation and specimen assignment methods. *Evolutionary Applications* **13**, 1451-1467 (2020).
47. Durand, J.-D., Hubert, N., Shen, K.-N. & Borsa, P. DNA barcoding grey mullets. *Reviews in Fish Biology and Fisheries* **27**, 233-243 (2017).
48. Alavi-Yeganeh, M. S., Khajavi, M. & Kimura, S. A new ponyfish, *Deveximentum mekranensis* (Teleostei: Leiognathidae), from the Gulf of Oman. *Ichthyological Research*, 1-8, <https://doi.org/10.1007/s10228-020-00794-y> (2021).
49. 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. (FAO Library, 2001).
50. Chen, W., Ma, X., Shen, Y., Mao, Y. & He, S. The fish diversity in the upper reaches of the Salween River, Nujiang River, revealed by DNA barcoding. *Scientific Reports* **5**, 1-12 (2015).
51. Guimarães-Costa, A. J. et al. Fish diversity of the largest deltaic formation in the Americas—a description of the fish fauna of the Parnaíba Delta using DNA Barcoding. *Scientific Reports* **9**, 1-8 (2019).
52. Hupało, K. et al. An urban Blitz with a twist: rapid biodiversity assessment using aquatic environmental DNA. *Environmental DNA* **3**, 200-213. (2020).
53. Zainal Abidin, D. H. & Noor Adelyna, M. A. in *Universities as Living Labs for Sustainable Development* 211-225 (Springer, 2020).
54. Ratnasingham, S. & Hebert, P. D. BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes* **7**, 355-364 (2007).
55. Benson, D. A. et al. GenBank. *Nucleic Acids Research* **46**, D41-D47 (2018).
56. Mansor, M. I. et al. Field guide to important commercial marine fishes of the South China Sea. (SEAFDEC/MFRDMD, 1998).
57. Nuruddin, A. A. & Isa, S. M. Trawl fisheries in Malaysia—issues, challenges and mitigating measures. (Fisheries Research Institute, Department of Fisheries Malaysia, 2013).
58. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547-1549 (2018).
59. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111-120 (1980).
60. Bouckaert, R. et al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* **10**, e1003537 (2014).

61. Edler, D., Klein, J., Antonelli, A. & Silvestro, D. raxmlGUI 2.0: a graphical interface and toolkit for phylogenetic analyses using RAxML. *Methods in Ecology and Evolution* (2020).
62. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott, B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **34**, 772-773 (2017).
63. Miller, M. A., Pfeiffer, W. & Schwartz, T. in Proceedings of the 2011 TeraGrid Conference: extreme digital discovery. 1-8 (2011).
64. Rambaut, A. FigTree v1.4.4. Available from: <http://tree.bio.ed.ac.uk/software/figtree/>
65. Ratnasingham, S. & Hebert, P. D. A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PloS one* **8**, e66213 (2013).
66. Pons, J. et al. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55**, 595-609 (2006).
67. Team, R. RStudio: integrated development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com> **42**, 14 (2015).

## Declarations

### Acknowledgements

We are very grateful to Norli Fauzani, Jamsari Amirul Firdaus, Siti Amalia Aisyah, Siti Zuliana, and Noor Zahidah for their help in sample collection. Special thanks to all the local fishermen of Merbok Estuary for their enthusiastic support in the sampling activities. This study was funded by the Ministry of Education (Department of Higher Education) through the Fundamental Research Grant Scheme (FRGS) [203/PCGSS/6711567], Universiti Sains Malaysia's Bridging Grant [304/PCGSS/6316237], and Ministry of Water, Land and Natural Resources Grant [UMT/IMB/2018/53307].

### Author contributions statement

Design and conception of experiments: DHZA NAMA SAMN SL. Data collection and specimens processing: DHZA SL. Barcoding data analyses: DHZA. Production and review of the manuscript: DHZA SAMN SL MAR NAJ NAMA.

### Competing Interest

The authors declare no competing interest.

## Tables

**Table 1.** List of fish species identified in Merbok Estuary and its adjacent marine waters through DNA barcoding.

ORDER, Family, <i>Species</i>	n	BOLD ID
<b>MYLIOBATIFORMES</b>		
<b>Dasyatidae</b>		
<i>Brevitrygon walga</i>	1	DBMR332-20
<i>Telatrygon zugei</i>	1	DBMR237-19
<b>Gymnuridae</b>		
<i>Gymnura poecilura</i>	1	DBMR036-19
<b>ORECTOLOBIFORMES</b>		
<b>Hemiscylliidae</b>		
<i>Chiloscyllium indicum</i>	1	DBMR335-20
<b>ANGUILLIFORMES</b>		
<b>Ophichthidae</b>		
<i>Pisodonophis cancrivorus</i>	1	DBMR001-19
<b>CLUPEIFORMES</b>		
<b>Chirocentridae</b>		
<i>Chirocentrus nudus</i>	1	DBMR331-20
<b>Clupeidae</b>		
<i>Anodontostoma chacunda</i>	1	DBMR341-20
<i>Escualosa thoracata</i>	3	DBMR342-20 - DBMR344-20
<i>Sardinella albella</i>	2	DBMR224-19, DBMR315-19
<b>Dussumieriidae</b>		
<i>Dussumieria albulina</i>	3	DBMR316-19 - DBMR318-19
<b>Engraulidae</b>		
<i>Setipinna taty</i>	1	DBMR009-19
<i>Stolephorus baweanensis</i>	1	DBMR011-19
<i>Stolephorus commersonii</i>	1	DBMR314-19
<i>Stolephorus dubiosus</i>	7	DBMR220-19 - DBMR223-19, DBMR311-19 - DBMR313-19
<i>Stolephorus indicus</i>	1	DBMR340-20
<i>Stolephorus tri</i>	1	DBMR010-19

<i>Thryssa hamiltonii</i>	4	DBMR012-19 - DBMR014-19, DBMR016-19
<i>Thryssa kammalensis</i>	4	DBMR018-19 - DBMR021-19
<i>Thryssa mystax</i>	2	DBMR015-19, DBMR017-19
<b>Pristigasteridae</b>		
<i>Ilisha melastoma</i>	6	DBMR022-19 - DBMR027-19
<i>Opisthopterus tardoore</i>	2	DBMR028-19 - DBMR029-19
<b>SILURIFORMES</b>		
<b>Ariidae</b>		
<i>Arius gagora</i>	3	DBMR273-19 - DBMR275-19
<i>Arius maculatus</i>	5	DBMR192-19 - DBMR196-19
<i>Hexanematichthys sagor</i>	3	DBMR198-19 - DBMR200-19
<i>Ketengus typus</i>	1	DBMR201-19
<i>Osteogeneiosus militaris</i>	4	DBMR202-19 - DBMR204-19, DBMR345-20
<i>Plicofollis argyropleuron</i>	3	DBMR205-19 - DBMR207-19
<i>Plicofollis layardi</i>	1	DBMR272-19
<i>Plicofollis polystaphylodon</i>	1	DBMR197-19
<b>Plotosidae</b>		
<i>Plotosus canius</i>	1	DBMR208-19
<b>AULOPIFORMES</b>		
<b>Synodontidae</b>		
<i>Saurida micropectoralis</i>	1	DBMR329-20
<b>BATRACHOIDIFORMES</b>		
<b>Batrachoididae</b>		
<i>Allenbatrachus grunniens</i>	1	DBMR336-20
<i>Batrachomoeus trispinosus</i>	1	DBMR002-19
<b>GOBIIFORMES</b>		
<b>Eleotridae</b>		
<i>Butis butis</i>	5	DBMR072-19, DBMR300-19 - DBMR301-19, DBMR322-19, DBMR323-19
<i>Butis humeralis</i>	8	DBMR065-19 - DBMR071-19, DBMR306-19

<i>Butis koilomatodon</i>	1	DBMR302-19
<b>Gobiidae</b>		
<i>Acentrogobius caninus</i>	1	DBMR082-19
<i>Boleophthalmus boddarti</i>	1	DBMR083-19
<i>Brachygobius aggregatus</i>	3	DBMR285-19 - DBMR287-19
<i>Exyrias puntang</i>	3	DBMR084-19, DBMR304-19, DBMR339-20
<i>Favonigobius gymnauchen</i>	2	DBMR085-19 - DBMR086-19
<i>Glossogobius aureus</i>	6	DBMR087-19 - DBMR092-19
<i>Hemigobius hoevenii</i>	1	DBMR296-19
<i>Psammogobius biocellatus</i>	1	DBMR305-19
<i>Pseudapocryptes elongatus</i>	2	DBMR093-19 - DBMR094-19
<i>Pseudogobius fulvicaudus</i>	3	DBMR297-19 - DBMR299-19
<i>Pseudogobius olorum</i>	2	DBMR294-19 - DBMR295-19
<i>Stigmatogobius sadanundio</i>	5	DBMR095-19 - DBMR096-19, DBMR291-19 -DBMR293-19
<i>Trypauchen pelaeos</i>	1	DBMR321-19
<i>Trypauchen vagina</i>	1	DBMR097-19
<i>Cryptocentrus</i> sp.	1	DBMR338-20
<b>ATHERINIFORMES</b>		
<b>Phallostethidae</b>		
<i>Neostethus lankesteri</i>	3	DBMR282-19 - DBMR284-19
<b>BELONIFORMES</b>		
<b>Adrianichthyidae</b>		
<i>Oryzias javanicus</i>	3	DBMR279-19 - DBMR281-19
<b>Belonidae</b>		
<i>Strongylura strongylura</i>	3	DBMR003-19 - DBMR005-19
<b>Hemiramphidae</b>		
<i>Hyporhamphus dussumieri</i>	1	DBMR219-19
<i>Hyporhamphus quoyi</i>	4	DBMR006-19 - DBMR008-19, DBMR324-19
<b>Zenarchopteridae</b>		

<i>Dermogenys sumatrana</i>	3	DBMR288-19 - DBMR290-19
<b>CARANGIFORMES</b>		
<b>Carangidae</b>		
<i>Alepes melanoptera</i>	2	DBMR043-19 - DBMR044-19
<i>Atule mate</i>	1	DBMR045-19
<i>Carangoides coeruleopinnatus</i>	3	DBMR046-19 - DBMR047-19, DBMR247-19
<i>Caranx ignobilis</i>	4	DBMR048-19 - DBMR051-19
<i>Caranx sexfasciatus</i>	1	DBMR346-20
<i>Megalaspis cordyla</i>	5	DBMR052-19 - DBMR056-19
<i>Scomberoides commersonianus</i>	1	DBMR060-19
<i>Scomberoides tala</i>	1	DBMR059-19
<i>Scomberoides tol</i>	2	DBMR057-19 - DBMR058-19
<i>Selaroides leptolepis</i>	2	DBMR328-20, DBMR330-20
<i>Trachinotus blochii</i>	2	DBMR061-19 - DBMR062-19
<i>Ulua mentalis</i>	1	DBMR246-19
<b>Cynoglossidae</b>		
<i>Cynoglossus arel</i>	1	DBMR268-19
<i>Cynoglossus puncticeps</i>	3	DBMR187-19, DBMR325-19, DBMR334-20
<i>Cynoglossus cynoglossus</i>	4	DBMR269-19 - DBMR271-19, DBMR319-19
<i>Cynoglossus lingua</i>	3	DBMR188-19 - DBMR190-19
<i>Cynoglossus oligolepis</i>	1	DBMR320-19
<b>Paralichthyidae</b>		
<i>Pseudorhombus elevatus</i>	1	DBMR327-20
<b>Latidae</b>		
<i>Lates calcarifer</i>	5	DBMR102-19 - DBMR106-19
<b>Polynemidae</b>		
<i>Eleutheronema tetradactylum</i>	7	DBMR137-19 - DBMR143-19
<i>Leptomelanosoma indicum</i>	1	DBMR144-19

<b>MUGILIFORMES</b>		
<b>Mugilidae</b>		
<i>Crenimugil buchmanani</i>	2	DBMR226-19, DBMR236-19
<i>Crenimugil crenilabis</i>	3	DBMR225-19 - DBMR228-19
<i>Osteomugil perusii</i>	2	DBMR229-19 - DBMR230-19
<i>Planiliza subviridis</i>	12	DBMR030-19 - DBMR035-19, DBMR231-19 - DBMR235-19, DBMR333-20
<b>PERCIFORMES</b>		
<b>Gerreidae</b>		
<i>Gerres filamentosus</i>	4	DBMR073-19 - DBMR074-19, DBMR081-19, DBMR248-19
<i>Gerres limbatus</i>	4	DBMR077-19 - DBMR080-19
<i>Gerres oyena</i>	2	DBMR075-19 - DBMR076-19
<b>Ambassidae</b>		
<i>Ambassis vachellii</i>	8	DBMR238-19 - DBMR245-19
<i>Ambassis interrupta</i>	3	DBMR040-19 - DBMR042-19
<i>Ambassis macracanthus</i>	5	DBMR037-19 - DBMR039-19, DBMR347-20 - DBMR348-20
<b>Haemulidae</b>		
<i>Pomadasys kaakan</i>	4	DBMR098-19 - DBMR101-19
<b>Lethrinidae</b>		
<i>Lethrinus lentjan</i>	3	DBMR128-19 - DBMR130-19
<b>Lutjanidae</b>		
<i>Lutjanus argentimaculatus</i>	1	DBMR136-19
<i>Lutjanus johnii</i>	2	DBMR131-19 - DBMR132-19
<i>Lutjanus russellii</i>	7	DBMR134-19 - DBMR135-19, DBMR250-19 - DBMR253-19, DBMR133-19
<b>Sciaenidae</b>		
<i>Dendrophysa russelii</i>	1	DBMR150-19
<i>Johnius</i> sp.	5	DBMR254-19 - DBMR257-19, DBMR259-19
<i>Nibea soldado</i>	3	DBMR151-19 - DBMR153-19
<i>Otolithes ruber</i>	5	DBMR154-19 - DBMR158-19

<i>Panna microdon</i>	1	DBMR260-19
<i>Pennahia anea</i>	1	DBMR266-19
<i>Pennahia ovata</i>	5	DBMR261-19 - DBMR265-19
<b>Serranidae</b>		
<i>Cephalopholis formosa</i>	1	DBMR278-19
<i>Epinephelus bleekeri</i>	1	DBMR277-19
<i>Epinephelus coioides</i>	5	DBMR159-19 - DBMR163-19
<i>Epinephelus heniochus</i>	1	DBMR276-19
<i>Epinephelus sexfasciatus</i>	1	DBMR352-20
<b>Sillaginidae</b>		
<i>Sillago sihama</i>	4	DBMR170-19 - DBMR173-19
<b>Sphyraenidae</b>		
<i>Sphyraena barracuda</i>	1	DBMR309-19
<i>Sphyraena jello</i>	2	DBMR174-19, DBMR310-19
<i>Sphyraena qenie</i>	1	DBMR267-19
<b>Platycephalidae</b>		
<i>Grammoplites scaber</i>	1	DBMR326-20
<i>Platycephalus indicus</i>	1	DBMR191-19
<b>Tetrarogidae</b>		
<i>Trichosomus trachinoides</i>	2	DBMR308-19, DBMR337-20
<b>Cichlidae</b>		
<i>Oreochromis mossambicus</i>	1	DBMR063-19
<b>ACANTHURIFORMES</b>		
<b>Drepaneidae</b>		
<i>Drepane punctata</i>	4	DBMR064-19, DBMR349-20 - DBMR351-20
<b>Leiognathidae</b>		
<i>Deveximentum ruconius</i>	2	DBMR124-19 - DBMR125-19
<i>Deveximentum indicium</i>	2	DBMR123-19, DBMR249-19
<i>Deveximentum hanedai</i>	1	DBMR126-19

<i>Eubleekeria jonesi</i>	1	DBMR107-19
<i>Leiognathus brevisrostris</i>	3	DBMR108-19 - DBMR110-19
<i>Leiognathus equula</i>	3	DBMR111-19 - DBMR113-19
<i>Nuchequula gerreoides</i>	9	DBMR114-19 - DBMR122-19
<b>Scatophagidae</b>		
<i>Scatophagus argus</i>	5	DBMR145-19 - DBMR149-19
<b>Siganidae</b>		
<i>Siganus fuscescens</i>	1	DBMR164-19
<i>Siganus javus</i>	5	DBMR165-19 - DBMR169-19
<b>SCOMBRIFORMES</b>		
<b>Stromatidae</b>		
<i>Pampus argenteus</i>	6	DBMR175-19 - DBMR180-19
<b>Trichiuridae</b>		
<i>Lepturacanthus savala</i>	3	DBMR184-19 - DBMR186-19
<b>CENTRARCHIFORMES</b>		
<b>Terapontidae</b>		
<i>Terapon jarbua</i>	2	DBMR182-19, DBMR307-19
<i>Terapon theraps</i>	2	DBMR181-19, DBMR183-19
<b>TETRAODONTIFORMES</b>		
<b>Tetraodontidae</b>		
<i>Arothron reticularis</i>	1	DBMR209-19
<i>Dichotomyctere fluviatilis</i>	1	DBMR210-19
<i>Dichotomyctere nigroviridis</i>	1	DBMR303-19
<i>Lagocephalus aff. lunaris</i>	1	DBMR213-19
<i>Lagocephalus lunaris</i>	5	DBMR211-19 - DBMR212-19, DBMR214-19 - DBMR216-19
<i>Takifugu oblongus</i>	1	DBMR217-19
<b>Triacanthidae</b>		
<i>Triacanthus nieuhofii</i>	1	DBMR218-19

**Table 2.** K2P divergence values from 350 analysed specimens with increasing taxonomic levels.

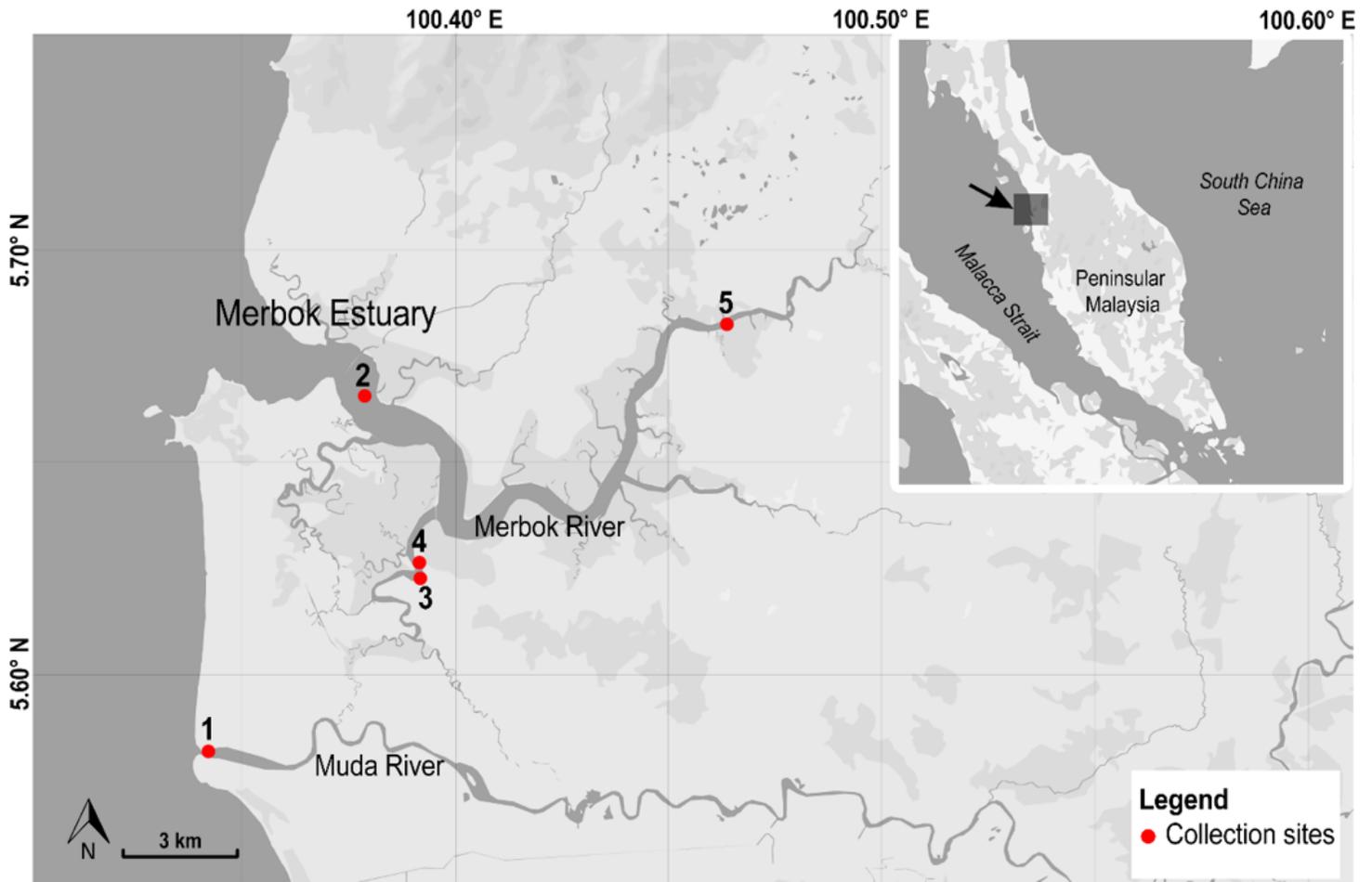
Category	n	Taxa	Comparisons	Minimum (%)	Mean (%)	Maximum (%)	SE (%)
Within species	285	75	532	0	0.85	16.66	0.01
Within genus	169	24	427	1.11	16.7	23.59	0.01
Within family	207	16	1197	0.16	18.17	26.02	0

SE: standard error

**Table 3.** List of morphological species with more than one MOTU or shared an MOTU. The summary statistics include the BIN of each species, their maximum intraspecific distance and distance to the nearest neighbour.

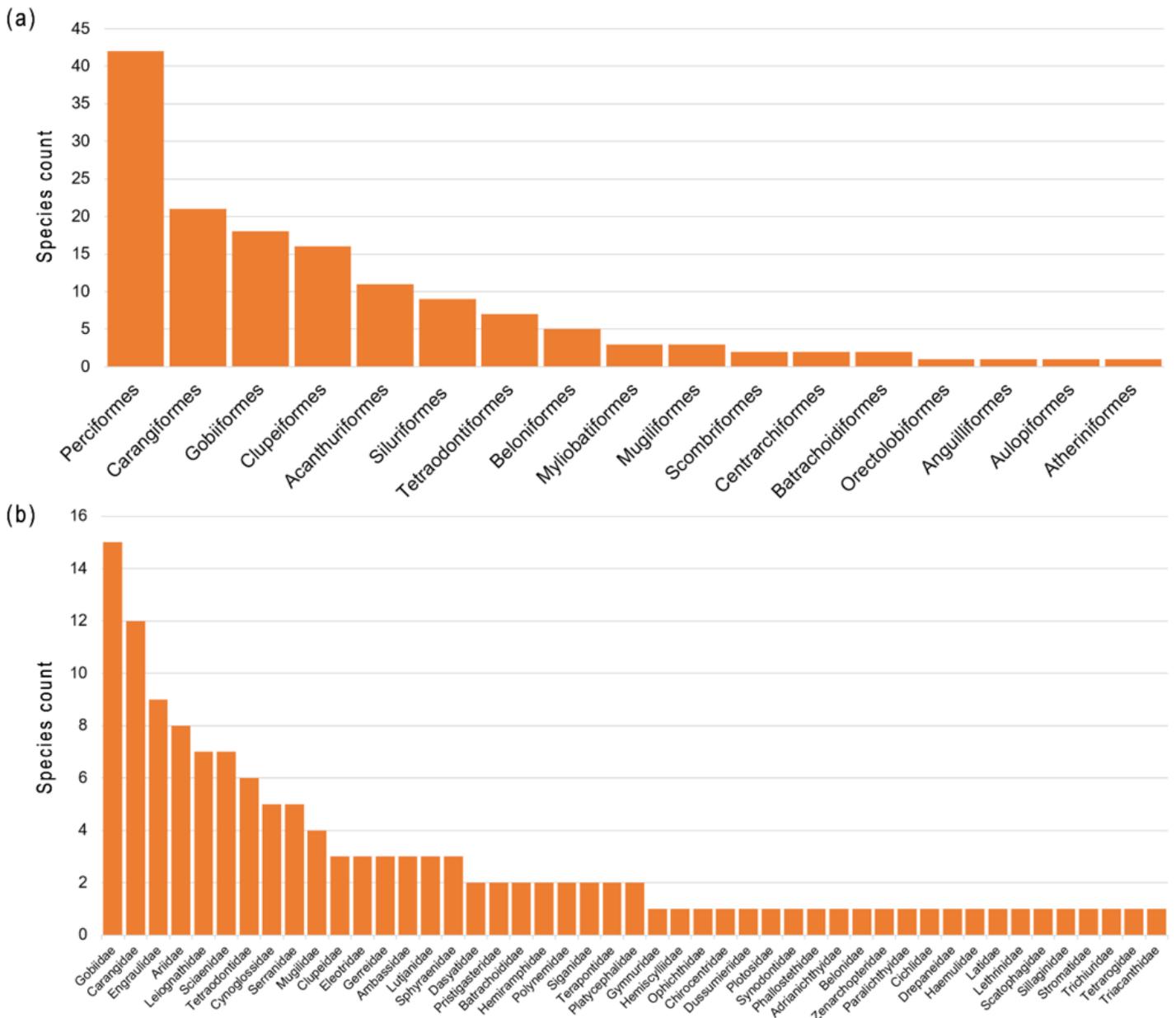
Species/ MOTUs (BIN)	Max. intraspecific distance (%)	Nearest neighbour distance (%)
Species with more than 1 MOTU/ BIN		
<i>Deveximentum indicium</i>	9.05	9.04
BOLD:ADZ6313	0	9.06
BOLD:AAF1238	0	8.41
<i>Eleutheronema tetradactylum</i>	16.66	19.40
BOLD:AAB8457	0	14.44
BOLD:AAB8458	0	14.44
<i>Gerres oyena</i>	4.29	18.49
BOLD:AAC1288	0	4.15
BOLD:AAC1290	0	4.15
<i>Lutjanus russellii</i>	4.12	14.24
BOLD:AAB2905	0.31	3.69
BOLD:AAB2904	0	3.69
<i>Osteomugil perusii</i>	14.24	15.90
BOLD:AAG3686	0.00	12.75
BOLD:AAW7354	0.00	12.75
<i>Planiliza subviridis</i>	13.44	15.30
BOLD:ABU7210	0.92	11.98
BOLD:ACC0823	0	11.98
Species with shared MOTU/ BIN		
<i>Alepes melanoptera</i>	0	0.16
<i>Caranx sexfasciatus</i>	0	0.16
BOLD:AAB5775	0.16	10.45
<i>Dichotomyctere cf. fluviatilis</i>	0	1.11
<i>Dichotomyctere nigroviridis</i>	0	1.11
BOLD:AAF2344	1.02	11.67

## Figures



**Figure 1**

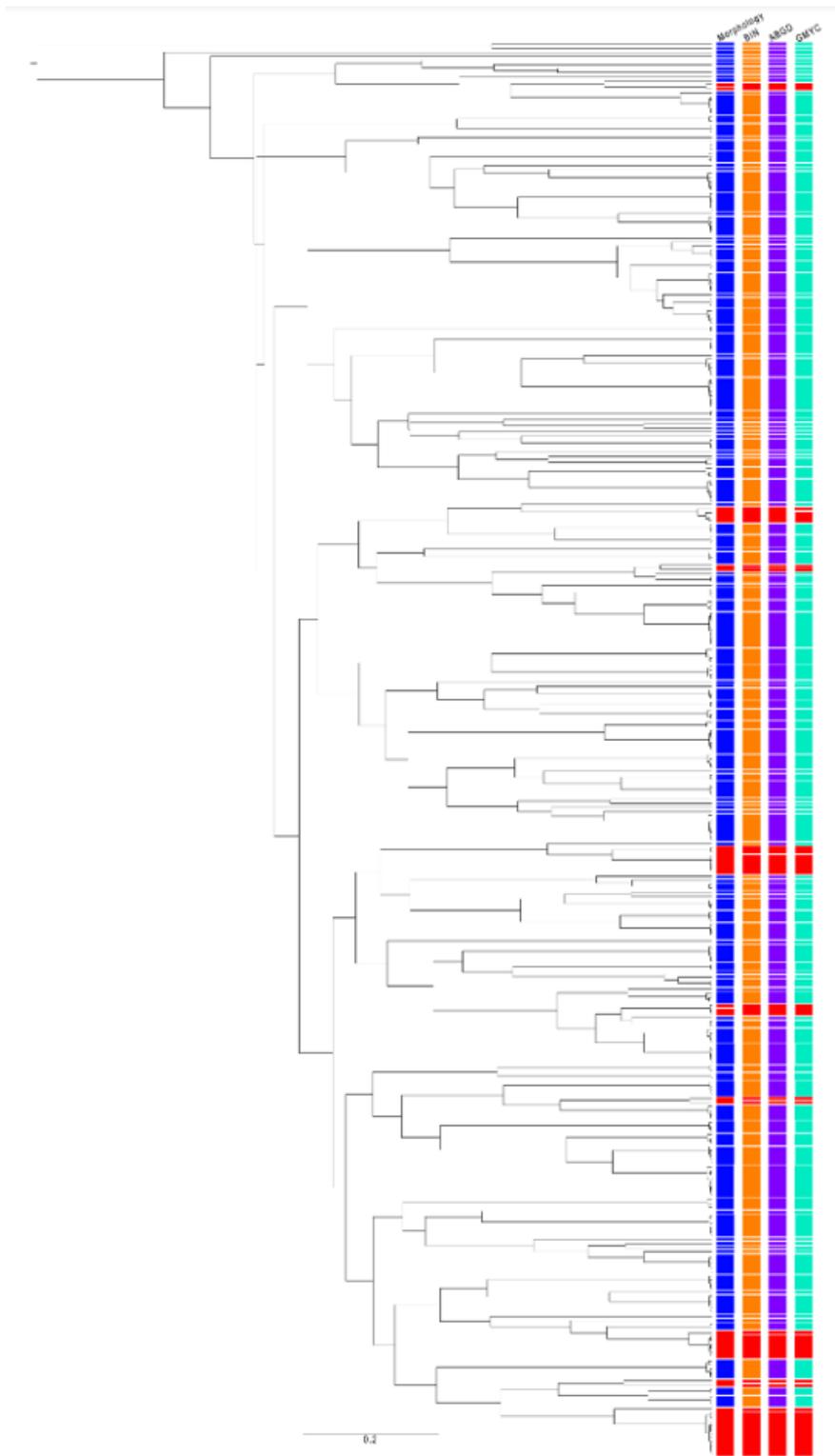
Sampling localities across the study area, which covers the Merbok Estuary (Merbok River) and Muda River. Sampling sites; 1: Kuala Muda Whispering Market, 2: Pompang Sungai Merbok, 3 and 4: Pompang Batu Lintang, 5: Semeling Bridge. Inset map shows the location of the study area within Peninsular Malaysia. Maps are generated using QGIS v.3.4.11 and edited in Adobe Photoshop CC 2019. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



**Figure 2**

Species count rankings according to (a) orders and (b) families recorded in this study.





**Figure 4**

Bayesian inference gene tree based on the 350 DNA barcodes with delineated MOTUs. Colour bars indicate delineated MOTUs by different methods – morphological characters (blue), BIN (orange), ABGD (purple), and GMYC (green). Red bars indicate MOTU discrepancies among the different schemes.

## Supplementary Files

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

- [DBMRManuscriptSupplementaryInformation.docx](#)