Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# DNA barcoding reveals species diversity, cryptic lineages, and new records of marine fishes in East Coast EEZ of Peninsular Malaysia

**Noorul Azliana Jamaludin**
Fisheries Research Institute Kampung Acheh
**Jamsari Amirul Firdaus Jamaluddin**
jamsariafj@gmail.com
Fisheries Research Institute Kampung Acheh
**Mohd Hans Afifi Hassan**
Fisheries Research Institute Kampung Acheh
**Hashim Shuhaimi**
Fisheries Research Institute Kampung Acheh
**Nurul Nadwa Abdul Fatah**
Fisheries Research Institute Kampung Acheh
**Nabilah Mustafha**
Fisheries Research Institute Kampung Acheh
**Nor Bariah Omaruddin**
Fisheries Research Institute Kampung Acheh
**Ryon Siow**
Fisheries Research Institute Kampung Acheh
**Haryati Abd Wahab**
Department of Fisheries Malaysia

**Research Article**

**Keywords:** Marine fishes, East Coast of Peninsular Malaysia, DNA barcode, COI, species delimitation

**Posted Date:** July 20th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10324186/v1

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

**Additional Declarations:** No competing interests reported.

# Abstract

The South China Sea is one of the world's most biodiverse shallow-water marine ecosystems, yet its true species richness remains underestimated owing to the continued discovery of new species, cryptic diversity, and unresolved taxonomic groups. Although molecular approaches have increasingly been applied to marine biodiversity studies, research in the South China Sea remains geographically biased toward its northern regions, with comparatively few investigations conducted in Malaysian waters. This study employed cytochrome c oxidase subunit I (COI) DNA barcoding to assess marine fish diversity within the Exclusive Economic Zone (EEZ) of the east coast of Peninsular Malaysia. A total of 508 specimens representing 225 species, 141 genera, 69 families, 25 orders, and two classes (Elasmobranchii and Actinopterygii) were analysed. Three species, *Engyprosopon xystrias, Ostorhinchus septemstriatus*, and *Synodus nigrotaeniatus*, are reported as new records for Malaysia and/or the South China Sea. Consensus species delimitation analyses identified 230 Molecular Operational Taxonomic Units (MOTUs), revealing cryptic diversity within *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi*, and *Netuma thalassina*. These findings demonstrate COI DNA barcoding complements traditional taxonomy, establishing a comprehensive COI-based reference library supporting biodiversity assessment, species identification, conservation planning, and sustainable management of marine fish resources in the region for research.


# Introduction

The South China Sea is approximately 3.5 million km² marginal sea located between the Pacific and Indian Oceans, spanning the subtropical and tropical zones of China, Malaysia, Brunei, Indonesia, the Philippines, Vietnam, and Thailand (Wang et al. 2014; Tang et al. 2023). Its depth ranges from shallow Sunda Shelf (less than 200 m) to the deep waters of the Manila Trench (more than 5000 m), and it is studded with numerous islets, atolls, and reefs, many of which are just awash at low tide, as well as extensive seagrass, mangrove coverage, and estuarine environments (Morton and Blackmore 2001).

Due to its location at the boundary between two oceans, dynamic Quaternary morphological changes, and tremendous diversity of natural ecosystems, the South China Sea supports highly productive biological resources, with more than 8,600 estimated species of plants and animals occurring in the region (Ng and Tan 2000), making it one of the world’s most diverse shallow-water marine areas (Morton and Blackmore 2001). This biodiversity encompasses not only species richness and intraspecific genetic diversity, but also habitat and functional diversity across ecosystems. Collectively, these biological resources provide a vital foundation for the livelihoods and well-being of local communities, supporting approximately 270 million people living within the coastal subregions of the South China Sea (Ke et al. 2022). However, there is growing concern over the overexploitation of these biological resources, particularly given the number of nations that depend on and utilize them, making the conservation and management of this species rich diversity essential for sustaining these resources. Information on a species, covering its occurrence, breeding, population structure, and demographic history, as well as its distribution patterns, habitat preferences, and responses to environmental change, is crucial for

understanding its ecology, evolution, and conservation needs (Gonzalez et al. 2024; Amos 2025; Ng et al. 2025). Therefore, an integrated and multidisciplinary approach incorporating modern techniques is needed to gather critical information in a race against time before it becomes too late to prevent stock depletion and species extinction (Ward et al. 2005; Xiong et al. 2022; Seijas-Díaz and Presa 2025).

Knowledge of marine fish diversity in the South China Sea has progressively expanded through both traditional taxonomic surveys (Vidthayanon 1999a, 1999b; Muda et al. 2001; Adrim et al. 2004; Matsunuma et al. 2011; Afiq-Firdaus et al. 2021; Li et al. 2022; Ng et al. 2024) and more recent molecular studies (Zhang 2011; Wang et al. 2012; Zhang and Hanner 2012; Thu et al. 2019; Xu et al. 2019, 2021; Azahar et al. 2022; Li et al. 2023; Zhang et al. 2023; Abidin et al. 2024; Jiang et al. 2024). Current assessments estimate that approximately 3,365 to 3,857 marine fish species, representing nearly 20% of the world's approximately 19,800 marine fish species, have been recorded from the region, highlighting the South China Sea as a major global biodiversity hotspot (Randall and Lim 2000; Froese and Pauly 2026). With approximately 21% of global marine fish species still remaining undescribed (Mora et al. 2008), together with the ongoing discovery of new species (Chang et al. 2025; Huang et al. 2025; Mat Jaafar et al. 2025), new distributional records (Yusri et al. 2021), and cryptic taxa (Abidin et al. 2024), the documented marine fish diversity of the South China Sea is expected to increase further. Nevertheless, comprehensive understanding of species composition and distribution in the region remains incomplete due to the limitations of morphology-based identification, particularly for cryptic and closely related taxa. Furthermore, the presence of unresolved taxa and gaps in genetic reference databases highlights the need for continued integrative research combining traditional taxonomy, molecular approaches, and ecological data to achieve a more comprehensive understanding of regional ichthyofaunal diversity.

Despite increasing research efforts, studies on marine fish diversity in the South China Sea have been dominated mainly by traditional taxonomic surveys, while molecular-based studies remain comparatively limited despite their growing application in recent years. In addition, most molecular studies have focused on the northern South China Sea, with fewer studies conducted in the southern region, particularly along the Malaysian coast. Knowledge of marine fish diversity in Malaysia has been documented through numerous pictorial guides, reference books, and research publications. Mohsin and Ambak (1996) reported 710 marine fish species from Malaysia and adjacent seas, while Ambak et al. (2012) estimated approximately 2,243 marine fish species in Malaysian waters. Chong et al. (2010) also documented 428 coral reef-associated fish species representing 109 families and 504 genera. In the broader South China Sea region, Vidthayanon (1999a, 1999b) recorded 300 fish species representing 89 families from the Gulf of Thailand and the east coast of Peninsular Malaysia, as well as 518 species representing 108 families from the waters of Sarawak, Sabah, and Brunei Darussalam. Matsunuma et al. (2011) later catalogued 441 marine and estuarine fish species from Terengganu, including four new records for the South China Sea. Similarly, Du et al. (2019) documented 314 species from 51 families in the Redang Archipelago, including five new records for Malaysian waters. Seah et al. (2020) reported 169 marine fish species from 62 families near Bidong Island, including nine new records for the South China Sea, while Seah et al. (2021) documented 256 species from 79 families in Kuantan, including 12 new regional records. Collectively, about 1,389 marine fish species associated with the South China Sea

have been recorded from Malaysian waters (Froese and Pauly 2026). More recently, molecular approaches have increasingly been applied in the Malaysian sector of the South China Sea. Abidin et al. (2024), using DNA taxonomy, identified 93 species representing 41 families along the east coast of Peninsular Malaysia, including two new records and evidence of cryptic species. Similarly, Azahar et al. (2022) DNA barcoded 83 species from 36 families in Pulau Tinggi, Johor, highlighting the value of molecular approaches for improving species identification and biodiversity assessment in the region.

Accurate fish species identification is essential for sustainable resource management because each species possesses unique biological and ecological characteristics, while understanding the factors that maintain high marine fish diversity is equally important (Li et al. 2023). Traditional taxonomy based on morphological traits remains widely used for species identification; however, it can be challenging for complex species groups due to phenotypic plasticity, cryptic species, hybridization, and convergent evolution, which may lead to misidentification (Human 2007; Vrijenhoek 2009; Cerutti-Pereyra et al. 2012). To overcome these limitations, DNA barcoding was introduced as a standardized molecular method for accurate species identification (Hebert et al. 2003). This technique uses a partial fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene (~ 650 bp) as a universal marker to distinguish species based on genetic variation (Ward et al. 2005). The COI gene is particularly suitable because of its high mutation rate, maternal inheritance, and lack of recombination, making it efficient and reliable for species-level identification (Xing et al. 2025). Since its introduction, DNA barcoding has become a rapid and cost-effective tool for resolving taxonomic uncertainties and detecting cryptic diversity across many taxa, including fishes, insects, crustaceans, birds, and mammals (Alam et al. 2024; Shneyer et al. 2025). Its effectiveness is further supported by the Barcode of Life Data Systems (BOLDSystems), which integrates DNA sequences with specimen and geographic information to support standardized species identification (Ratnasingham and Hebert 2007, 2013). DNA barcoding has contributed greatly to biodiversity assessment, environmental monitoring, food authentication, forensic studies, invasive species detection, and environmental DNA (eDNA) research (Staats et al. 2016; Ruppert et al. 2019; Filonzi et al. 2023; Jamaluddin et al. 2022; Suhailan et al. 2025). Although challenges such as incomplete reference databases and insufficient sampling still exist, DNA barcoding remains a powerful complementary tool to traditional taxonomy and has significantly improved the accuracy and efficiency of biodiversity research and conservation efforts.

In this study, specimens of marine fishes were collected from the waters within the Exclusive Economic Zone (EEZ) of the east coast of Peninsular Malaysia to investigate species diversity and distribution using DNA barcoding. This approach enables accurate species identification and provides valuable insights into regional fish diversity. The resulting DNA barcode reference dataset serves as an important scientific foundation for biodiversity assessment, species identification, and resource management in the South China Sea and Malaysia.

## Methodology

# Sampling and taxonomic identification

Sampling was carried out across 63 sites along the east coast of Peninsular Malaysia within the Exclusive Economic Zone (EEZ) of the South China Sea between 22 June and 15 August 2025 (Fig. 1; Supplementary Material 1). Fish specimens were collected using a single fishing vessel equipped with bottom trawl gear, operating at depths ranging from 7.6 m to 60.2 m depending on site conditions. Additional specimens were collected on 23 August 2025 from three fish landing sites in the Straits of Johor: Kampung Pasir Gudang Baru, Kong Kong, and Pengerang. Following collection, specimens were placed in labelled plastic bags and transported on ice to the sorting station. Specimens were identified by Fisheries Department personnel to the lowest possible taxonomic level using standard identification guides and taxonomic references (Matsunuma et al. 2011; Ambak et al. 2012; Ahmad Ali et al. 2018). Taxonomic names and classifications were subsequently verified and updated using the California Academy of Sciences' Eschmeyer's Catalog of Fishes database (Fricke et al. 2025) and WoRMS database (WoRMS Editorial Board 2026). Following identification, specimens were labelled and photographed, and tissue samples (fin clips) were preserved in 95% ethanol and stored at − 20°C. The remaining specimens were fixed in 10% formalin and later transferred to 70% ethanol for long-term preservation. All voucher specimens were deposited in the ichthyological collection of the Fisheries Research Institute, Kampung Acheh, Perak.

## DNA extraction, amplification and sequencing

A total of 508 specimens were selected from the 1,131 documented individuals to represent the full range of morphological taxonomic diversity for subsequent molecular analyses (Supplementary Material 1). Genomic DNA was extracted using the PrimeWay Genomic DNA Extraction Kit (1st BASE, Singapore) in accordance with the manufacturer's instructions. A partial fragment of the cytochrome c oxidase subunit I (COI) gene was amplified in a 25 µL reaction mixture containing 12.5 µL of 2× MyTaq™ Red Reaction Buffer (Bioline GmbH, Germany), 0.5 µL of each primer (5 µM), and 10.5 µL of sterile distilled water. Three primer pairs were used: FishF1/FishR1 and FishF2/FishR2 (Ward et al. 2005), and Johni 6F/Johni 916R for *Johnius* spp. (Lo et al. 2017). PCR amplification was performed with an initial denaturation at 95°C for 4 min, followed by 30 cycles of 94°C for 30 s, 47.9°C for 50 s, and 72°C for 1 min, and a final extension at 72°C for 7 min. The resulting PCR products were sequenced using Sanger sequencing at Apical Scientific Sdn. Bhd. (Selangor, Malaysia) with the BigDye Terminator v3.1 Cycle Sequencing Kit on an ABI 3730xl DNA Analyzer (Applied Biosystems, Thermo Fisher Scientific, USA).

## Data analysis and estimation of MOTUs

Chromatograms were edited, and consensus COI sequences for each specimen were generated by assembling forward and reverse reads using MEGA v10.1.7 (Kumar et al. 2018). Sequence alignment was performed using ClustalW implemented in MEGA v10.1.7. All data and sequences from this study have been deposited in the Barcode of Life Data System (BOLD) database (https://v4.boldsystems.org) under the project "Assessment of Coastal Marine Demersal Resources on the East Coast of Peninsular Malaysia" (project code: FRIKA).

Maximum likelihood (ML) phylogenetic reconstruction was used to assess species monophyly, implemented in RAxML v8.0 (Stamatakis 2014) via the graphical interface raxmlGUI v2.0 (Silvestro and Michalak 2012). Phylogenetic trees were reconstructed under the general time-reversible model with gamma-distributed rate heterogeneity (GTR + Γ), with node support assessed using 1,000 rapid bootstrap replicates, and visualized in FigTree v1.4.4 ((http://tree.bio.ed.ac.uk/software/figtree/).

Barcode gap analysis was performed using the BOLD analytical toolkit to assess taxonomic discrimination by comparing intraspecific genetic distances with nearest-neighbour distances for each species. The analysis included only sequences ≥ 500 bp and was conducted using the pairwise distance model with sequence alignment performed using the Kalign algorithm (Lassmann and Sonnhammer 2005).

DNA-based species delimitation was performed on the dataset using three independent approaches, including Assemble Species by Automatic Partitioning (ASAP) (Puillandre et al. 2021), Refined Single Linkage (RESL) as implemented in BOLD (Ratnasingham and Hebert 2013), and the multi-rate Poisson Tree Processes (mPTP) model (Kapli et al. 2017), to delineate molecular operational taxonomic units (MOTUs). The ASAP analysis was performed on the ASAP tool within the SpartExplorer platform (https://spartexplorer.mnhn.fr/) using the simple distance (p-distance) metric. The partition with the lowest score was selected as the final species delimitation result. RESL analysis was conducted on the BOLD Systems platform using the “cluster sequences” function with default parameters, which assigns sequences to Barcode Index Numbers (BINs). The resulting BIN assignments, evaluated using the BIN Discordance tool in BOLD Systems, were used to delineate molecular operational taxonomic units (MOTUs) and assess taxonomic congruence within the dataset. The mPTP delimitation was conducted using the mPTP web interface (https://mptp.h-its.org/#/tree) using the MrBayes and RAxML trees as input. All other parameters were set to default values. Subsequently, morphological evidence and phylogenetic relationships were integrated to evaluate these MOTUs and determine the final species delimitations.

## Results

# Sequence composition and Species Diversity

In this study, 508 fish sequences were generated from the east coast of Peninsular Malaysia’s Exclusive Economic Zone (EEZ). Sequence lengths ranged from 571 to 655 base pairs, and all sequences were free of indels and premature stop codons. The nucleotide composition averaged 29.6% (T), 28.3% (C), 23.8% (A), and 18.3% (G). Sampling effort varied among taxa, with specimen representation ranging from 1 to 14 individuals per species.

Based on morphological and molecular diagnostic characteristics, a total of 225 species, 141 genera, 69 families, and 25 orders across two classes of Elasmobranchii (rays and sharks) and Actinopterygii (bony fishes) were identified. Of the 225 species molecularly analysed, 68 were represented by a single

specimen, whereas 96 species were represented by three or more individuals, including 28 species that were well represented with at least five individuals each.

Out of the total, 12 species could only be identified to the genus level, while one specimen was identified to the subfamily level due to difficulties in morphological identification, high genetic divergence (> 2%) from the closest reference sequences, and inconsistent species matches within existing databases, which collectively complicated accurate taxonomic assignment. The assemblage was dominated by Carangidae (20 species), followed by Leiognathidae (15 species), Nemipteridae (14 species), Sciaenidae (11 species), and Apogonidae (10 species), while most other families were represented by fewer than five species. *Etroplus suratensis*, a species commonly associated with brackish-water habitats, was the only introduced species recorded in the present study. Cross-validation against available databases and published references confirmed that all recorded species are commonly distributed in Malaysian waters and the South China Sea, with the exception of *Engyprosopon xystrias, Ostorhinchus septemstriatus*, and *Synodus nigrotaeniatus*, which represent new records for Malaysia and/or the South China Sea.

Assessment of the identified to species level (212 species) against the International Union for Conservation of Nature (IUCN) Red List database (https://www.iucnredlist.org/) revealed that 167 species (78.8%) are classified as Least Concern. However, 14 species (6.6%) were of conservation concern. This group includes ten sharks and rays: *Rhynchobatus australiae* (Critically Endangered, CR), *Gymnura zonura* (Endangered, EN), *Maculabatis gerrardi* (EN), *Chiloscyllium plagiosum* (Near Threatened, NT), *Chiloscyllium punctatum* (NT), *Rhizoprionodon oligolinx*, (NT), *Telatrygon zugei* (Vulnerable, VU), *Brevitrygon heterura* (VU), *Gymnura poecilura* (VU), and *Hemigaleus microstoma* (VU), and four threatened bony fishes: *Choerodon schoenleinii* (NT), *Protonibea diacanthus* (NT), *Scomberomorus commerson* (NT), and *Rastrelliger brachysoma* (VU). In addition, 31 species were either Not Evaluated (19 species) or classified as Data Deficient (12 species). The full classification, IUCN conservation status, number of specimens barcoded, and associated BOLD and BIN IDs are presented in Supplementary Material 1.

# Genetic divergence, phylogeny, and species delimitation

In species represented by two or more DNA barcode sequences (n = 146), mean intraspecific K2P distances ranged from 0% to 4.49%, with most species exhibiting values below 1%, while maximum intraspecific distances ranged from 0% to 6.07% (Fig. 2A, B). Across all species (n = 212), minimum interspecific distances to the nearest neighbour (min. NN) ranged from 2.24% to 20.1% and were clearly separated from intraspecific variation, with interspecific divergence exceeding intraspecific divergence in most cases, supporting effective species discrimination using COI barcodes (Fig. 2A, B).

Analysis of the barcode gap revealed that all morphospecies were positioned above the 1:1 line, indicating that minimum nearest-neighbor distances exceeded maximum intraspecific divergences (Fig. 2C). Most morphospecies were clustered in the first quadrant (Q1), exhibiting low intraspecific divergence and clear separation from their nearest neighbors, consistent with a distinct barcode gap. In contrast, *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi*, and *Netuma thalassina* were positioned in the second quadrant (Q2), displaying elevated intraspecific

divergences ranging from 2.70% to 6.07% while maintaining substantially higher nearest-neighbor distances. This pattern suggests the presence of deeply divergent lineages and potential cryptic diversity within these taxa. This study also generated unique BINs for two taxa, *Ostorhinchus septemstriatus* (BOLD:AHK1679) and an unidentified species within the family Aploactinidae (BOLD:AHL3550).

The maximum likelihood (ML) tree generated from the complete barcode dataset supported the distinctiveness of the morphologically identified species clusters, including singleton taxa, with most barcode groupings receiving strong bootstrap support (≥ 87%). However, *Neotrygon varidens* and *N. malaccensis* did not form clearly separated monophyletic clusters in the ML analysis (Supplementary Material 2).

The Molecular Operational Taxonomic Units (MOTUs) identified using three species delimitation methods (ASAP, RESL, and mPTP) yielded varying numbers of MOTUs (Fig. 3). Numbers of delimited MOTUs were 222 for ASAP, 230 for RESL, and 232 for mPTP (Fig. 3, Table S4). The ASAP analysis grouped *Neotrygon varidens* and *N. malaccensis, Atropus atropos* and *A. mentalis, Pterocaesio chrysozona* and *P. pisang*, as well as *Rastrelliger kanagurta* and *R. brachysoma* into single MOTUs despite their morphological distinctiveness and interspecific genetic divergences exceeding the commonly applied 2% threshold. In contrast, *Netuma thalassina* was split into two distinct MOTUs. Additional MOTUs were also recovered in *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii*, and *Maculabatis gerrardi*, increasing the total number of MOTUs identified by RESL to 230. Meanwhile, mPTP recovered 232 MOTUs, including two additional MOTUs detected in *Elates ransonnettii* and *Synodus nigrotaeniatus*. Considering morphology, MOTU delimitation schemes, the typical 2% sequence divergence threshold for fishes, and the presence of a barcode gap, the consensus delimitation analysis recovered 230 MOTUs from 225 morphologically identified species. *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi*, and *Netuma thalassina* were each represented by two MOTUs, indicating incongruences between MOTU and species boundaries and suggesting the presence of cryptic diversity. The examination of the spatial distribution of haplotypes further supported this pattern, however, no consistent geographic barrier was observed across all five species, as some taxa exhibited geographically structured networks whereas others showed sympatric or mixed lineage distributions. In particular, both lineages of *Turrum coeruleopinnatum* and *Nemipterus peronii* occurred sympatrically across the sampled coastal regions (Fig. 3).

## Discussion

# Species Diversity

This study represents one of the most comprehensive molecular surveys conducted to establish a reliable DNA barcode reference library for marine fishes within the Exclusive Economic Zone (EEZ) of Peninsular Malaysia in the South China Sea, documenting 225 morphospecies across 126 genera, 69 families, 28 orders, and two classes. The previous DNA barcode checklist of fish species present from Malaysian waters of the South China Sea, established by Abidin et al. (2024) and Azahar et al. (2022)

recorded 93 species representing 41 families and 83 species representing 36 families, respectively. However, only 50 species (22.2%) overlapped with Abidin et al. (2024), while 16 species (7.1%) overlapped with Azahar et al. (2022). Collectively, these studies have contributed DNA barcodes for 335 marine fish species, with the highest diversity recorded in the families Carangidae (22 species), Lutjanidae (14 species), and Nemipteridae (14 species). At the genus level, *Nemipterus* was the most species-rich genus with 12 species, followed by *Lutjanus, Epinephelus*, and *Ostorhinchus* with eight species each. Despite these advances, many marine fish species in the region remain genetically undocumented, and substantial taxonomic gaps persist within existing DNA barcode reference libraries. Furthermore, the currently barcoded species represent only a fraction of the estimated 1,389 marine fish species reported from Malaysian waters of the South China Sea (Froese and Pauly 2026), highlighting the considerable potential for future DNA barcoding initiatives, biodiversity assessments, and species discoveries in the region. The large discrepancies observed between the present and previous studies may be attributed to differences in sampling strategies, including fish collection methods (e.g., bottom trawls versus spinning rods, small hand-nets, and gillnets), habitat coverage (e.g., inshore versus offshore and coral reef environments), and sampling effort, while taxonomic uncertainties may also contribute to species misidentifications and inconsistencies among studies (Chang et al. 2017).

# Genetic Divergence, species delimitation and Cryptic Lineages

DNA barcoding-based species identification is largely centered on patterns of sequence divergence (Xiao et al. 2022). Hebert et al. (2004) proposed that species can be distinguished when interspecific genetic divergence substantially exceeds intraspecific divergence, with a 10-fold difference suggested as a general benchmark for animal species identification. In this study, the mean interspecific divergence was 39 times higher than the mean intraspecific divergence, indicating strong genetic differentiation among the marine fish species examined.

Consequently, the presence of a barcode gap, characterized by a clear separation between intraspecific and interspecific genetic distances, is considered a reliable criterion for species discrimination (Meyer and Paulay 2005). In this study, all morphospecies exhibited distinct barcode gaps, with no overlap between intra- and interspecific genetic distances, supporting robust species delimitation consistent with morphological taxonomy. The clear separation observed may partly reflect the relatively limited geographic scope of sampling, the low representation of closely related taxa, and the presence of numerous species represented by a single specimen.

In fish identification, genetic distance thresholds are among the most widely used approaches for species delimitation. A 2% genetic distance threshold is commonly applied in fishes, whereby individuals with divergences below 2% are generally considered conspecific, whereas divergences exceeding 2% may indicate distinct species (Hubert and Hanner 2015). In addition, genetic divergences ranging from 1% to 2% may reflect population genetic structuring within a species across different geographic regions (Nunes et al. 2025). In the present study, the mean intraspecific genetic distance was 0.22%, which falls

within the range reported for marine fishes in the South China Sea (0.15–0.25%) (Xiao et al. 2022; Zhang et al. 2023) and is slightly lower than values reported from other geographic regions (reviewed in Xu et al. 2021). However, six species, namely *Siganus fuscescens, Scomberomorus commerson, Nibea soldado, Rastrelliger kanagurta, Synodus nigrotaeniatus, and Samaris cristatus*, exhibited relatively high intraspecific divergences ranging from 1% to 2%. These elevated divergence values may indicate the presence of population structuring within the study region. Such patterns could have been shaped by historical geomorphological events associated with the Sunda Shelf during the Quaternary period, when repeated sea-level fluctuations promoted population isolation and subsequent genetic differentiation (Voris 2000; Gaither et al. 2011; Lim et al. 2021; Goulding and Dayrat 2023).

Among the 225 OTUs investigated in this study, five species, namely *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi*, and *Netuma thalassina*, exhibited maximum intraspecific genetic distances exceeding 2%, whereas the remaining species showed divergences below this threshold, consistent with typical intraspecific variation. Similar patterns of elevated intraspecific divergence have been reported in other marine fish studies (Thu et al. 2023; Abidin et al. 2024) and were also observed in early DNA barcoding studies of marine fishes (Ward et al. 2005). Although these elevated divergences may indicate the presence of cryptic diversity, a universal genetic divergence threshold is difficult to apply across different taxonomic groups and barcode regions. As a result, fixed thresholds (e.g., 2%) may lead to under- or over-splitting due to lineage-specific evolutionary histories and variation in divergence rates among taxa. Therefore, comparative analyses should be used alongside threshold-based approaches to improve species delimitation accuracy and validate candidate cryptic lineages (Neigel et al. 2007).

To further evaluate species boundaries, species delimitation analyses were conducted using a distance-based method (ASAP), a network-based method (RESL), and a coalescent tree-based method (mPTP), which recovered between 222 and 232 MOTUs across the dataset depending on the method applied. The ASAP produced the most conservative estimate (222 MOTUs), whereas mPTP recovered the highest number of MOTUs (232), with RESL identifying 230 MOTUs. Despite minor differences among these methods, the majority of morphologically identified species were consistently recovered as single molecular entities, indicating strong congruence between morphological and molecular evidence. However, five species, namely *Pomadasys maculatus*, *Turrum coeruleopinnatum*, *Nemipterus peronii*, *Maculabatis gerrardi*, and *Netuma thalassina*, were consistently split into multiple MOTUs by most delimitation approaches, corresponding to their elevated intraspecific divergences (2.70–6.07%) and supporting the presence of deeply divergent lineages within currently recognized species. Conversely, a small number of taxa were merged into fewer MOTUs than morphologically recognized species, especially in ASAP. This result likely reflects the distance-partitioning nature of ASAP, which identifies species boundaries based on the overall distribution of pairwise genetic distances and barcode-gap structure across the dataset rather than applying a fixed divergence threshold (Puillandre et al. 2021; Hubert et al. 2024). For example, *Neotrygon varidens* and *N. malaccensis* exhibited the lowest interspecific p-distance (2.24%) and were grouped into a single MOTU by ASAP despite being recognized as distinct species within the recently revised *Neotrygon kuhlii* complex (Borsa et al. 2016; Loh et al.

2023). This result suggests potential under-splitting and highlights the limitations of single-locus delimitation methods, emphasizing the need for integrative taxonomic approaches to improve species boundary inference (Dellicour and Flot 2015).

The concordance between elevated genetic divergence, barcode-gap analyses, haplotype network structure, and multiple species delimitation approaches suggests the presence of candidate cryptic species within *Pomadasys maculatus*, *Turrum coeruleopinnatum*, *Nemipterus peronii*, *Maculabatis gerrardi*, and *Netuma thalassina*. However, because the delimitation analyses were based solely on mitochondrial COI data, the resulting MOTUs should be considered preliminary species hypotheses rather than definitive species boundaries. Integrative taxonomic studies incorporating morphological, ecological, biogeographic, and nuclear genetic data are therefore necessary to validate these putative lineages.

# New records and non-native species

The blackstripe lizardfish, *Synodus nigrotaeniatus* (Allen et al. 2017), was described from North Sulawesi in the Celebes Sea and was later reported from Jakarta Bay, Indonesia, in the Java Sea (Wibowo et al. 2024), West Papua in the Ceram Sea (based on BOLD records), and Palawan in the Sulu Sea (Bemis et al. 2023). In the present study, the species was recorded from the coast of Pahang, representing a significant expansion of its known distribution range and the first record of *S. nigrotaeniatus* from both Malaysian waters and the South China Sea. Interestingly, analysis of COI sequences generated in this study together with publicly available sequences from the BOLD Systems revealed the presence of two BINs in the region, showing a mean p-distance divergence of 1.4%. These BINs were associated with BOLD:ACU5526 (Celebes Sea, Sulu Sea, and Ceram Sea) and BOLD:AGE8639 (Java Sea), suggesting that the South China Sea region may represent an overlap zone between the two lineages. *Engyprosopon xystrias* (Hubbs, 1915) was first described from the vicinity of the Vincennes Strait, Japan, and is widely distributed throughout the Indo-West Pacific region. Within the South China Sea, the species has previously been reported from the coasts of China and Vietnam. Its occurrence along the coast of Pahang in the present study represents the southernmost record within the South China Sea and the first record of the species from Malaysian waters. Similarly, *Ostorhinchus septemstriatus* (Günther 1880), previously known from the eastern Indian Ocean and western Pacific regions, was also collected from the coast of Pahang, representing the first record of the species from Malaysian waters. However, the observed p-distance divergence of 2.76% from *O. septemstriatus* specimens from Queensland, Australia (BOLD:AAN3092) suggests the need for further taxonomic examination of the species recorded in this study.

*Etroplus suratensis*, commonly known as the green chromide, was the only introduced euryhaline species recorded in this study and was collected from the Straits of Johor. The species is considered native to Sri Lanka and possibly India (Daniels 2002; Lever 1996; Welcomme 1988) and is believed to have been introduced into Malaysia and Singapore through aquaculture activities and the aquarium trade, followed by escape into natural waters (Lever 1996). Although generally regarded as less destructive than some other introduced cichlids, *E. suratensis* may compete with native fishes for resources such as food and

breeding sites. Therefore, feral populations of *E. suratensis* in the region should be closely monitored and further investigated (Kasim et al. 2025; Ng and Tan 2010).

# Conservation Significance of Marine Fishes in the South China Sea

The Indo-West Pacific region, particularly the Coral Triangle and South China Sea, is recognized as a global hotspot for shark and ray diversity, supporting exceptionally high species richness and endemism, although the region is increasingly facing significant conservation threats (Pimiento et al. 2023). Consistent with previous studies on sharks and rays in the region (Loh et al. 2023; Clark-Shen et al. 2023), a high proportion of threatened or Near Threatened species was recorded in the present study, with 10 of the 12 elasmobranch species classified under threatened categories in the IUCN Red List, comprising five shark species and five stingray species. Several of these species, including *Rhynchobatus australiae*, *Hemigaleus microstoma*, *Rhizoprionodon oligolinx*, *Maculabatis gerrardi*, and *Gymnura poecilura*, are commercially important and commonly associated with coastal and trawl fisheries throughout the South China Sea region. Among the threatened species recorded, *Rhynchobatus australiae* is also listed under the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) Appendix II, indicating that international trade of the species is regulated to ensure utilization is not detrimental to its survival. The high proportion of threatened species observed reflects the ongoing global decline of elasmobranch populations, with nearly one-third of shark and ray species currently threatened with extinction (Dulvy et al. 2021). Major contributing factors include overexploitation, incidental capture in commercial fisheries, recreational fishing pressure, habitat degradation, and the international shark fin trade, where mislabelled and unregulated products continue to accelerate population declines (Bineesh et al. 2017; Clark-Shen et al. 2023). The sustainability and management of these resources are further complicated by the transboundary nature of the South China Sea, which poses challenges in achieving coordinated conservation and fisheries management among bordering countries and local communities.

In addition to elasmobranchs, several commercially important bony fishes recorded in this study were also classified as threatened or Near Threatened, including *Choerodon schoenleinii*, *Protonibea diacanthus*, *Scomberomorus commerson*, and *Rastrelliger brachysoma*. These species are heavily exploited by coastal and commercial fisheries throughout the Indo-Pacific region, raising increasing conservation concerns associated with overfishing and fisheries pressure.

Risk assessment and biodiversity conservation become increasingly challenging in the presence of hidden genetic diversity. Nine out of the eleven species exhibiting moderate to high genetic divergence (> 1%) are currently classified as either ‘Least Concern’ (*Siganus fuscescens*, *Nibea soldado*, *Rastrelliger kanagurta*, *Samaris cristatus*, *Pomadasys maculatus*, *Turrum coeruleopinnatum*, *Nemipterus peronii*, and *Netuma thalassina*) or ‘Data Deficient’ (*Synodus nigrotaeniatus*) under the IUCN (2026) criteria, suggesting that the actual conservation risks faced by these taxa may be underestimated due to insufficient taxonomic resolution. Overlooking cryptic diversity may therefore compromise conservation

strategies, particularly for economically important or locally adapted lineages possessing distinct ecological adaptations (Feckler et al. 2014). This issue is especially critical in regions experiencing increasing anthropogenic pressures, including pollution, coastal development, and overfishing (Ramírez-Ortiz et al. 2025). Consequently, failure to recognize cryptic diversification may result in the irreversible loss of unique genetic diversity, ultimately affecting biodiversity, species persistence, and the functional stability and resilience of marine ecosystems (Laikre et al. 2016).

Furthermore, a considerable number of species recorded in this study were classified as Data Deficient (12 species) or Not Evaluated (19 species), highlighting substantial gaps in taxonomic, ecological, and conservation knowledge. Such limitations complicate conservation planning, fisheries management, and future IUCN reassessments. Therefore, strengthened regional cooperation, standardized biodiversity monitoring, and the integration of molecular tools such as DNA barcoding are essential for improving species identification, facilitating species discovery, expanding DNA barcode reference libraries, supporting sustainable fisheries management, and enhancing long-term conservation strategies for threatened marine fishes in the region.

# Conclusions

This study revealed hidden diversity and significantly reduced knowledge gaps by establishing a robust DNA barcode reference library for marine fishes, particularly from Malaysia and the South China Sea, as well as surrounding waters. A total of 225 species representing 141 genera, 69 families, and 25 orders across the classes Elasmobranchii and Actinopterygii were successfully recorded and analyzed. The findings also indicated the presence of potential cryptic diversity within *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi*, and *Netuma thalassina*. This barcode reference library provides an important molecular resource that will facilitate future ichthyological research, including taxonomy, biodiversity assessment, conservation planning, and fisheries management. Overall, the study highlights the importance of integrating molecular approaches into marine biodiversity research and advocates for more comprehensive regional investigations to better understand, monitor, and conserve marine fish diversity in the region.

# Declarations

**CRediT authorship contribution statement**

**Jamsari Amirul Firdaus Jamaluddin**; Writing - original draft, conceptualization, methodology, investigation. **Noorul Azliana Jamaludin:** Funding acquisition, conceptualization, methodology, supervision. **Mohd Hans Afifi Hassan**: Conceptualization, investigation. **Hashim Shuhaimi**: Conceptualization, investigation. **Nurul Nadwa Abdul Fatah**: Conceptualization, investigation. **Nabilah Mustafha**: Conceptualization, investigation. **Nor Bariah Omaruddin**: Conceptualization, investigation. **Ryon Siow:** Supervision, project administration, funding acquisition. **Haryati Abd Wahab**:

Conceptualization, funding acquisition. All authors review and approved the draft and final version for publication.

**Conflict of Interest**

The authors report no declarations of interest.

**Ethical Approval**

Not applicable.

**Acknowledgements**

We would like to acknowledge the Ministry of Agriculture and Food Security Malaysia, through the Capture Fisheries Resources Division, Department of Fisheries Malaysia, for providing research funding under the Twelfth Malaysia Plan (RMK-12) (Development Allocation), Project for Strengthening National Fisheries Resources (P21300040170301), for the implementation of the project “Assessment of Coastal Demersal Resources on the East Coast of Peninsular Malaysia 2025”. We also extend our gratitude to Mr. Abdul Rahman bin Abdul Majid for his assistance in sample collection and species identification. Special thanks are also extended to the staff of the Marine Capture Fisheries Research Division, Fisheries Research Institute Kampung Acheh, and the Department of Fisheries Malaysia who were directly or indirectly involved in this project. Additionally, we acknowledge the assistance of OpenAI ChatGPT (GPT-4), a large language model, for language editing and formatting support during the preparation of this manuscript.

# References

1. Abidin DHZ, Nor SAM, Seah YG, Ali MS, Jamaluddin JAF, Rahim MA, Zulkifly BA, Tan NS, Zain MP, Jaafar KM TNAM (2024) An Odyssey of Integrative Taxonomy Unveils Marine Fish Diversity, New Records and Cryptic Species in Malaysian Waters. Zool Stud 63:e30. https://doi.org/10.6620/ZS.2024.63-30
2. Adrim M, Chen IS, Chen ZP, Lim KKP, Tan HH, Yusof Y, Jaafar Z (2004) Marine fishes recorded from the Anambas and Natuna Islands, South China Sea. Raffles Bull Zool Supplement No 11:117–130
3. Afiq-Firdaus MA, Lim CS, Che Din MS, Sheikh Abdul Kadir ST, Mohd Azmi Ramasamy A, Ab Manaf N, Mohd Sharol Ali M, Lachs L, Wright HE, Bachok Z (2021) Coral reef fish community at Pulau Bidong, Terengganu, South China Sea. J Sustain Sci Manag 16(5):48–66. https://doi.org/10.46754/jssm.2021.07.004
4. Ahmad Ali, Lim APK, Zakaria NA, Noor MS, Krajangdara T (2018) Marine fishes and crustaceans of the Southeast Asian region (SEAFDEC/MFRDMD/SP/40). Southeast Asian Fisheries Development Center, Marine Fishery Resources Development and Management Department, Samutprakan, Thailand

5. Alam M, Abbas K, Usmani N, Raza MT, Bhat MR, Mustafa M, Husain A (2024) A comprehensive review on DNA barcoding for species identification across diverse taxa. Munis Entomol Zool 19(2):1057–1072
6. Allen GR, Erdmann MV, Peristiwady T (2017) *Synodus nigrotaeniatus*, a new species of lizardfish (Aulopiformes: Synodontidae) from Indonesia. J Ocean Sci Foundation 26:59–67. http://dx.doi.org/10.5281/zenodo.546124
7. Ambak MA, Mat Isa M, Zakaria MZ, Abd Ghaffar M (2012) Fishes of Malaysia, 2nd edn. Penerbit Universiti Malaysia Terengganu, Terengganu, Malaysia
8. Amos R (2025) A Critical Analysis of the Global Biodiversity Framework. J Int Wildl Law Policy 28(2):123–192. https://doi.org/10.1080/13880292.2025.2539577
9. Azahar MA, Rosli NAS, Ab Rahman NS, Badrul Hisham NS, Solehin SN, Mohamed Rehan A, Zainal MZ, Abdul Latiff MAB, Giat SY, Kamarudin KR (2022) Species checklist and DNA barcoding of marine fishes from Pulau Tinggi, Johor, Malaysia. J Sustain Sci Manag 17(11):45–55. https://doi.org/10.46754/jssm.2022.11.006
10. Bemis KE, Girard MG, Santos MD, Carpenter KE, Deeds JR, Pitassy DE, Flores NAL, Hunter ES, Driskell AC, Macdonald KS 3rd, Weigt LA, Williams JT (2023) Biodiversity of Philippine marine fishes: A DNA barcode reference library based on voucher specimens. Sci Data 10(1):411. https://doi.org/10.1038/s41597-023-02306-9
11. Bineesh KK, Gopalakrishnan A, Akhilesh KV, Sajeela KA, Abdussamad EM, Pillai NGK, Basheer VS, Jena JK, Ward RD (2017) DNA barcoding reveals species composition of sharks and rays in the Indian commercial fishery. Mitochondrial DNA Part DNA Mapp Sequencing Anal 28(4):458–472. https://doi.org/10.3109/19401736.2015.1137900
12. Borsa P, Shen KN, Arlyza IS, Hoareau TB (2016) Multiple cryptic species in the blue-spotted maskray (Myliobatoidei: Dasyatidae: *Neotrygon* spp.): An update. CR Biol 339(9–10):417–426. https://doi.org/10.1016/j.crvi.2016.07.004
13. Cerutti-Pereyra F, Meekan MG, Wei NW, O'Shea O, Bradshaw CJ, Austin CM (2012) Identification of rays through DNA barcoding: an application for ecologists. PLoS ONE 7(6):e36479. https://doi.org/10.1371/journal.pone.0036479
14. Chang CH, Shao KT, Lin HY, Chiu YC, Lee MY, Liu SH, Lin PL (2017) DNA barcodes of the native ray-finned fishes in Taiwan. Mol Ecol Resour 17(4):796–805. https://doi.org/10.1111/1755-0998.12601
15. Chang CW, Shen KN, Chen IS (2025) A new species of dwarf goby of *Trimma* (Teleostei: Gobiidae) from Taiping Island, South China Sea. Zootaxa 5738(1):8–12. https://doi.org/10.11646/zootaxa.5738.1.4
16. Chong VC, Lee PK, Lau CM (2010) Diversity, extinction risk and conservation of Malaysian fishes. J Fish Biol 76(9):2009–2066. https://doi.org/10.1111/j.1095-8649.2010.02685.x
17. Clark-Shen N, Chin A, Arunrugstichai S, Labaja J, Mizrahi M, Simeon B, Hutchinson N (2023) Status of Southeast Asia's marine sharks and rays. Conserv Biology: J Soc Conserv Biology 37(1):e13962. https://doi.org/10.1111/cobi.13962

18. Daniels RJR (2002) Freshwater Fishes of Peninsular India. University, Hyderabad
19. Dellicour S, Flot JF (2015) Delimiting Species-Poor Data Sets using Single Molecular Markers: A Study of Barcode Gaps, Haplowebs and GMYC. Syst Biol 64(6):900–908. https://doi.org/10.1093/sysbio/syu130
20. Du J, Loh KH, Hu W, Zheng X, Amri AY, Ooi JLS, Ma Z, Rizman-Idid M, Chan AA (2019) An updated checklist of the marine fish fauna of Redang Islands, Malaysia. Biodivers Data J 7:e47537. https://doi.org/10.3897/BDJ.7.e47537
21. Dulvy NK, Pacoureau N, Rigby CL, Pollom RA, Jabado RW, Ebert DA, Finucci B, Pollock CM, Cheok J, Derrick DH, Herman KB, Sherman CS, VanderWright WJ, Lawson JM, Walls RHL, Carlson JK, Charvet P, Bineesh KK, Fernando D, Ralph GM, Matsushiba JH, Hilton-Taylor C, Fordham SV, Simpfendorfer CA (2021) Overfishing drives over one-third of all sharks and rays toward a global extinction crisis. Curr Biol 31(21):4773–4787e8. https://doi.org/10.1016/j.cub.2021.08.062
22. Feckler A, Zubrod JP, Thielsch A, Schwenk K, Schulz R, Bundschuh M (2014) Cryptic species diversity: An overlooked factor in environmental management? J Appl Ecol 51(4):958–967. https://doi.org/10.1111/1365-2664.12246
23. Filonzi L, Ardenghi A, Rontani PM, Voccia A, Ferrari C, Papa R, Bellin N, Nonnis Marzano F (2023) Molecular Barcoding: A Tool to Guarantee Correct Seafood Labelling and Quality and Preserve the Conservation of Endangered Species. Foods 12(12):2420. https://doi.org/10.3390/foods12122420
24. Fricke R, Eschmeyer WN, Van der Laan R (eds) (2025) Eschmeyer's catalog of fishes: Genera, species, references. http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp. Accessed January 2025
25. Froese R, Pauly D (eds) (2026) FishBase. World Wide Web electronic publication. www.fishbase.org. Accessed February 2026
26. Gaither MR, Bowen BW, Bordenave TR, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT (2011) Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. BMC Evol Biol 11:189. https://doi.org/10.1186/1471-2148-11-189
27. Gonzalez F, Barria P, Ponce F, Mora S (2024) In: Genetics (ed) Population Genetic Structure of Marine Fishes. IntechOpen, London. https://doi.org/10.5772/intechopen.112694
28. Goulding TC, Dayrat B (2023) The Coral Triangle and Strait of Malacca are two distinct hotspots of mangrove biodiversity. Sci Rep 13:15793. https://doi.org/10.1038/s41598-023-42057-6
29. Günther A (1880) Report on the shore fishes procured during the voyage of H. M. S. Challenger in the years 1873–1876. In: Report on the scientific results of the voyage of H. M. S. Challenger during the years 1873-76. Vol. 1, Part 6. pp 1–82. https://www.biodiversitylibrary.org/part/145825
30. Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of birds through DNA barcodes. PLoS Biol 2(10):e312. https://doi.org/10.1371/journal.pbio.0020312

31. Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc B 270(1512):313–321. https://doi.org/10.1098/rspb.2002.2218
32. Huang H, Chen J, Ke Z, Zhang C (2025) *Branchiostegus sanae*, a new species of deepwater tilefish (Eupercaria, Branchiostegidae) from the South China Sea. ZooKeys 1227:129–142. https://doi.org/10.3897/zookeys.1227.130512
33. Hubbs CL (1915) Flounders and soles from Japan collected by the United States Bureau of Fisheries steamer Albatross in 1906. Proceedings of the United States National Museum 48(2082):449–496. https://doi.org/10.5479/si.00963801.48-2082.449
34. Hubert N, Hanner R (2015) DNA barcoding, species delineation and taxonomy: A historical perspective. DNA Barcodes 3:44–58. https://doi.org/10.1515/dna-2015-0006
35. Hubert N, Phillips JD, Hanner RH Delimiting Species with Single-Locus DNA Sequences. Methods in Molecular Biology (, Clifton NJ (2024) 2744, 53–76. https://doi.org/10.1007/978-1-0716-3581-0_3
36. Human BA (2007) A taxonomic revision of the catshark genus *Haploblepharus* Garman 1913 (Chondrichthyes: Carcharhiniformes: Scyliorhinidae). Zootaxa 1451(1):1–40. https://doi.org/10.11646/zootaxa.1451.1.1
37. Jamaluddin JAF, Lavoué S, Alshari NFMAH, Ghazali SZB, Ahmad A, Tan A, Leng CL, Nor SAM (2022) Reassessing fish diversity of Penang Island's freshwaters (northwest Peninsular Malaysia) through a molecular approach raises questions on its conservation status. Biodivers Conserv 31:1551–1576. https://doi.org/10.1007/s10531-022-02409-6
38. Jiang C, Liu F, Qin J, Hubert N, Kang B, Huang L, Yan Y (2024) DNA barcode reference library of the fish larvae and eggs of the South China Sea: taxonomic effectiveness and geographic structure. BMC Ecol Evol 24(1):132. https://doi.org/10.1186/s12862-024-02316-0
39. Kapli P, Lutteropp S, Zhang J, Kobert K, Pavlidis P, Stamatakis A, Flouri T (2017) Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics 33(11):1630–1638. https://doi.org/10.1093/bioinformatics/btx025
40. Kasim NS, Mat Jaafar TNAM, Tan MP, Mohd Yusoff NIS, Alshari NFMAH, Mohd Nur NF, Mohd Sharol A, Sheikh Abdul Kadir ST, Lavoué S, Abidin DHZ, Rahim MA, Jamaludin NA, Ahmad A, Aqmal-Naser M, Ghazali SZ, Seah YG, Mohd Nor SA (2025) Ichthyofauna of the largest wetlands of East Peninsular Malaysia and its adjacent waters: The Setiu Wetlands. Thalassas 41:117. https://doi.org/10.1007/s41208-025-00870-y
41. Ke C, Gu YG, Huang X, Wu W, Ma Z, Song C (2022) Editorial: Fisheries ecological environment in South China Sea. Front Environ Sci 10:981443. https://doi.org/10.3389/fenvs.2022.981443
42. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096
43. Laikre L, Lundmark C, Jansson E, Wennerström L, Edman M, Sandström A (2016) Lack of recognition of genetic biodiversity: International policy and its implementation in Baltic Sea marine protected areas. Ambio 45(6):661–680. https://doi.org/10.1007/s13280-016-0776-7

44. Lassmann T, Sonnhammer EL (2005) Kalign–an accurate and fast multiple sequence alignment algorithm. BMC Bioinformatics 6:298. https://doi.org/10.1186/1471-2105-6-298
45. Lever C (1996) Naturalized Fishes of the World. Academic, London
46. Li H, Chen Y, Li X, Zhou P, Xiong X (2023) Assessing larval fish diversity and conservation needs in the Luzon Strait using DNA barcoding. Front Mar Sci 10:1268399. https://doi.org/10.3389/fmars.2023.1268399
47. Li Y, Chen Z, Zhang J (2022) Fish Composition and Diversity of Four Coral Reefs in the South China Sea Based on Hand-Line Catch. J Mar Sci Eng 10(1):38. https://doi.org/10.3390/jmse10010038
48. Lim HC, Habib A, Chen WJ (2021) Comparative phylogeography and phylogeny of Pennah croakers (Teleostei: Sciaenidae) in Southeast Asian waters. Genes 12(12):1926. https://doi.org/10.3390/genes12121926
49. Lo PC, Liu SH, Nor SAM, Chen WJ (2017) Molecular exploration of hidden diversity in the Indo-West Pacific sciaenid clade. PLoS ONE 12(4):e0176623. https://doi.org/10.1371/journal.pone.0176623
50. Loh KH, Lim KC, Then AY, Adam S, Leung AJ, Hu W, Bong CW, Wang A, Sade A, Musel J, Du J (2023) Advancing DNA Barcoding to Elucidate Elasmobranch Biodiversity in Malaysian Waters. Animals 13(6):1002. https://doi.org/10.3390/ani13061002
51. Mat Jaafar TNAM, Seah YG, Imamura H, Motomura H, Matsunuma M (2025) *Rogadius azizahae*, a new flathead (Perciformes: Platycephalidae) from the South China Sea. J Fish Biol 107(5):1702–1713. https://doi.org/10.1111/jfb.70135
52. Matsunuma M, Motomura H, Matsuura K, Shazili NAM, Ambak MA (eds) (2011) Fishes of Terengganu – East Coast of Malay Peninsula, Malaysia. National Museum of Nature and Science, Universiti Malaysia Terengganu, and Kagoshima University Museum, Tokyo/Terengganu/Kagoshima
53. Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 3(12):e422. https://doi.org/10.1371/journal.pbio.0030422
54. Mohsin AKM, Ambak MA (1996) Marine fishes and fisheries of Malaysia and neighbouring countries. University Pertanian Malaysia, Serdang, Malaysia
55. Mora C, Tittensor DP, Myers RA (2008) The completeness of taxonomic inventories for describing the global diversity and distribution of marine fishes. Proc R Soc B 275(1631):149–155. https://doi.org/10.1098/rspb.2007.1315
56. Morton B, Blackmore G (2001) South China Sea. Mar Pollut Bull 42(12):1236–1263. https://doi.org/10.1016/S0025-326X(01)00240-5
57. Muda O, Isa MM, Vinh CT, Nasir MTM, Zainal A (2001) Fish taxonomic studies in the South China Sea, Area IV: Vietnamese waters. In: Proceedings of the Fourth Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area IV: Vietnamese Waters, 18–20 September 2000. Bangkok, Thailand: Secretariat, Southeast Asian Fisheries Development Center. pp 41–54
58. Neigel JE, Domingo A, Stake JL (2007) DNA barcoding as a tool for coral reef conservation. Coral Reefs 26(3):487–499. https://doi.org/10.1007/s00338-007-0248-4

59. Ng CZW, Reis-Santos P, Gonzalez JG, Gillanders BM, Saleh MF, Ong JJL (2025) A non-linear statistical framework to investigate changes in life history patterns within and among fish populations. Rev Fish Biol Fish 35:2065–2080. https://doi.org/10.1007/s11160-025-09993-0
60. Ng PKL, Tan KS (2000) The state of marine biodiversity in the South China Sea. Raffles Bull Zool Supplement No 8:3–7
61. Ng SL, Liu HW, Mediodia DP, Lin YT, Lee CH, Lin CF, Huang SP, Wu SM, Tung CR, Ho HC, Lin CH (2024) An updated checklist of fishes of Dongsha Island, Taiwan, northern South China Sea. ZooKeys 1220:175–242. https://doi.org/10.3897/zookeys.1220.131100
62. Ng TH, Tan HH (2010) The introduction, origin and life-history attributes of the non-native cichlid *Etroplus suratensis* in the coastal waters of Singapore. J Fish Biol 76(9):2238–2260. https://doi.org/10.1111/j.1095-8649.2010.02642.x
63. Nunes DBSM, Pontes AI, dos Santos LF, Calado LL, Barros-Neto LF, Lustosa-Costa SY, Lima SMQ, Jacobina UP (2024) DNA barcoding highlights taxonomic uncertainties and cryptic lineages of São Francisco River basin fishes. Neotropical Ichthyol 22(4):e240059. https://doi.org/10.1590/1982-0224-2024-0059
64. Pimiento C, Albouy C, Silvestro D, Di Marco M, Bruno JF, Carlson JK, Ceballos G, Dulvy NK, Harrison SP, Jetz W, MacIvor JS, Queiroz N, Griffin JN (2023) Functional diversity of sharks and rays is highly vulnerable and supported by unique species and locations worldwide. Nat Commun 14:7691. https://doi.org/10.1038/s41467-023-43212-3
65. Puillandre N, Brouillet S, Achaz G (2021) ASAP: Assemble Species by Automatic Partitioning. Mol Ecol Resour 21(2):609–620. https://doi.org/10.1111/1755-0998.13281
66. Ramírez-Ortiz G, Valencia-Méndez O, Hernández L, González-Mendoza T, López-Pérez A (2025) Conservation of fish diversity in protected sites and adjacent fishing areas of the Southern Mexican Pacific Ocean. PLoS ONE 20(6):e0324155. https://doi.org/10.1371/journal.pone.0324155
67. Randall JE, Lim KKP (eds) (2000) A checklist of the fishes of the South China Sea. The Raffles Bulletin of Zoology Supplement No. 8:569–667
68. Ratnasingham S, Hebert PD (2007) BOLD: The Barcode of Life Data System. Mol Ecol Notes 7(3):355–364. http://www.barcodinglife.org https://doi.org/10.1111/j.1471-8286.2007.01678.x
69. Ratnasingham S, Hebert PD (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. PLoS ONE 8(7):e66213. https://doi.org/10.1371/journal.pone.0066213
70. Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Glob Ecol Conserv 17:e00547. https://doi.org/10.1016/j.gecco.2019.e00547
71. Seah YG, Ali MS, Ghaffar MA, Mat Jaafar TNAM (2021) Marine fishes of Kuantan, Malaysia Biodiversity Information System (MYBIS). Penerbit Universiti Malaysia Terengganu, Terengganu, Malaysia
72. Seah YG, Mat Jaafar TNAM, Ali MS (2020) Field guide to trawl fishes near Bidong Island. Penerbit Universiti Malaysia Terengganu, Terengganu, Malaysia

73. Seijas-Díaz I, Presa P (2025) Towards a more integrative paradigm in fisheries assessment: Genetic reference points. Front Mar Sci 12:1524309. https://doi.org/10.3389/fmars.2025.1524309
74. Shneyer VS, Rodionov AV (2025) 20 years of DNA barcoding – Achievements and problems. Biochem (Mosc) 90(11):1602–1619. https://doi.org/10.1134/S0006297925602977
75. Silvestro D, Michalak I (2012) raxmlGUI: A graphical front-end for RAxML. Organisms Divers Evol 12(4):335–337. https://doi.org/10.1007/s13127-011-0056-0
76. Staats M, Arulandhu AJ, Gravendeel B, Holst-Jensen A, Scholtens I, Peelen T, Prins TW, Kok E (2016) Advances in DNA metabarcoding for food and wildlife forensic species identification. Anal Bioanal Chem 408(17):4615–4630. https://doi.org/10.1007/s00216-016-9595-8
77. Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313. https://doi.org/10.1093/bioinformatics/btu033
78. Suhailan MAA, Fakhruddin NF, Jamaluddin JAF, Jamaludin NA, Hurzaid A (2025) DNA barcoding and biodiversity of commercially important fishes (Actinopterygii) from Middle Bank, Penang, Malaysia: A first report. IOP Conference Series: Earth and Environmental Science 1508(1):012002. https://doi.org/10.1088/1755-1315/1508/1/012002
79. Tang D, Sun Z, Su G (2023) Geological environment in the South China Sea. J Oceanol Limnol 41(2):403–408. https://doi.org/10.1007/s00343-023-3403-2
80. Thu PT, Huang WC, Chou TK, Van Quan N, Van Chien P, Li F, Shao KT, Liao TY (2019) DNA barcoding of coastal ray-finned fishes in Vietnam. PLoS ONE 14(9):e0222631. https://doi.org/10.1371/journal.pone.0222631
81. Vidthayanon C (1999a) Species composition and diversity of fishes in the South China Sea, Area I: Gulf of Thailand and east coast of Peninsular Malaysia. In: Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I: Gulf of Thailand and Peninsular Malaysia, 24–26 November 1997. Bangkok, Thailand: Training Department, Southeast Asian Fisheries Development Center. pp 172–240
82. Vidthayanon C (1999b) Species composition and diversity of fishes in the South China Sea, Area II: Sarawak, Sabah and Brunei Darussalam waters. In: Proceedings of the Second Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area II: West Coast of Sabah, Sarawak and Brunei Darussalam, 14–15 December 1998. Kuala Lumpur, Malaysia: Training Department, Southeast Asian Fisheries Development Center. pp 224–242
83. Voris HK (2000) Maps of Pleistocene sea levels in Southeast Asia: Shorelines, river systems and time durations. J Biogeogr 27(5):1153–1167. https://doi.org/10.1046/j.1365-2699.2000.00489.x
84. Vrijenhoek RC (2009) Cryptic species, phenotypic plasticity, and complex life histories: Assessing deep-sea faunal diversity with molecular markers. Deep Sea Res Part II 56(19–20):1713–1723. https://doi.org/10.1016/j.dsr2.2009.05.016
85. Wang P, Li Q, Tian J (2014) Pleistocene paleoceanography of the South China Sea: Progress over the past 20 years. Mar Geol 352:381–396. https://doi.org/10.1016/j.margeo.2014.03.003

86. Wang ZD, Guo YS, Liu XM, Fan YB, Liu CW (2012) DNA barcoding South China Sea fishes. Mitochondrial DNA 23(5):405–410. https://doi.org/10.3109/19401736.2012.710204
87. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD (2005) DNA barcoding Australia's fish species. Philos Trans R Soc B 360(1462):1847–1857. https://doi.org/10.1098/rstb.2005.1716
88. Welcomme RL (1988) International introductions of inland aquatic species. FAO Fisheries Technical Paper 294. Rome: FAO
89. Wibowo K, Cahyani NKD, Larashati S, Afrisal M, Wahyudewantoro G, Oktaviyani S, Kusumah RV (2024) Range extension of blackstripe lizardfish, *Synodus nigrotaeniatus* (Actinopterygii, Aulopiformes, Synodontidae), from Jakarta Bay. Indonesia Check List 20(3):601–606. https://doi.org/10.15560/20.3.601
90. WoRMS Editorial Board (2026) World Register of Marine Species. https://www.marinespecies.org/. Accessed 16 April 2026
91. Xiao Y, Li C, Wang T, Lin L, Guo J, Quan Q, Liu Y (2022) DNA barcoding revealing the parrotfish (Perciformes: Scaridae) diversity of the coral reef ecosystem of the South China Sea. Sustainability 14(22):15386. https://doi.org/10.3390/su142215386
92. Xing B, Lin L, Wu Q (2025) Application of mitochondrial genomes to species identification and evolution. Electron J Biotechnol 76:39–48. https://doi.org/10.1016/j.ejbt.2025.04.001
93. Xiong F, Shu L, Zeng H, Gan X, He S, Peng Z (2022) Methodology for fish biodiversity monitoring with environmental DNA metabarcoding: The primers, databases and bioinformatic pipelines. Water Biology Secur 1(1):100007. https://doi.org/10.1016/j.watbs.2022.100007
94. Xu L, Van Damme K, Li H, Ji Y, Wang X, Du F (2019) A molecular approach to the identification of marine fish of the Dongsha Islands (South China Sea). Fish Res 213:105–112. https://doi.org/10.1016/j.fishres.2019.01.011
95. Xu L, Wang X, Van Damme K, Huang D, Li Y, Wang L, Ning J, Du F (2021) Assessment of fish diversity in the South China Sea using DNA taxonomy. Fish Res 233:105771. https://doi.org/10.1016/j.fishres.2020.105771
96. Yusri Y, Seah YG, Izarenah MR, Lee JN (2021) First record of spotted knifejaw, *Oplegnathus punctatus* (Temminck & Schlegel, 1844) (Oplegnathidae) in the southern South China Sea. Check List 17(4):1195–1198. https://doi.org/10.15560/17.4.1195
97. Zhang J, Xu L, Du F, Tang Q, Wang L, Ning J, Huang D, Li Y, Liu S, Wang X (2023) DNA barcoding of marine fish species in the waters surrounding Hainan Island, northern South China Sea. Front Mar Sci 10:1249073. https://doi.org/10.3389/fmars.2023.1249073
98. Zhang J (2011) Species identification of marine fishes in China with DNA barcoding. Evidence-Based Complement Altern Med 2011:978253. https://doi.org/10.1155/2011/978253
99. Zhang J, Hanner R (2012) Molecular approach to the identification of fish in the South China Sea. PLoS ONE 7(2):e30621. https://doi.org/10.1371/journal.pone.0030621

# Figures

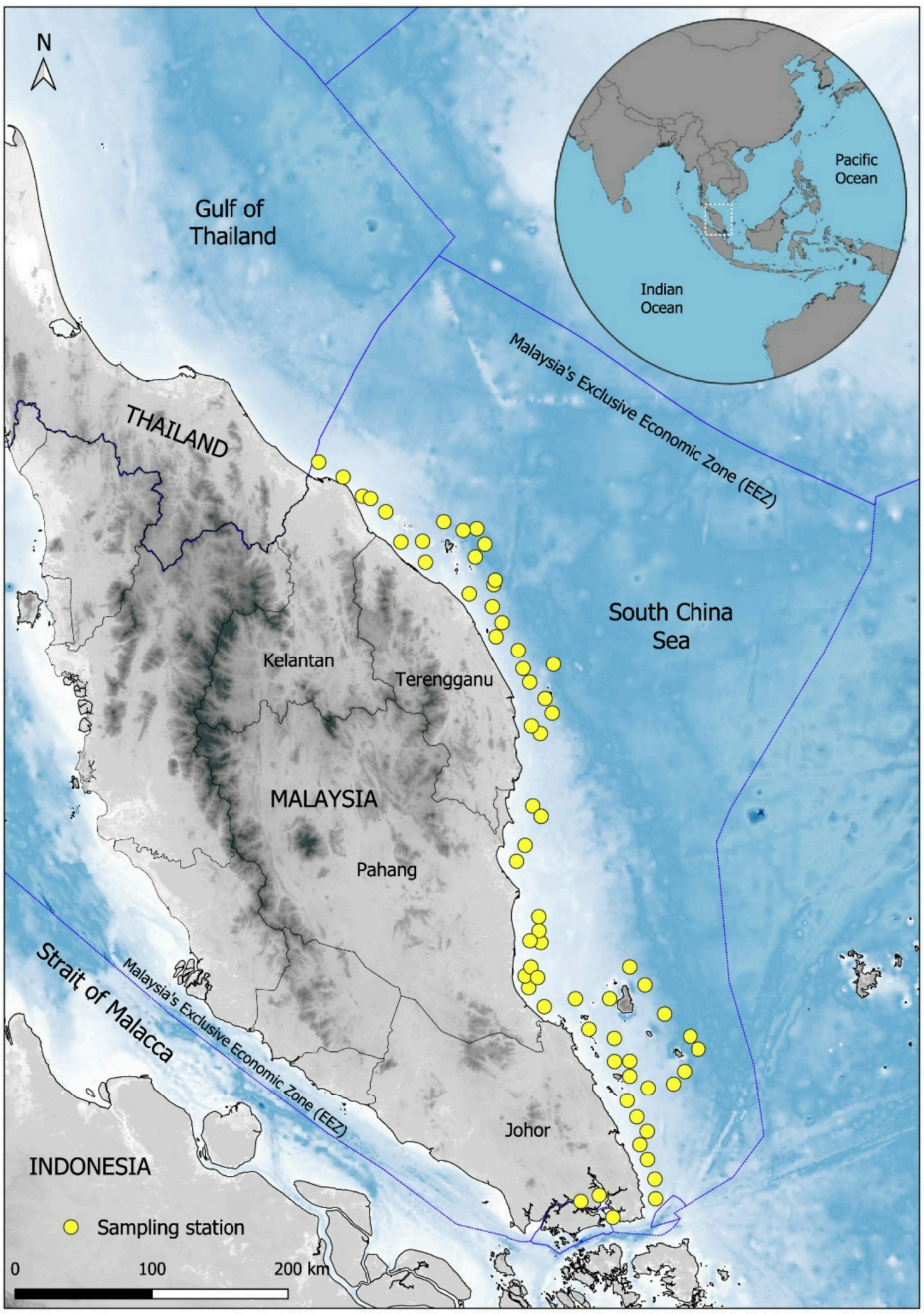



**Figure 1**

Map showing the distribution of survey stations (yellow circles) within the Exclusive Economic Zone (EEZ) of the east coast of Peninsular Malaysia.

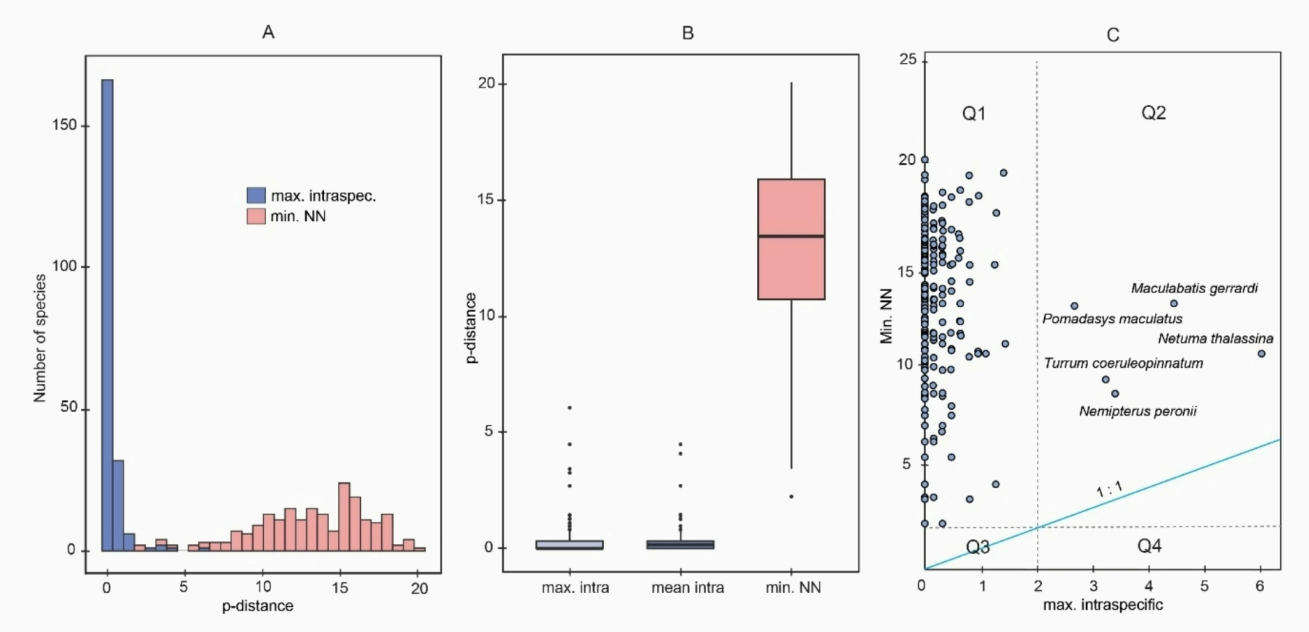
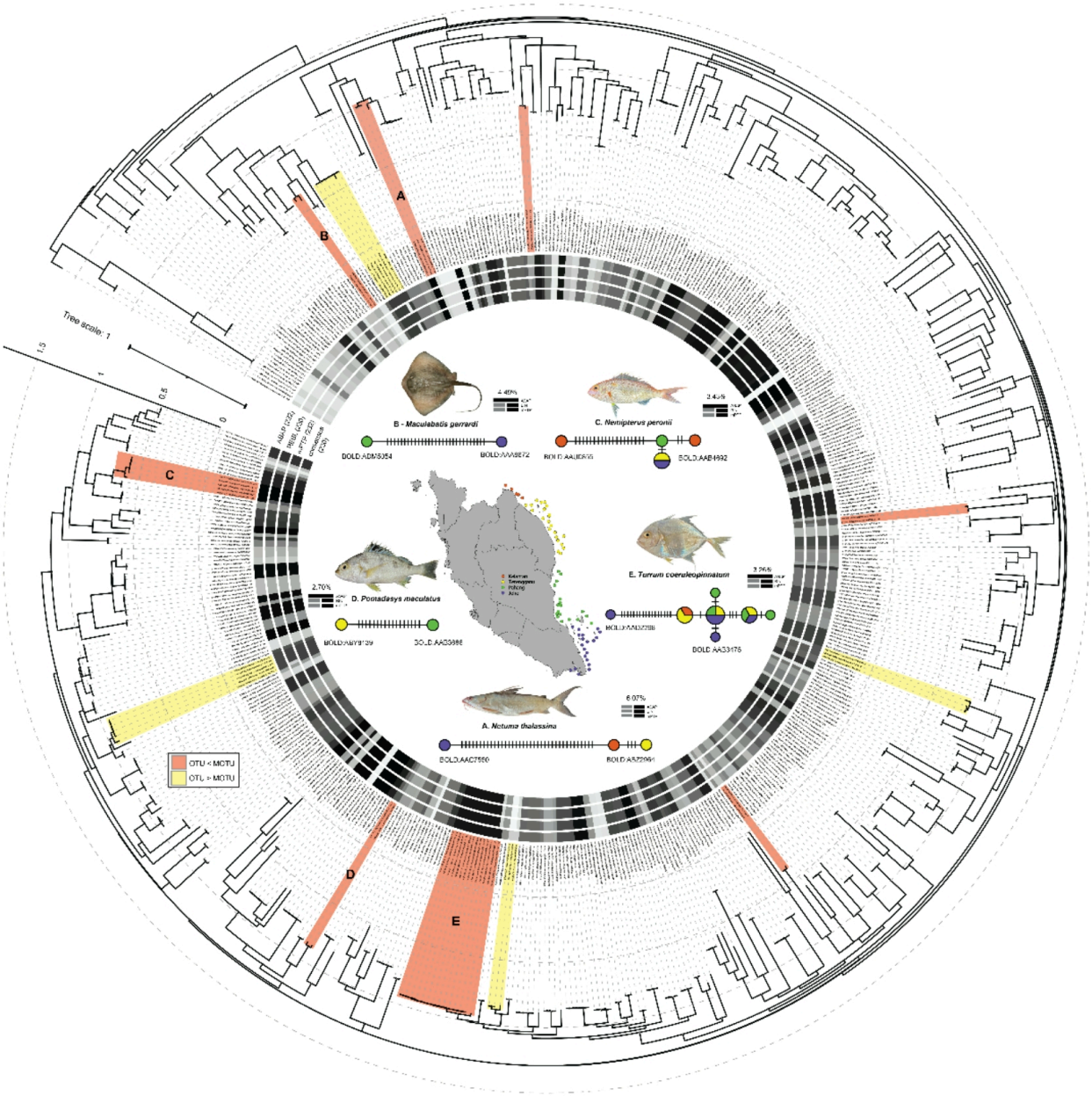



**Figure 2**

Intraspecific and nearest-neighbor genetic distances calculated using the p-distance model. (A) Histograms showing the distribution of maximum intraspecific distances (max. intra) and minimum nearest-neighbor distances (min. NN) across taxa. (B) Boxplots of mean intraspecific (mean intra), maximum intraspecific (max. intra), and minimum nearest-neighbor distances (min. NN). (C) Scatterplot of maximum intraspecific versus minimum interspecific (nearest-neighbor) distances, illustrating the barcoding gap. The dashed line represents the 2% divergence threshold. Species in Quadrant 1 (Q1) show congruence between morphological and molecular identification. Quadrant 2 (Q2) indicates high intraspecific divergence, suggesting potential cryptic species. Quadrant 3 (Q3) includes taxa with low interspecific divergence, possibly reflecting recent divergence, hybridization, or synonymy. Quadrant 4 (Q4) shows high intraspecific but low interspecific divergence, indicating discrepancies that warrant further taxonomic investigation.




## Figure 3

Maximum Likelihood (ML) phylogenetic tree showing relationships among 225 marine fish species from the Exclusive Economic Zone (EEZ) of the east coast of Peninsular Malaysia. Molecular Operational Taxonomic Units (MOTUs) were delimited using ASAP, RESL, and mPTP, methods, including a consensus delimitation. Highlighted sections indicate five cryptic species (>1 BIN). The inset displays a haplotype network for *Pomadasys maculatus, Turrum coeruleopinnatum, Nemipterus peronii, Maculabatis gerrardi,* and*Netuma thalassina*.