Zoological Studies **63:**30 (2024)

An Odyssey of Integrative Taxonomy Unveils Marine Fish Diversity, New Records and Cryptic Species in Malaysian Waters

Danial Hariz Zainal Abidin[1], Siti Azizah Mohd. Nor[2,6], Giat Seah Ying[3,4], Mohd Sharol Ali[3], Jamsari Amirul Firdaus Jamaluddin[6], Masazurah A Rahim[5], Nur Syafiqa Zulkifly[3], Min Pau Tan[2], Khaironizam Md Zain[6], and Tun Nurul Aimi Mat Jaafar[3,*]

[1]*Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, 43600, Selangor, Malaysia. E-mail: geneprodigy@gmail.com (Zainal Abidin)*

[2]*Institute of Climate Adaptation and Marine Biotechnology, Universiti Malaysia Terengganu, 21030, Kuala Nerus, Terengganu, Malaysia. E-mail: s.azizah@umt.edu.my (Nor); mptan@umt.edu.my (Tan)*

[3]*Faculty of Fisheries and Food Science, Universiti Malaysia Terengganu, 21030, Kuala Nerus, Terengganu, Malaysia. *Correspondence: E-mail: tun_aimi@umt.edu.my (Mat Jaafar)*
E-mail: ygseah@umt.edu.my (Ying); sharol@umt.edu.my (Ali); nursyafiqazulkifly97@gmail.com (Zulkifly); tun_aimi@umt.edu.my (Mat Jaafar)

[4]*South China Sea Repository & Reference Centre, Institute Oceanography and Environment, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia*

[5]*Fisheries Research Institute, 11960, Batu Maung, Penang, Malaysia. E-mail: masazurah@dof.gov.my (A Rahim)*

[6]*School of Biological Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia. E-mail: jamsari_85@yahoo.com (Jamaluddin); khaironizam@usm.my (Md Zain)*

(Received 27 December 2023 / Accepted 26 June 2024 / Published -- 2024)
Communicated by Ryuji Machida

ORCiD:
Danial Hariz Zainal Abidin: https://orcid.org/0000-0002-2654-6306
Siti Azizah Mohd. Nor: https://orcid.org/0000-0002-9890-9612
Ying Giat Seah: https://orcid.org/0000-0002-2976-4448
Jamsari Amirul Firdaus Jamaluddin: https://orcid.org/0000-0002-3792-6873
Min Pau Tan: https://orcid.org/0000-0003-0149-5247
Khaironizam Md Zain: https://orcid.org/0000-0002-4230-0779
Tun Nurul Aimi Mat Jaafar: https://orcid.org/0000-0002-3032-0545

This study elucidates the species diversity of marine fishes in the Exclusive Economic Zone (EEZ) of Peninsular Malaysia (PM) using an integrative approach combining DNA barcoding and morphological identification. Our focus was on demersal surveys conducted on the east coast of PM in the South China Sea. We re-evaluated the diversity of 475 specimens across 93 putative species (92 barcoded morphospecies), from 16 orders and 41 families, including two IUCN vulnerable species. A total of two species - *Saurida isarankurai* and *Oxyurichthys auchenolepis* are presented

as new record and three species - *Nemipterus balinensoides*, *Gymnothorax reevesii* and *Synodus hoshinonis* as first specimen-based record in Malaysian waters. Cytochrome *c* oxidase subunit I (*COI*) sequence analyses delineated 95 consensus Molecular Operational Taxonomic Units (MOTUs), exceeding morphological diversity. Interestingly, the barcode analysis revealed several MOTUs delimited within one morphologically identified fish species, with both intraspecific and interspecific genetic divergences exceeding 2%, indicating substantial intraspecific genetic divergence within species groups or the existence of morphologically cryptic species within our dataset. These findings highlight the complexity of species delimitation and the value of genetic methods. Our study provides valuable insights into marine fish diversity from the east coast of Peninsular Malaysia and enhances our understanding of genetic diversity, distribution, and conservation needs of ecosystems through DNA barcoding. By integrating DNA barcoding with morphology, we present a comprehensive framework for future research to develop conservation and management strategies for Malaysia's marine biodiversity. The expansion of the genetic barcode database generated in this study will facilitate future molecular taxonomy research.

Key words: Peninsular Malaysia, DNA barcoding, Exclusive Economic Zone (EEZ), South China Sea, Marine fishes, Demersal survey, Cryptic species

Citation: Zainal Abidin DH, Nor SAM, Ying GS, Ali MS, Jamaluddin JAF, A Rahim M, Zulkifly NS, Tan MP, Md Zain K, Mat Jaafar TNA. 2024. An Odyssey of Integrative Taxonomy Unveils Marine Fish Diversity, New Records and Cryptic Species in Malaysian Waters. Zool Stud **63:**30.

BACKGROUND

Understanding the biological diversity within marine ecosystems is of paramount importance for the conservation and sustainable management of these critical habitats. Notably, fish biodiversity plays a crucial role in maintaining ecosystem functionality, economic well-being, and food security (Mora et al. 2011). Situated in the South China Sea, the Exclusive Economic Zone (EEZ) of Peninsular Malaysia (PM), which spans an area of nearly 130,000 km^2 (Nadira et al. 2019), is a productive hotspot for marine biodiversity and hosts an impressive diversity of fish species (Allen 2008; Myers et al. 2000). These vibrant fish communities contribute significantly to the region's ecological equilibrium and form a vital part of Malaysia's economy and local livelihoods (Teh and Pauly 2018). To date, there is no large-scale and comprehensive ichthyofaunal assessment been done in this area, except for few smaller scale regional studies (Chong et al. 2010; Du et al. 2019; Matsunuma et al. 2011; Motomura et al.; Seah et al. 2020; 2021). Nonetheless, our

understanding of this biodiversity remains partial, largely hindered by the challenges of morphological identification.

The South China Sea and surrounding Southeast Asian waters are home to an incredible diversity of marine fish species. With over 3,300 recorded species from more than 250 families (Allen et al. 2000; Froese and Pauly 2023), this biodiversity hotspot harbours high diversity of marine ichthyofauna. Malaysia, Indonesia, the Philippines and Vietnam account for most of this regional diversity (Allen and Adrim 2003). Recent molecular studies utilising DNA barcoding have identified more than 116 putative fish species from major fishing regions in the South China Sea (Xu et al. 2021). Additionally, a four-year (2015-2018) survey of demersal fish communities of the South China Sea revealed a remarkably diverse assemblage of over 250 fish species inhabiting its northern continental slope waters (Zhang et al. 2022). While there is no definitive current estimate, experts estimate the total fish species in this region of reaching >3,700 species across the entire South China Sea basin, reflecting diverse neritic, oceanic and coral reef habitats (Pauly and Liang 2020). However, our understanding of this exceptional biodiversity, especially of the highly diverse fish community, remains insufficient.

Several factors contribute to the exceptional marine fish diversity observed in this region. The wide range of habitats, including over 160,000 km^2 of coral reefs, mangroves, and estuaries, provides unique ecological niches that promote population differentiation and ultimately speciation over time (Burke and Selig 2002). This habitat heterogeneity supports and accelerates divergence among fish assemblages (Burke and Selig 2002). Complex circulation patterns and seasonal monsoons also enhance productivity and connectivity between populations within this area (Fang et al. 2010). Nonetheless, major knowledge gaps remain regarding the taxonomy, biogeography, ecology, and population genetics of South China Sea fishes. Many cryptic and undescribed species are likely to exist (Hou et al. 2018; Mat Jaafar et al. 2012; Puckridge et al. 2013). Connectivity patterns between regional fishery stocks are poorly resolved, hampering spatial management (Pauly and Liang 2020). In addition, there is a lack of data on vulnerable and data-deficient species within this important region (Huang et al. 2016). Thus, targeted biodiversity assessments, DNA barcoding, and range-wide ecological studies are urgently needed to better characterise and conserve the marine biodiversity in this region.

Over the past two decades, DNA barcoding has emerged as a formidable tool for species identification and biodiversity assessment. By analysing a short sequence from a standardized region of the genome, typically the mitochondrial cytochrome oxidase I gene (*COI*) in animals, DNA barcoding facilitates rapid, precise, and cost-effective species identification (Hebert et al., 2003). The generation of DNA barcodes (species-specific sequences) provides diagnostic markers that supplement classical morphological taxonomy and accelerate taxonomic identification and

discoveries, but they are not intend to replace it (DeSalle et al. 2005). Its efficacy in delimiting marine and freshwater fishes has been demonstrated in regional barcoding studies, including those conducted in the Indian Ocean (Lakra et al. 2011), among Indo-Pacific coral reef fishes (Hubert et al. 2012), Indonesian freshwater species (Hubert et al. 2015), and in the west coast of PM (Zainal Abidin et al. 2021). When coupled with morphological techniques, barcoding offers comprehensive and reliable insights into biodiversity (Ward et al. 2009), unveiling cryptic and potentially novel species (Mat Jaafar et al. 2012; Seah et al. 2017; Zainal Abidin et al. 2021).

Leveraging on these advancements, this study aims to elucidate the marine fish diversity within Peninsular Malaysia's east coast EEZ by integrating DNA barcoding and morphological identification. The findings contribute to building a comprehensive genetic reference database of local fish diversity, serving as a valuable resource for future environmental DNA (eDNA) metabarcoding assessments – an increasingly important method for monitoring marine biodiversity (Alshari et al. 2021; Zainal Abidin et al. 2022). By providing a robust genetic baseline, our study enables effective utilization of eDNA techniques to track changes in fish communities and assess impacts of environmental factors or management strategies. We delve into the genetic diversity of regional fish populations, offering valuable insights to facilitate their preservation and sustainable management. Ultimately, this study underscores the power of amalgamating traditional and contemporary methodologies to decode complex marine biodiversity. It establishes a comprehensive framework, encompassing genetic reference data and morphological records, to guide future research and decision-making processes for conserving and managing Malaysia's marine biodiversity and fishery resources.

MATERIALS AND METHODS

Sample collection

A total of 475 fish specimens were collected during demersal surveys conducted from May to July 2016 in the Exclusive Economic Zone (EEZ) along the East Coast of PM (ECPM). The surveys were organized by the Department of Fisheries Malaysia with the use of bottom trawls onboard the research vessel MV SEAFDEC II. Sampling was performed at 41 stations distributed within the EEZ (Fig. 1). The sample was collected using a bottom trawl with a 40 mm cod end mesh net. The trawl sampling lasted for 60 minutes at a speed of 3.2 knots, covering 3 nautical miles. Information on sampling locations (geographical coordinates), collection data, taxonomy and

details of voucher specimens can be found in the online project dataset implemented in the Barcode of Life Database (BOLD) under the project code 'DBEEZ'.

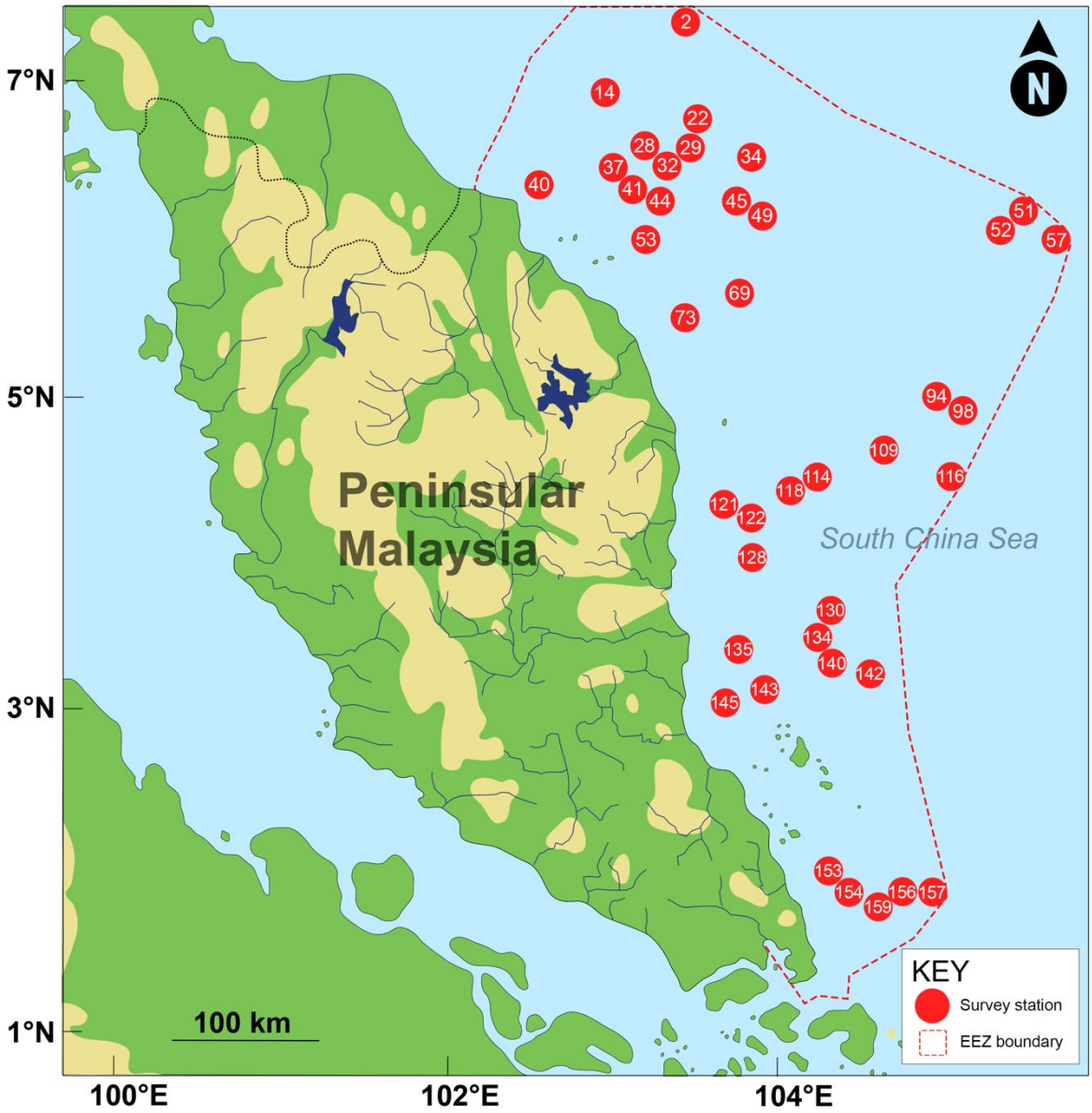

Fig. 1. Map showing the locations of survey stations where fish samples were collected within Peninsular Malaysia's Exclusive Economic Zone (EEZ) in the South China Sea.

Sample processing and morphological identification

Fresh specimens were immediately photographed and tissue sampled for DNA analysis. Fin clips were preserved in 90% ethanol for DNA extraction. Whole voucher specimens were fixed in 10% formalin for one week before long-term storage in 70% ethanol. All specimens were catalogued and deposited at the South China Sea Repository and Reference Centre, Universiti

Malaysia Terengganu. Initial morphological identification of specimens utilised established taxonomic keys (Carpenter and Niem 1999a; 1999b; 2001a; 2001b). An initial species checklist was built based on the morphological identification. Species names were verified based on Eschmeyer's Catalogue of Fishes (Fricke et al. 2023). Ordinal and familial classifications follow van der Laan et al. (2023). Details on the specimens and species identified in this study are provided in table 1. Where possible, at least three specimens per morphospecies were selected for DNA analysis to capture intraspecific morphological variability.

Table 1. Morphologically identified fish species (OTU) from the survey of demersal fishes in the Exclusive Economic Zone (EEZ) of PM. This table lists the common name, sample/museum ID, Barcode of Life Database (BOLD) ID for COI gene sequences, number of specimens examined (n) and IUCN conservation status for each morphologically defined OTU

ORDER, Family, Species	Common Name	Sample/ Museum ID	BOLD ID	*n*	IUCN
TORPEDINIFORMES					
Narcinidae					
Narcine brevilabiata	Shortlip electric ray	EEZ 253	DBEEZ122-23	1	VU
ACANTHURIFORMES					
Siganidae					
Siganus canaliculatus	White-spotted spinefoot	EEZ 384	DBEEZ036-23	1	LC
Siganus fuscescens	Mottled spinefoot	EEZ 382, EEZ 383	DBEEZ034-23, DBEEZ035-23	2	LC
Nemipteridae					
Nemipterus balinensoides[4]	Dwarf threadfin bream	EEZ 119, EEZ 121	DBEEZ108-23, DBEEZ109-23	2	LC
Nemipterus bathybius	Yellowbelly threadfin bream	EEZ 112	DBEEZ107-23	1	LC
Nemipterus nematophorus	Doublewhip threadfin bream	EEZ 032, EEZ 033, EEZ 034	DBEEZ104-23, DBEEZ105-23, DBEEZ106-23	3	LC
Nemipterus nemurus	Redspine threadfin bream	EEZ 411	DBEEZ121-23	1	LC
Nemipterus thosaporni	Palefin threadfin bream	EEZ 030, EEZ 126, EEZ 188, EEZ 189	DBEEZ103-23, DBEEZ110-23, DBEEZ111-23, DBEEZ112-23	4	LC
Nemipterus virgatus	Golden threadfin bream	EEZ 281, EEZ 282	DBEEZ113-23, DBEEZ114-23	2	VU
Scolopsis taenioptera	Lattice monocle bream	EEZ 380, EEZ 381	DBEEZ115-23, DBEEZ116-23	2	LC
Scolopsis vosmeri	Whitecheek monocle bream	EEZ 405, EEZ 406, EEZ 409	DBEEZ118-23, DBEEZ119-23, DBEEZ120-23	3	LC
Pentapodus setosus	Butterfly whiptail	EEZ 397	DBEEZ117-23	1	LC
Lutjanidae					
Lutjanus lutjanus	Bigeye snapper	EEZ 047, EEZ 049, EEZ 259	DBEEZ137-23, DBEEZ139-23, DBEEZ143-23	3	LC
Lutjanus xanthopinnis	Yellowfin snapper	EEZ 048, EEZ 127, EEZ 128	DBEEZ138-23, DBEEZ141-23, DBEEZ142-23	3	DD
Lutjanus vitta	Brownstripe red snapper	EEZ 052	DBEEZ140-23	1	LC
Lethrinidae					
Gymnocranius elongatus	Forktail large-eye bream	EEZ 387	DBEEZ145-23	1	LC
Gymnocranius griseus	Grey large-eye bream	EEZ 414	DBEEZ149-23	1	LC
Lethrinus genivittatus	Longspine emperor	EEZ 388, EEZ 389, EEZ 390	DBEEZ146-23, DBEEZ147-23, DBEEZ148-23	3	LC
Leiognathidae					
Photopectoralis bindus	Orangefin ponyfish	EEZ 169, EEZ 170, EEZ 171	DBEEZ150-23, DBEEZ151-23, DBEEZ152-23	3	NE
Haemulidae					
Diagramma pictum	Painted sweetlips	EEZ 283, EEZ 284, EEZ 323	DBEEZ165-23, DBEEZ166-23, DBEEZ167-23	3	NE
Gerreidae					

Pentaprion longimanus	Longfin mojarra	EEZ 199, EEZ 200, EEZ 201	DBEEZ174-23, DBEEZ175-23, DBEEZ176-23	3	LC
Chaetodontidae					
Coradion chrysozonus	Goldengirdled coralfish	EEZ 291, EEZ 292, EEZ 293	DBEEZ187-23, DBEEZ188-23, DBEEZ189-23	3	LC
ANGUILLIFORMES					
Muraenidae					
Gymnothorax reevesii[4]	Reeve's moray	EEZ 475	DBEEZ123-23	1	LC
Gymnothorax longinquus	Yellow-gilled reef-eel	EEZ 425	DBEEZ217-23	1	LC
AULOPIFORMES					
Synodontidae					
Saurida isarankurai[1]	Shortjaw saury	EEZ 066, EEZ 068, EEZ 069, EEZ 102, EEZ 264	DBEEZ009-23, DBEEZ010-23, DBEEZ011-23, DBEEZ012-23, DBEEZ019-23	5	LC
Saurida longimanus	Longfin lizardfish	EEZ 263	DBEEZ018-23	1	LC
Saurida undosquamis	Brushtooth lizardfish	EEZ 130	DBEEZ013-23	1	LC
Synodus hoshinonis[4]	Blackear lizardfish	EEZ 165, EEZ 167, EEZ 168	DBEEZ014-23, DBEEZ015-23, DBEEZ016-23	3	LC
Trachinocephalus myops	Snakefish	EEZ 345, EEZ 441, EEZ 442	DBEEZ020-23, DBEEZ021-23, DBEEZ022-23	3	LC
BLENNIIFORMES					
Blenniidae					
Xiphasia setifer	Hairtail blenny	EEZ 355, EEZ 416, EEZ 454	DBEEZ203-23, DBEEZ204-23, DBEEZ205-23	3	LC
CARANGIFORMES					
Soleidae					
Liachirus melanospilus	Carpet sole	EEZ 079, EEZ 080, EEZ 299, EEZ 300	DBEEZ026-23, DBEEZ027-23, DBEEZ029-23, DBEEZ030-23	4	LC
Zebrias zebra[3]	Zebra sole	EEZ 275, EEZ 276	DBEEZ028-23, DBEEZ093-23	2	NE
Zebrias quagga	Fringefin zebra sole	EEZ 426, EEZ 427, EEZ 428	DBEEZ031-23, DBEEZ032-23, DBEEZ033-23	3	LC
Samaridae					
Samaris cristatus	Cockatoo righteye flounder	EEZ 471, EEZ 472, EEZ 473	DBEEZ058-23, DBEEZ059-23, DBEEZ060-23	3	LC
Paralichthyidae					
Pseudorhombus javanicus	Javan flounder	EEZ 297	DBEEZ094-23	1	LC
Pseudorhombus dupliciocellatus	Ocellated flounder	EEZ 474	DBEEZ096-23	1	LC
Pseudorhombus pentophthalmus	Fivespot flounder	EEZ 298	DBEEZ095-23	1	LC
Menidae					
Mene maculata	Moonfish	EEZ 016, EEZ 017, EEZ 018	DBEEZ134-23, DBEEZ135-23, DBEEZ136-23	3	NE
Cynoglossidae					
Cynoglossus puncticeps[3]	Speckled tonguesole	EEZ 431, EEZ 432, EEZ 433	DBEEZ178-23, DBEEZ179-23, DBEEZ180-23	3	LC
Cynoglossus kopsii	Tonguesole	EEZ 296	DBEEZ177-23	1	
Citharidae					
Laiopteryx novaezeelandiae	Yellow-dabbled flounder	EEZ 104, EEZ 105, EEZ 106	DBEEZ181-23, DBEEZ182-23, DBEEZ183-23	3	LC
Carangidae					
Seriolina nigrofasciata	Blackbanded trevally	EEZ 228	DBEEZ193-23	1	LC
Turrum coeruleopinnatum	Coastal trevally	EEZ 229, EEZ 238, EEZ 239	DBEEZ194-23, DBEEZ196-23, DBEEZ197-23	3	LC
Uraspis helvola	Whitetongue jack	EEZ 230	DBEEZ195-23	1	LC
CLUPEIFORMES					
Chirocentridae					
Chirocentrus dorab	Dorab wolf-herring	EEZ 021, EEZ 022, EEZ 023	DBEEZ184-23, DBEEZ185-23, DBEEZ186-23	3	NE
GOBIIFORMES					

Gobiidae					
Oxyurichthys auchenolepis[1]	Scaly-nape tentacle goby	EEZ 061, EEZ 062, EEZ 063	DBEEZ168-23, DBEEZ169-23, DBEEZ170-23	3	NE
Yongeichthys nebulosus	Shadow goby	EEZ 455, EEZ 456, EEZ 457	DBEEZ171-23, DBEEZ172-23, DBEEZ173-23	3	LC
HOLOCENTRIFORMES					
Holocentridae					
Sargocentron rubrum	Redcoat	EEZ 356	DBEEZ164-23	1	LC
KURTIFORMES					
Apogonidae					
Rhabdamia sp.	Cardinalfish	EEZ 084, EEZ 085	DBEEZ210-23, DBEEZ211-23	2	
Ostorhinchus nigrocincta	Blackbelt cardinalfish	EEZ 082	DBEEZ209-23	1	NE
Ostorhinchus fasciatus	Broadbanded cardinalfish	EEZ 001, EEZ 154, EEZ 155, EEZ 156	DBEEZ207-23, DBEEZ212-23, DBEEZ213-23, DBEEZ214-23	4	LC
Jaydia truncata	Flagfin cardinalfish	EEZ 013	DBEEZ208-23	1	NE
LOPHIIFORMES					
Ogcocephalidae					
Halieutaea stellata[2]	Smoky seabat	EEZ 247, EEZ 248, EEZ 267	DBEEZ100-23, DBEEZ101-23, DBEEZ102-23	3	LC
Lophiidae					
Lophiomus setigerus[5]	Blackmouth angler	EEZ 278	DBEEZ144-23	1	LC
PERCIFORMES					
Uranoscopidae					
Uranoscopus cognatus	Two-spined yellow-tail stargazer	EEZ 096, EEZ 097, EEZ 231	DBEEZ001-23, DBEEZ002-23, DBEEZ003-23	3	NE
Caesionidae					
Dipterygonotus balteatus	Mottled fusilier	EEZ 257	DBEEZ017-23	1	LC
Synanceiidae					
Choridactylus multibarbus	Orangebanded stingfish	EEZ 334, EEZ 347	DBEEZ023-23, DBEEZ024-23	2	LC
Inimicus cuvieri	Longsnout stinger	EEZ 420	DBEEZ025-23	1	NE
Serranidae					
Epinephelus areolatus	Areolate grouper	EEZ 241, EEZ 242, EEZ 244, EEZ 245	DBEEZ038-23, DBEEZ039-23, DBEEZ040-23, DBEEZ041-23	4	LC
Epinephelus sexfasciatus	Sixbar grouper	EEZ 234, EEZ 367, EEZ 368	DBEEZ037-23, DBEEZ045-23, DBEEZ046-23	3	LC
*Epinephelus heniochus**	Bridled grouper	EEZ 245-A	No barcode, only voucher specimen		
Diploprion bifasciatum	Barred soapfish	EEZ 319, EEZ 320, EEZ 321	DBEEZ042-23, DBEEZ043-23, DBEEZ044-23	3	LC
Cephalopholis boenak	Chocolate hind	EEZ 419	DBEEZ047-23	1	LC
Scorpaenidae					
Brachypterois serrulata	Sawcheek scorpionfish	EEZ 136, EEZ 137, EEZ 138	DBEEZ048-23, DBEEZ049-23, DBEEZ050-23	3	NE
Scorpaenopsis neglecta	Yellowfin scorpionfish	EEZ 333, EEZ 336, EEZ 338	DBEEZ051-23, DBEEZ052-23, DBEEZ054-23	3	LC
Pterois russelli	Plaintail turkeyfish	EEZ 337, EEZ 339	DBEEZ053-23, DBEEZ055-23	2	LC
Neomerinthe procurva	Curvedspine scorpionfish	EEZ 354	DBEEZ057-23	1	NE
Synanceiidae					
Ablabys sp.	Waspfish	EEZ 346	DBEEZ056-23	1	NE
Platycephalidae					
Kumococius rodericensis	Spiny flathead	EEZ 093	DBEEZ064-23	1	LC
Kumococius sp.		EEZ 225, EEZ 226	DBEEZ068-23, DBEEZ069-23	2	
Rogadius pristiger	Thorny flathead	EEZ 422, EEZ 423	DBEEZ074-23, DBEEZ075-23,	2	LC
Rogadius sp.		EEZ 094, EEZ 095, EEZ 424	DBEEZ065-23, DBEEZ066-23, DBEEZ076-23	3	
Thysanophrys chiltonae	Longsnout flathead	EEZ 358, EEZ 359, EEZ 360	DBEEZ070-23, DBEEZ071-23, DBEEZ072-23	3	LC

Insidiator macracanthus	Large-spined flathead	EEZ 421	DBEEZ073-23	1	LC
Elates ransonnettii	Dwarf flathead	EEZ 449, EEZ 450, EEZ 451	DBEEZ077-23, DBEEZ078-23, DBEEZ079-23	3	NE
Pinguipedidae					
Parapercis bicoloripes	Sandperch	EEZ 053, EEZ 054, EEZ 055, EEZ 059, EEZ 060	DBEEZ080-23, DBEEZ081-23, DBEEZ082-23, DBEEZ083-23, DBEEZ084-23	5	NE
Parapercis filamentosa	Threadfin sandperch	EEZ 348, EEZ 436, EEZ 437	DBEEZ087-23, DBEEZ091-23, DBEEZ092-23	3	NE
Parapercis displospilus	Doublespot grubfish	EEZ 349, EEZ 350	DBEEZ088-23, DBEEZ089-23	2	NE
Parapercis xanthozona	Yellowbar sandperch	EEZ 287, EEZ 332, EEZ 415	DBEEZ085-23, DBEEZ086-23, DBEEZ090-23	3	LC
Labridae					
Iniistius evides	Blackspot razorfish	EEZ 042, EEZ 043, EEZ 044, EEZ 285, EEZ 288, EEZ 290	DBEEZ155-23, DBEEZ156-23, DBEEZ157-23, DBEEZ161-23, DBEEZ162-23, DBEEZ163-23	6	LC
Iniistius trivittatus	Blue-razor wrasse	EEZ 036, EEZ 037	DBEEZ153-23, DBEEZ154-23	2	DD
Choerodon typus	Blue-banded wrasse	EEZ 144, EEZ 145, EEZ 146	DBEEZ158-23, DBEEZ159-23, DBEEZ160-23	3	LC
Apistus carinatus	Ocellated waspfish	EEZ 446, EEZ 447	DBEEZ215-23, DBEEZ216-23	2	LC
SCOMBRIFORMES					
Ariommatidae					
Ariomma indicum	Indian driftfish	EEZ 232	DBEEZ206-23	1	NE
SILURIFORMES					
Plotosidae					
Plotosus lineatus	Striped eel catfish	EEZ 177, EEZ 178, EEZ 179	DBEEZ061-23, DBEEZ062-23, DBEEZ063-23	3	NE
SYNGNATHIFORMES					
Mullidae					
Upeneus moluccensis[6]	Goldband goatfish	EEZ 269, EEZ 270, EEZ 271	DBEEZ127-23, DBEEZ128-23, DBEEZ129-23	3	LC
Upeneus sulphureus	Sulphur goatfish	EEZ 224, EEZ 215, EEZ 216, EEZ 217	DBEEZ067-23, DBEEZ124-23, DBEEZ125-23, DBEEZ126-23	4	LC
Upeneus tragula	Freckled goatfish	EEZ 444	DBEEZ130-23	1	LC
Centriscidae					
Centriscus scutatus	Grooved razor-fish	EEZ 304, EEZ 305, EEZ 306	DBEEZ190-23, DBEEZ191-23, DBEEZ192-23	3	LC
Callionymidae					
Dactylopus dactylopus	Fingered dragonet	EEZ 352, EEZ 353	DBEEZ198-23, DBEEZ199-23	2	LC
Callionymus recurvispinnis	Belcher's dragonet	EEZ 460, EEZ 461, EEZ 462	DBEEZ200-23, DBEEZ201-23, DBEEZ202-23	3	NE
TETRAODONTIFORMES					
Tetraodontidae					
Lagocephalus suezensis	Pufferfish	EEZ 210, EEZ 211, EEZ 212	DBEEZ004-23, DBEEZ005-23, DBEEZ006-23	3	LC
Torquigener gloerfelti	Pufferfish	EEZ 403, EEZ 404	DBEEZ007-23, DBEEZ008-23	2	LC
Ostraciidae					
Ostracion nasus	Shortnose boxfish	EEZ 398, EEZ 399, EEZ 402	DBEEZ097-23, DBEEZ098-23, DBEEZ099-23	3	NE
Monacanthidae					
Paramonacanthus pusillus	Faintstripe filefish	EEZ 149, EEZ 150, EEZ 151	DBEEZ131-23, DBEEZ132-23, DBEEZ133-23	3	LC

IUCN - LC: Least Concern; VU: Vulnerable; NE: Not Evaluated; DD: Data Deficient. [1] New record in Malaysia. [2] New record in Peninsular Malaysia (PM). [3] New record in East Coast of PM. [4] New specimen-based record in Malaysia. [5] New specimen-based record in PM. [6] New specimen-based record in East Coast of PM. *Morphospecies without COI sequence (barcode).

DNA analyses

Genomic DNA was isolated from specimens using the standard phenol-chloroform extraction protocol (Sanbrook et al. 2001). DNA purity and concentration were quantified with a microvolume UV spectrophotometer (Quawell Q300, Quawell, CA) and stored at -20°C until further use. A ~650 bp fragment of the mitochondrial COI gene was PCR amplified using universal teleost primers by Ward et al. (2005):
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. Thermal cycling conditions were: initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C for 30 sec, 48°C for 50 sec, and 72°C for 1 min, ending with a final extension at 72°C for 10 min. Negative controls lacking template were included. Amplified PCR products were visualized by 2% agarose gel electrophoresis. Successful amplicons were purified and Sanger sequenced bidirectionally by a commercial provider (Apical Scientific Sdn. Bhd.) using the ABI PRISM 3730XL automated sequencer and the ABI PRISM BigDye terminator cycle sequencing kit v3.1 (Applied Biosystems, Foster City, CA).

Phylogenetic reconstruction and automatic species delimitation

The chromatogram traces from each sequenced sample were visually inspected before alignment in Geneious Prime v2030.1.2 (Biomatters Ltd., Auckland, NZ). The forward and reverse sequences were proofread, aligned, and examined for any deletions, insertions, or stop codons using the same software.
A total of 215 cytochrome c oxidase subunit I (COI) barcode sequences were generated in this study. Complete data are accessible under the project 'DBEEZ: DNA Barcoding – EEZ Offshore Demersal Survey' in the Barcode of Life Database (BOLD) database (Ratnasingham and Hebert 2007). To evaluate taxon discrimination, pairwise genetic distances were calculated within and between species, genera, and families utilizing the Kimura 2-parameter (K2P) substitution model (Kimura 1980) in the BOLD analysis tools. The barcode gap analysis was computed for all sequences excluding singleton species, within the same BOLD analysis platform. Phylogenetic relationships were inferred using Bayesian Inference (BI) in BEAST 2 (Bouckaert et al. 2014) and

Maximum Likelihood (ML) in raxmlGUI (Edler et al. 2021). IQ-TREE v2.2.0 (Kalyaanamoorthy et al. 2017) identified the General Time Reversible with empirical base frequencies, invariant sites and gamma distribution with 4 categories (GTR+F+I+G4) as the optimal evolutionary model for our dataset, as implemented in the CIPRES portal (Miller et al. 2011). The BI analysis employed a relaxed molecular clock and birth-death tree prior with empirical base frequencies and four gamma categories. Two independent Markov Chain Monte Carlo (MCMC) runs of 40 million generations were sampled every 1000 generations, discarding the first 20% as burn-in. Convergence was assessed in Tracer v1.7.2 (Rambaut et al. 2018) (ESS > 200) before combining runs in LogCombiner, as integrated in the BEAST 2 package. The final BI tree was constructed in TreeAnnotator (Rambaut and Drummond 2013). The ML analysis was performed using 1000 rapid bootstrap replicates under the GTR+I+G model. Resulting phylogenies were visualized and edited in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Five different sequence-based methods were used to delimit the Molecular Operational Taxonomic Units (MOTUs) (=putative species) from our dataset - (1) Refined Single Linkage (RESL), (2) Automatic Barcode Gap Discovery (ABGD), (3) Assemble Species by Automatic Partitioning (ASAP), (4) Generalized Mixed Yule Coalescent (GMYC), and (5) Bayesian Poisson Tree Process (bPTP). Employing five different automatic molecular species delimitation methods, each with distinct underlying assumptions, enabled the exploration of the reliability of MOTU partitioning in this study (Luo et al. 2018).

The first analysis was done within the BOLD platform using the RESL algorithm (Ratnasingham and Hebert 2013) to assign sequences to a dedicated Barcode Index Numbers (BIN). Next, the ABGD (Puillandre et al. 2012) 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.

The ASAP (Puillandre et al. 2021) analysis was performed on https://bioinfo.mnhn.fr/abi/public/asap/asapweb.html using the Kimura 2-parameter (K2P) substitution model. ASAP calculates a score for different MOTU partitioning scenarios above a threshold of 3% genetic distance. Lower ASAP scores indicate more optimal MOTU partitioning. Finally, both GMYC (Pons et al. 2006) and bPTP (Zhang et al. 2013) methods were employed with the fully resolved, BI ultrametric tree using only unique haplotypes (see above for the reconstruction method). The haplotype dataset was built in collapsing all 215 individual COI sequences into 181 unique haplotype sequences using FaBox v1.61 (Villesen 2007). A single-threshold GMYC analysis was run in RStudio (Allaire 2012) with the ‘splits’ package (Fujisawa

and Barraclough 2013). The bPTP analysis was performed on the GMYC web server (https://speciesh-its.org/gmyc/). The final species delimitation scheme was determined based on the consensus of the five molecular delimitation methods. Species initially identified by morphological characteristics are referred to as morphospecies (operational taxonomic unit = OTU), while putative species delimited by DNA barcodes are referred to as molecular operational taxonomic units (MOTUs).

To further investigate potential cryptic diversity in three OTUs showing deep mitochondrial divergences - *Rogadius pristiger*, *Kumococius rodericensis*, and *Upeneus sulphureus* - we compiled an expanded barcode sequences (COI) dataset incorporating publicly available verified sequences from the Barcode of Life Data System (BOLD). Additional sequences were sourced across the Indo-West Pacific, including Indonesia, Vietnam, Bangladesh, South China, Australia, and the United Arab Emirates (Fig. 5). Figure 5 provides the BOLD IDs and GenBank accession numbers for all utilised sequences. Phylogenetic reconstruction on this dataset was performed using maximum likelihood (ML) analysis with the methods described previously. Incorporating publicly available barcodes provides geographical context and allows comparison to our OTU lineages to delineate species boundaries and evolutionary relationships.

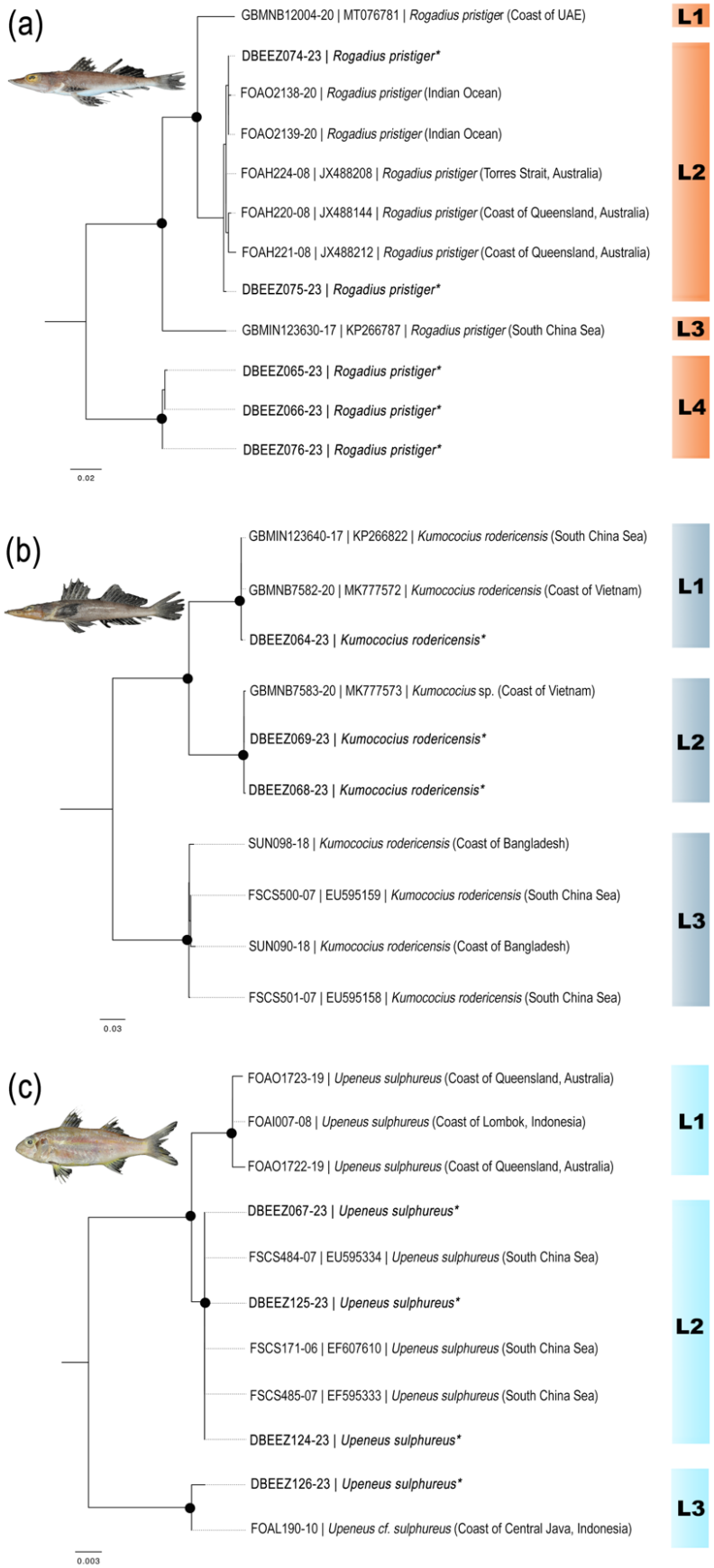

Fig. 5. Maximum likelihood phylogenetic tree based on expanded *COI* sequences retrieved from the Barcode of Life Data System (BOLD) for (a) *Rogadius pristiger*, (b) *Kumococius rodericensis*, and (c) *Upeneus sulphureus*. Sequences marked with * are from this study. Colored bars delineate distinct phylogenetic lineages (L1-L4). In the final dataset, all

sequences within L4 in (a) and L2 (b) are designated as *Rogadius* sp. and *Kumococius* sp., respectively. BOLD identification numbers and GenBank accession numbers are provided for each included sequence. Black circles indicate node support greater than 85%.

RESULTS

Fish diversity

This study obtained partial (~650 bp) mitochondrial cytochrome c oxidase subunit I (COI) sequences from 217 out of 475 specimens collected during demersal trawl surveys across the East Coast PM Exclusive Economic Zone (EEZ). The successfully sequenced specimens represented 92 morphological species belonging to 68 genera, 41 families, 16 orders, and two classes – Elasmobranchii and Actinopterygii (Table 1). One morphospecies, the bridled grouper (*Epinephelus heniochus*), failed to be barcoded but was retained in the final species checklist (Table 1). In the initial morphological identification, all species were identified to species level based on morphological taxonomy, except *Rhabdamia* sp. and *Ablabys* sp.which could only be identified to the genus level.

The order Perciformes displayed the highest species richness with 26 species, accounting for 28.2% of the total species identified. Acanthuriformes and Carangiformes were the next most speciose orders with 21 species (22.8%) and 14 species (15.2%), respectively (Fig. 2; Table 1). Within Perciformes, Platycephalidae and Scorpaenidae exhibited the greatest diversity with five species each, followed by Pinguipedidae and Serranidae with four species per family. The family Nemipteridae of the order Acanthuriformes contained the highest number of species at nine. Based on International Union for Conservation of Nature (IUCN) Red List assessments, two species in our dataset were classified as vulnerable - the shortlip electric ray (*Narcine brevilabiata*) and the golden threadfin bream (*Nemipterus virgatus*). The remaining species were categorized as least concern, data deficient, or not evaluated.

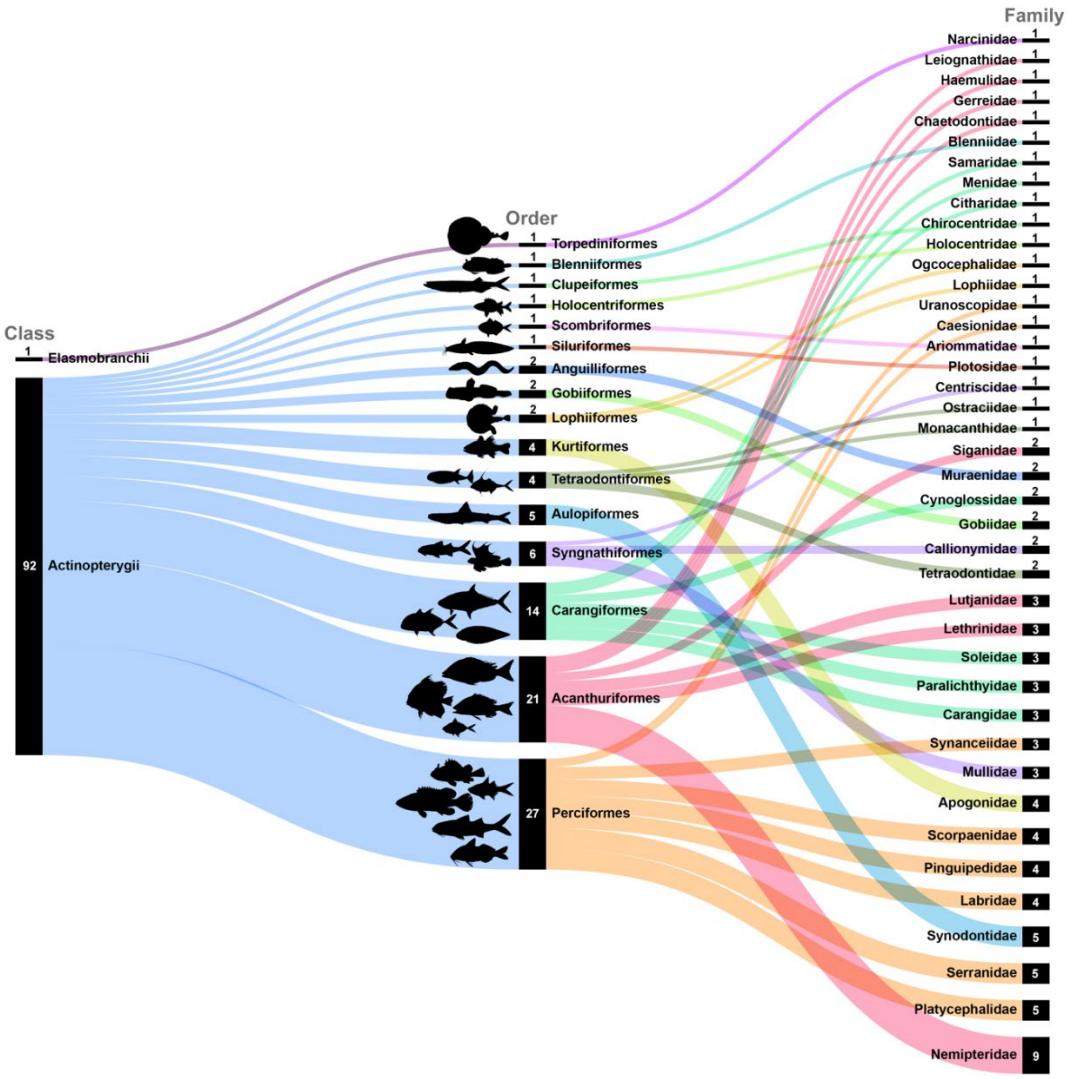

Fig. 2. Number of fish species recorded in this study, ranked by class, order, and family. The black columns represent the taxonomic richness within each group.

To cross-validate our findings and identify potential new records, we analysed several published species checklists from the waters of the east coast of PM (Chong et al. 2010; Du et al. 2019; Matsunuma et al. 2011; Motomura et al.; Seah et al. 2020; 2021). Of the 92 species identified, two, namely *Saurida isarankurai* Shindo & Yamada, 1972, and *Oxyurichthys auchenolepis* Bleeker, 1876, were recorded for the first time in Malaysia. Both species are uncommon in Malaysian waters. *Saurida isarankurai*, a benthic species, is typically found in the Northwest Pacific and Southwestern Pacific regions (Froese and Pauly 2023). Previous reports indicate its occurrence in neighbouring areas such as China, India, Thailand, Vietnam, Indonesia, the Philippines and Papua New Guinea (GBIF.org 2023). Our findings extend the known distribution range of *S. isarankurai* to the southern South China Sea on PM. On the other hand, *O. auchenolepis*, a member of the family Gobiidae, is distributed in the western Central Pacific from China to northern Australia (Froese and Pauly 2023). Interestingly, despite our thorough literature survey, *O. auchenolepis* has not yet been documented in Malaysian waters. We also identified a batfish *Halieutaea stellata*

(Vahl, 1797) as a new PM record, with previous documentation only from Sarawak, East Malaysia (GBIF.org 2023). Additionally, we provide the first Malaysian specimen-based records for three species: *Nemipterus balinensoides*, *Gymnothorax reevesii* and *Synodus hoshinonis*.

DNA-based delimitation

The 217 generated *COI* barcode sequences had lengths exceeding 650 base pairs, with no insertions, deletions, or stop codons detected. Nucleotide composition analysis revealed mean percentages of 18.63% guanine (G), 22.43% cytosine (C), 23.59% adenine (A), and 29.72% thymine (T). Over half of the species (67%, 62 species) were represented by multiple specimens, while 30 were singletons (Table 1). The mean number of specimens per species was 2.35. As expected, genetic divergence values based on the Kimura 2-parameter (K2P) model increased with taxonomic level: average within-species divergence was 0.77% (SE = 0.01), within-genus divergence was 17.14% (SE = 0.02), and within-family divergence was 23.09% (SE = 0.01) (Table 2).

Table 2. K2P divergence values from 217 analysed specimens with increasing taxonomic levels

Category	n	Taxa	Comparisons	Minimum (%)	Mean (%)	Maximum (%)	SE (%)
Within species	187	62	208	0	0.77	13.2	0.01
Within genus	94	15	233	6.70	17.14	26.57	0.02
Within family	121	12	423	14.01	23.09	27.86	0.01

SE: standard error.

Intraspecific genetic distances based on the K2P model exceeded the standard 2% threshold (Hebert et al. 2003b; Hubert et al. 2008) in three species: *Rogadius pristiger* (13.20%), *Kumococius rodericensis* (12.79%), and *Upeneus sulphureus* (2.67%) (Table 3). Barcode gap analysis showed that all species represented by multiple sequences had non-overlapping intra- and interspecific divergence, supporting species delimitation. The majority of sequences fell into quartile 2 (Q2) of the distribution, where intraspecific divergence was below 2% and interspecific divergence was above 2% (Fig. 3). The three aforementioned OTUs (*i.e.*, *Rogadius pristiger*, *Kumococius rodericensis* and *Upeneus sulphureus*) fell into quartile 3 (Q3), meaning that both intraspecific and interspecific divergence exceeded 2%.

Table 3. List of morphological species comprising two MOTUs (=BINs). The summary statistics include the BIN of each MOTU, their maximum intraspecific distance and distance to the nearest neighbour (*i.e.*, minimum interspecific distance)

OTU/MOTUs	Max. intraspecific distance (%)	Nearest neighbour distance (%)
Species comprising two MOTUs		
Rogadius pristiger	13.2	17.84
BOLD:AFH2200	0.33	11.44
BOLD:AAD3175	0.49	11.44
Kumococius rodericensis	12.79	18.54
BOLD:ADL2558	0	11.44
BOLD:AED3908	0.16	11.44
Upeneus sulphureus	2.67	11.98
BOLD:AAB6466	0	2.61
BOLD:AAM2094	0	2.61

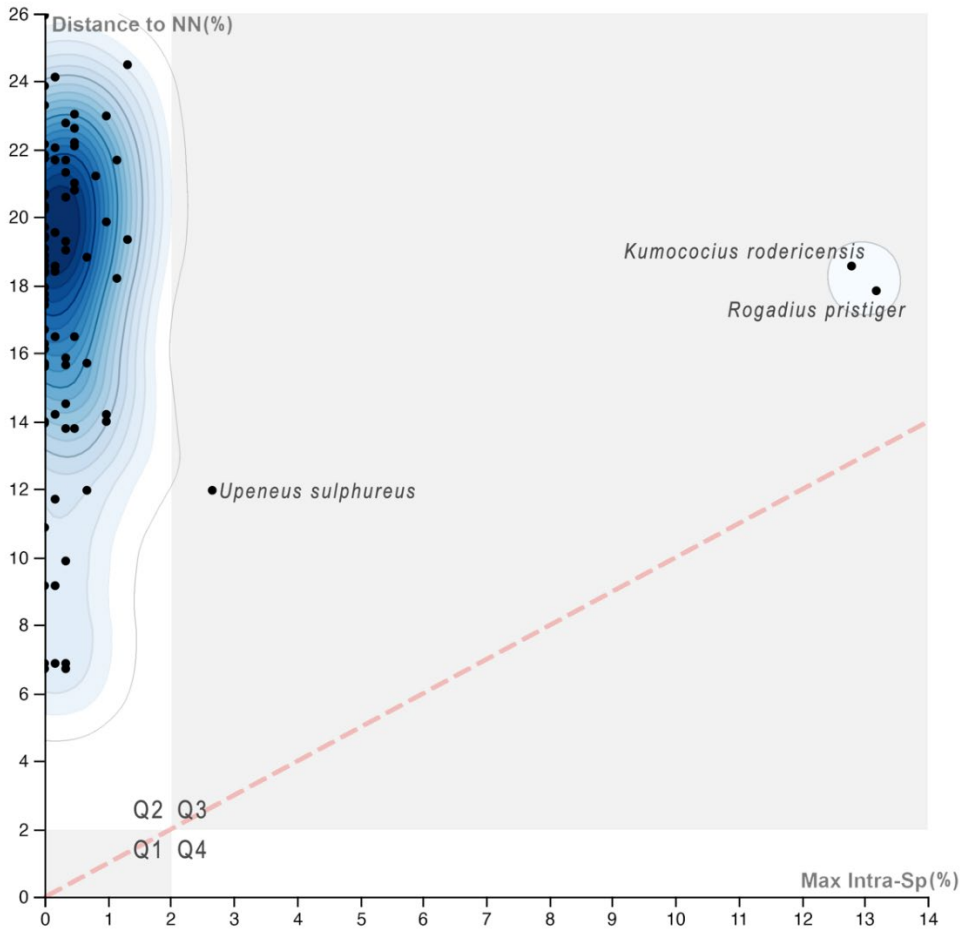

Fig. 3. The scatter plot shows the overlap of the maximum intraspecific distances compared to the interspecific distances (nearest neighbour, NN) of all species excluding singleton species. The contour plot displays the estimated density of each cluster of plots. A threshold of 2% divergence was set as the heuristic threshold for species delimitation, indicated by the division of quartile one and two (Q1 and Q2). Plots above the red line suggest the presence of a barcoding gap.

The Bayesian Inference (BI) and Maximum Likelihood (ML) phylogenies (Fig. 4) were fully resolved with minimal topology differences. Overall, nodal support were higher in the ML tree, and thus utilised to visualise the molecular operational taxonomic unit (MOTU) delimitation outcomes. The five MOTU delimitation methods (RESL, ABGD, ASAP, GMYC, and bPTP) were generally consistent in MOTU counts, though all exceeded the initial 92 morphology-based species estimate. The RESL analysis clustered sequences into 95 MOTUs (Barcode Index Numbers or BINs), independently of taxonomic assignment. Comparing input taxonomy against RESL-BIN designations validates concordance between barcode clusters and species labels. Here, we performed this validation by contrasting specimen taxonomy with other records in their respective BINs, including from external BOLD database submissions. ABGD analysis identified 94 MOTUs across initial partitions for all substitution models at prior intraspecific divergence values of 0.0010-0.0599. ASAP, GMYC, and bPTP analyses congruently recovered 95 MOTUs. Figure 4 highlights incongruences between MOTUs and morphological taxonomy (red bars), detailed in Table 3.

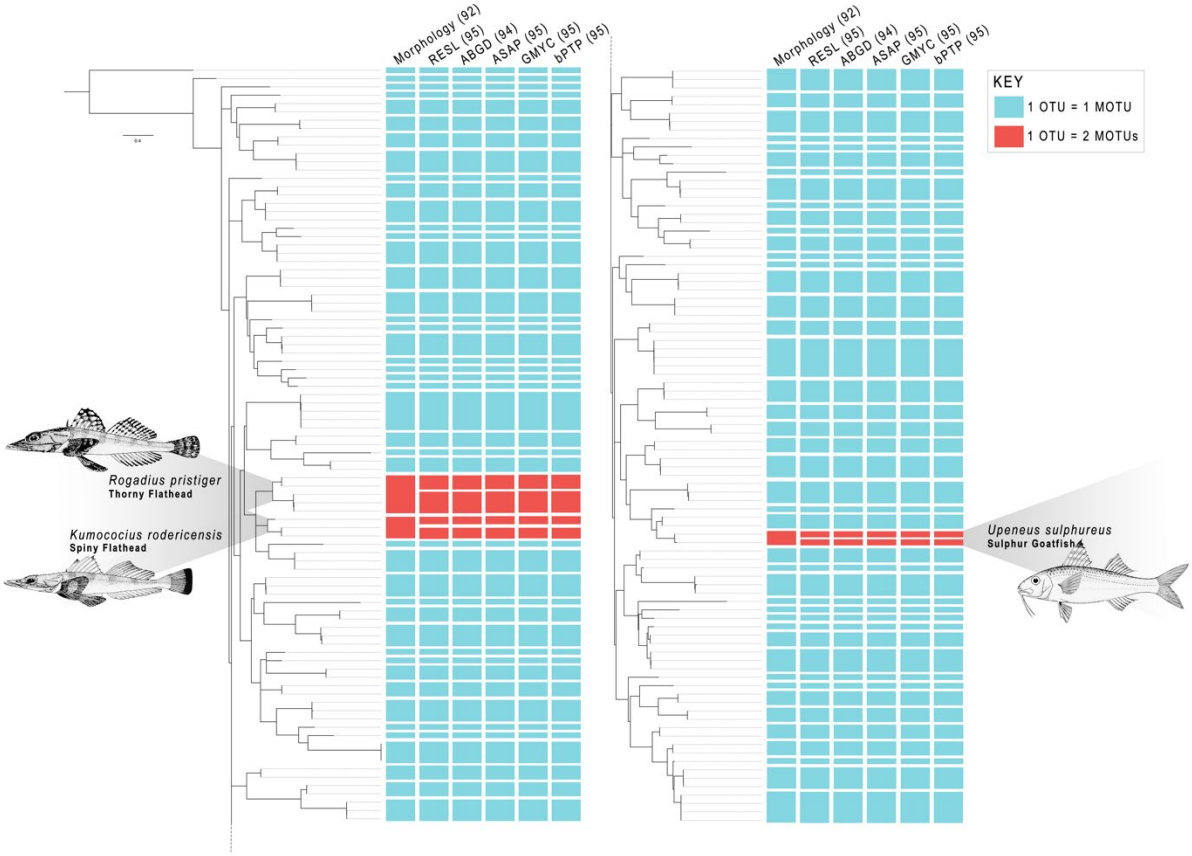

Fig. 4. Maximum likelihood phylogenetic tree based on the 217 DNA barcodes with delineated MOTUs. Coloured bars show (from left to right): morphological species, MOTUs delimited by RESL, ABGD, ASAP, GMYC and bPTP species delimitation schemes. Red bars indicate discrepancies between the different analyses (morphological-genetic discrepancies). The fish illustrations were retrieved from FAO guidebook (Bruin et al. 1995).

DISCUSSION

MOTU delimitation and barcode gap

Previous DNA barcoding studies (*e.g.*, Ramirez et al. 2017; Sholihah et al. 2020; Zainal Abidin et al. 2021) have emphasised the importance of employing MOTU delimitation approaches, as each is based on different assumptions and may yield varying MOTU estimates for a given dataset. The rationale for a multi-method approach is that comparing outcomes across algorithms with distinct underlying models provides greater confidence in delimitation and reveals cryptic diversity (when present) not apparent from morphology alone. In this study, we contrasted results from five automated delimitation methods - RESL, ABGD, ASAP, GMYC, and bPTP against our morphology-based species taxonomy. Despite differing computational algorithms, all five approaches yielded congruent results, delimiting 94-95 MOTUs compared to the 92 initial morphospecies. The higher MOTU estimates indicate the likely presence of at least two cryptic species in our dataset based on the automated delimitation analyses. This highlights the ability of DNA barcoding to reveal potential taxonomic underestimation relative to morphological demarcation alone. By integrating multiple delimitation approaches, our study provides a robust, conservative framework for delineating putative species clusters corroborated across distinct algorithms (Collins and Cruickshank 2013). Such integrated delimitation facilitates the discovery of cryptic diversity in understudied tropical marine faunas that may be underestimated by morphological taxonomy.

A key premise of DNA barcoding is detecting a 'barcode gap' between maximum intraspecific and minimum interspecific genetic distances (nearest neighbor distances - NN) (Hubert and Hanner 2015). Substantial barcode gap within a morphological species provide evidence of cryptic diversity (April et al. 2011). However, barcode overlap between two morphospecies indicates either intraspecific forms or shared ancestral polymorphisms and/or hybridization with introgression (Hubert and Hanner 2015; Zainal Abidin et al. 2021). In such cases, expanding to a multigene approach can help resolve taxonomic status, as mitochondrial introgression may obscure species boundaries (Dupuis et al. 2012). The inclusion of nuclear markers to complement mitochondrial DNA barcodes allows robust reciprocal monophyly testing to confirm or reject potential hybridization (Miralles and Vences 2013).

Our barcode gap analysis revealed that the majority of sequences were positioned in quartile 2 (Q2) of the distribution (Fig. 3), exhibiting less than 2% intraspecific but greater than 2% interspecific divergences. The non-overlapping intra- and interspecific genetic distances observed

for most taxa suggest these species have achieved reciprocal monophyly, with DNA barcode clusters concordant with current morphological taxonomy (Hubert and Hanner 2015). The presence of a barcode gap provides support for clear species delimitation, as coalescence and divergence between species exceeds variation within species in these cases. However, three OTUs - *Rogadius pristiger*, *Kumococius rodericensis*, and *Upeneus sulphureus* - were positioned in quartile 3 (Q3), indicating both intra- and interspecific distances exceeded 2%. Such deep intraspecific divergences may reflect potential cryptic diversity or shared ancestral polymorphisms between closely related species, as reported in earlier barcoding studies (Hou et al. 2018; Hubert et al. 2012; Mat Jaafar et al. 2012). These taxa likely encompass multiple distinct lineages that were undetected by morphological diagnoses (Hubert and Hanner 2015). Additional integrative taxonomic analysis incorporating morphologic, molecular, and biogeographic evidence is warranted to fully delimit species boundaries and clarify taxonomic status for these three OTUs exhibiting mitochondrial barcode overlaps.

The mean intraspecific K2P divergence (0.77%) was 20-fold lower than the mean congeneric divergence (17.14%), reflecting the expected increase in genetic divergence at higher taxonomic levels. However, both estimates exceed typical values reported for marine fishes in previous barcoding studies. Most assessments show intraspecific divergences of 0.25-0.39% and congeneric divergences of 4.56-9.93% (Lakra et al. 2011; Landi et al. 2014; Mecklenburg et al. 2011; Ward et al. 2005; Xu et al. 2021). Notwithstanding, higher average values have been documented among Indo-Pacific coral reef fishes (intraspecific = 1.06%; congeneric = 15.34%) (Hubert et al. 2012), native ray-finned fishes in Taiwan (intraspecific = 1.51%; congeneric = 15.24%) (Chang et al. 2017), and Malaysian mangrove fishes (intraspecific = 0.85%; congeneric = 16.70%) (Zainal Abidin et al. 2021). Elevated genetic divergence observed likely stems from taxonomic underestimation rather than exceptional within-species variation for these understudied tropical fish faunas (Zemlak et al. 2009). Tropical ecosystems harbor extensive cryptic diversity (Hubert et al. 2017) – genetically distinct lineages that are morphologically indistinguishable, representing overlooked species (Bickford et al. 2007). Incorrect grouping of such cryptic species into a single taxon results in deep intraspecific divergences and inflated genetic variance. Molecular taxonomy consistently uncovers higher species numbers compared to morphology-based approaches in hyperdiverse tropical regions, highlighting limitations of traditional taxonomy (Zainal Abidin et al. 2022; Ghazali et al. 2023).

Deep intraspecific divergence and cryptic diversity

DNA barcode clusters frequently conflict with morphology-based taxonomy in tropical ecosystems, reflecting higher rates of polyphyly and paraphyly compared to temperate regions (Hubert and Hanner 2015). Our dataset revealed three OTUs with elevated mitochondrial divergences (>2%) that may represent overlooked cryptic diversity - *Rogadius pristiger*, *Kumococius rodericensis*, and *Upeneus sulphureus*. To further investigate these taxa, we compiled an expanded cytochrome *c* oxidase subunit I (*COI*) dataset incorporating publicly available sequences from the Barcode of Life Data System (BOLD). Comparing our lineages against those from across the Indo-Pacific provides biogeographical context and helps delineate species boundaries, revealing cryptic diversity and evolutionary relationships.

The first two cases involved two flathead OTUs - *Rogadius pristiger* and *Kumococius rodericensis*. Our dataset of five *R. pristiger* specimens formed four distinct COI lineages when analysed with public sequences (Fig. 5a). Two barcodes clustered with sequences from the Indian Ocean, Torres Strait, and Queensland, Australia. While remaining three of our barcodes formed a divergent clade (L4, Fig. 5a) representing a novel Barcode Index Number (BIN) in BOLD, evidencing probable cryptic species. Integration of global databases revealed geographical variants and relationships within this morphospecies. No significant variation was recorded within lineages, yet inter-lineage divergence reached 14.1% between lineages 3 (L3) and 4 (L4) (Supplementary Table 1; Fig. 5a), corroborating taxonomic underestimation. Notably, our East Coast PM specimens in lineage 2 (L2) showed 8.0% divergence (Supplementary Table 1) from a South China Sea individual, despite overlapping distributions. After thorough consideration, all sequences within lineage 4 (L4) are designated as *Rogadius* sp. (Table 1).

In the second case, the expanded dataset of ten *K. rodericensis* specimens, including three of our barcodes, formed three distinct COI lineages when analyzed with BOLD sequences (Fig. 5b). One of our barcodes was clustered with specimens from the South China Sea and the Vietnamese coast, while the other two interestingly formed a common lineage with an unidentified spiny flathead sample (i.e. *Kumococius* sp.) from Vietnam. The third lineage (L3; $n = 4$) included sequences from the coast of Bangladesh and the South China Sea. Despite minimal within-lineage variation (0 - 1.0%), substantial inter-lineage divergences up to 23.5% were observed (Supplementary Table 1), indicating cryptic diversity that is morphologically undetectable within this dataset. Such intraspecific polyphyly despite similar morphology likely reflects taxonomic underestimation of species number in hyperdiverse tropical ecosystems (Hubert et al., 2012). Unlike previous barcoding studies on tropical marine fishes (*e.g.*, DiBattista et al. 2016; Steinke et al. 2009), our flatheads lineages do not exhibit allopatric distributions. The overlapping distributions of our lineages instead reject the hypothesis that they simply represent geographically isolated populations with independent evolutionary trajectories. Sympatric cryptic taxa likely reflect

incomplete taxonomy and limitations of traditional morphological diagnoses rather than just biogeographic variance (Steinke et al. 2009). Consequently, after careful consideration, all sequences within lineage 2 (L2) are designated as *Kumococius* sp. (Table 1).

The commercially exploited platycephalid flatheads (Perciformes: Platycephalidae) of the Indo-Pacific exemplify the challenges of coping with long-standing fishing pressure in tropical ecosystem due to obscured cryptic lineages (Imamura and Knapp 2009; Puckridge et al. 2013). The substantial mitochondrial divergences within the flathead OTUs likely signify genuine overlooked diversity rather than anomalous variation, congruent to cryptic diversity patterns reported in other studies. For instance, DNA barcoding uncovered approximately 10% divergence between western and eastern Australian forms of the morphospecies *Platycephalus marmoratus* (Puckridge 2006), providing robust evidence of an unrecognized species that was subsequently described as *P. orbitalis* (Imamura & Knapp, 2009). Another comprehensive assessment of *P. indicus* across the Indo-West Pacific revealed eight remarkably divergent lineages separated by up to 16.37% genetic distance, unveiling extensive cryptic diversity (Puckridge et al., 2013). The large intraspecific divergences correspond to geographically isolated lineages, reflecting recognised biogeographic barriers in the studied region (Puckridge et al., 2013). Moreover, multiple taxonomic work on platycephalids has focused on Australian waters, delineating new flathead species in the region (Imamura 2007; Imamura and Gomon 2010; Knapp and Imamura 2004). This further supports the likelihood that the distinct genetic lineages uncovered in our study may represent undescribed cryptic species within the currently recognized taxa.

Expanded analysis of the *Upeneus sulphureus* OTU included 11 *COI* sequences, with four from our dataset (Fig. 5c). Three phylogenetic lineages were observed, with three of our barcodes clustering with South China Sea samples in lineage 2 (L2), reflecting shared genetic pool on a regional scale. Intriguingly, this lineage only diverged by 0.7% from the Australian and Indonesian sequences in lineage 1 (L1) (Supplementary Table 1, Fig. 5c), despite broad geographical separation. However, one barcode formed a divergent lineage, clustering with an uncertain *Upeneus* cf. *sulphureus* specimen from Indonesia at 3.0% distance from other lineages. Similar slightly elevated within species divergence (2.02%) was also documented in *U. sulphureus* from China (Zhang et al., 2011). Although intraspecific divergence was moderate in this OTU (2.67%, Table 3), comparison against publicly available data provided a biogeographical context to infer probable cryptic diversity within our specimens.

OTU – MOTU discordance

Several factors contribute to the discordances between morphospecies (OTUs) and molecular operational taxonomic units (MOTUs) observed in this study. The taxonomic impediment disproportionately impacts the megadiverse tropical ecosystems, where extreme diversity coupled with limited taxonomists results in incomplete alpha taxonomy (Bini et al. 2006). Furthermore, the slower turnover and older age of tropical marine species increase the manifestation of ancestral polymorphisms, leading to deep mitochondrial divergence despite taxonomic integrity (Hubert and Hanner 2015; Rabosky et al. 2018). As seen here, such polyphyly has been attributed to cryptic speciation in fishes, specifically within the tropical South China Sea, validating barcoding method in identifying known species and highlighting potential new ones (Landi et al. 2014; Lara et al. 2010; Ward et al. 2008). Our integrated taxonomic framework, combining morphology, DNA barcoding, and phylogeography, elucidates evolutionary lineages and biogeography to uncover overlooked diversity. Incorporating verified data from public sources such as the Barcode of Life Data System (BOLD) and GenBank facilitates insight into tropical biodiversity across spatial scales (Sholihah et al. 2020; Zainal Abidin et al. 2022). Although DNA barcoding is not a substitute for morphological taxonomy, it is a powerful supplementary tool for identifying species and guiding future taxonomic research. Our study lays the groundwork for such investigations into the drivers of cryptic diversity patterns in Malaysian demersal fishes.

Towards the establishment of a comprehensive DNA barcoding library of the fish community in the EEZ of Peninsular Malaysia

Accurate identification of organisms is essential for assessing ecosystem status which is now well acknowledged to require the integration of morphological and molecular techniques (Bourlat et al. 2013). DNA barcoding enables rapid biodiversity assessment, although contingent on availability of comprehensive libraries for comparative analysis (Ward et al. 2005). Hebert et al. (2003a) highlighted the efficacy of the mitochondrial COI gene for delineating species boundaries, suggesting its discriminatory power could enable species identification from sequences alone. Our study highlights the ability of barcoding to reveal potential cryptic diversity, although a formal downstream taxonomic analysis is still required. This study provides the first comprehensive DNA barcoding assessment of demersal fishes in the waters of Peninsular Malaysia's Exclusive Economic Zone (EEZ) in the South China Sea. Characterising biodiversity in this vast but under-researched marine region is critical for national monitoring and management of fisheries resources, as over 50% of the species examined are commercially exploited (DoF 2000; 2018). Under the United Nations Convention on the Law of the Sea (UNCLOS), EEZs which extends 200 nautical miles define the marine areas and sovereign rights of the coastal states over the resources (Poling

2013). Safeguarding EEZ biodiversity and ecosystems is therefore crucial for Malaysia's national food security, economy, and heritage.

DNA barcoding relies on constructing comprehensive reference libraries to enable sequence-based species identification (Ratnasingham and Hebert 2007). While public databases like BOLD and GenBank accelerate insights, localised curated repositories provide more practical foundations tailored to regional biota (Bemis et al. 2023; Zainal Abidin et al. 2021). For instance, our Malaysia-focused library better elucidates biogeographic patterns in these demersal fishes. Beyond species discovery, such resources have diverse applications from food authentication to conservation (Chin et al. 2016; Zainal Abidin et al. 2022). Ongoing barcoding efforts should engage taxonomists to integrate multiple lines of evidence for robust species delimitation, especially in understudied tropical ecosystems. Our study establishes an integrative taxonomic framework combining morphology, DNA barcoding, and phylogenetics to elucidate cryptic diversity, thus facilitating species discovery, and provide insights into evolutionary lineages among Malaysian demersal fishes. The 92 morphospecies and distinct mitochondrial lineages uncovered highlight underestimated diversity and represent candidate species for description. The curated DNA barcode library provides a foundation for conservation and sustainable use of these commercially valuable fish stocks (Knebelsberger et al. 2014). As climate change, overfishing, and other stressors rapidly impact tropical marine biodiversity, continuous assessments are imperative (Pecl et al. 2014). Comprehensive barcoding surveys like this research are crucial for monitoring, managing, and protecting Malaysia's invaluable marine living resources.

CONCLUSIONS

This study uses an integrative taxonomic approach that combines DNA barcoding and morphological identification to elucidate fish diversity in the Exclusive Economic Zone (EEZ) of Peninsular Malaysia's east coast. We reassessed 475 demersal fishes comprising 92 putative species and 16 orders, including two vulnerable IUCN species. The DNA barcoding cytochrome c oxidase subunit I (COI) gene revealed 95 consensus Molecular Operational Taxonomic Units (MOTUs) for all automated delimitation methods. Interestingly, several MOTUs within a morphospecies had over 2% intra- and interspecific genetic divergence, indicating either deep intraspecific variation or cryptic species. These results highlight the complexity of species delimitation and the value of genetic methods. Our study provides important insights into east coast fish diversity, improves understanding of genetic distribution and conservation needs, and creates a comprehensive framework combining barcoding and morphology to inform future research and management

strategies for Malaysia's marine biodiversity. The expanded genetic barcode database will facilitate ongoing and future molecular taxonomic studies of Malaysian ichthyofauna. Overall, this first large-scale analysis of east coast PM demersal fishes demonstrates the ability of barcoding to illuminate diversity and reveal hidden divergences indicative of reproductive isolation and cryptic speciation. Our contribution to public databases enables species identification by experts and laypeople alike, and offers a wide range of potential applications. Further taxonomic research is warranted, as evident by the knowledge gaps highlighted in this study.

Acknowledgments: We thank the Department of Fisheries Malaysia for the opportunity to participate in the 2016 National Demersal Trawl Survey and for facilitating the sample collection. We also thank the SEAFDEC Training Department in Bangkok for their assistance with samples collection during the survey. Finally, we wish to extend our appreciation to Mrs. Salwani Abdullah and Mr. Muhammad Wan Hanafi, whose assistance in sample collection and processing made this work possible. Danial Hariz Z.A. was supported by a postdoctoral fellowship (JPNP.AUPE002) from Universiti Sains Malaysia (August 2023 – March 2024).

Availability of data and materials: Voucher specimens are available as described in the text. All the COI sequences determined in this study have been uploaded in BOLD under the public project – DBEEZ: DNA Barcoding EEZ Offshore Demersal Survey and deposited in GenBank (Accession nos. OR918571–OR918786; Table S2).

Authors' contributions: DHZA designed experiments, conducted analyses, data visualisation, and wrote the manuscript. SAMN and TNAMJ provided funding, participated in data collection and analysis. TNAMJ, JAFJ, MSA, YGS, MAR, NSZ, MPT, and KMZ collected specimens, performed morphological identification and laboratory analysis. All authors contributed in the review of the manuscript.

Competing interest: The authors declare no competing interest.

Consent for publication: Not applicable.

Ethics approval consent to participate: This project adhered to relevant national and international guidelines and did not involve any endangered or protected fish species. All fish specimens were collected during the East Coast PM(ECPM) Exclusive Economic Zone (EEZ) demersal trawl survey, for which the authors obtained the necessary permits and approvals.

REFERENCES

Allaire J. 2012. RStudio: Integrated development environment for R.**770:**165–171. Available at: http://www.rstudio.com/.

Allen GR. 2008. Conservation hotspots of biodiversity and endemism for Indo-pacific coral reef fishes. Aquatic Conservation: Marine and Freshwater Ecosystems **18:**541–556. doi:10.1002/aqc.880.

Allen GR, Adrim M. 2003. Coral reef fishes of Indonesia. Zool Stud **42:**1–72.

Allen GR, Amaoka K, Anderson Jr WD, Bellwood DR, Bohlke EB, Bradbury MG, Carpenter KE, Caruso JH, Cohen AC, Cohen DM. 2000. A checklist of the fishes of the South China Sea. Raffles B Zool **8:**569–667.

April J, Mayden RL, Hanner RH, Bernatchez L. 2011. Genetic calibration of species diversity among North America's freshwater fishes. Biological Sciences **108:**10602–10607. doi:10.1073/pnas.1016437108.

Alshari NFMAH, Ahmad SZ, Azlan A, Lee YH, Azzam G, Nor SAM. 2021. Metabarcoding of fish larvae in Merbok River reveals species diversity and distribution along its mangrove environment. Zool Stud **60:**76. doi:10.6620/ZS.2021.60-6.

Bemis KE, Girard MG, Santos MD, Carpenter KE, Deeds JR, Pitassy DE, Flores NAL, Hunter ES, Driskell AC, Macdonald III KS. 2023. Biodiversity of Philippine marine fishes: A DNA barcode reference library based on voucher specimens. Scientific Data **10:**411. doi:10.1038/s41597-023-02306-9.

Bickford D, Lohman DJ, Sodhi NS, Ng PK, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. Trends Ecol Evol **22:**148–155. doi:10.1016/j.tree.2006.11.004.

Bini LM, Diniz-Filho JAF, Rangel TF, Bastos RP, Pinto MP. 2006. Challenging wallacean and linnean shortfalls: Knowledge gradients and conservation planning in a biodiversity hotspot. Divers Distrib **12:**475–482. doi:10.1111/j.1366-9516.2006.00286.x.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A software platform for bayesian evolutionary analysis. PLoS Comp Biology **10:**e1003537. doi:10.1371/journal.pcbi.1003537.

Bourlat SJ, Borja A, Gilbert J, Taylor MI, Davies N, Weisberg SB, Griffith JF, Lettieri T, Field D, Benzie J. 2013. Genomics in marine monitoring: New opportunities for assessing marine health status. Mar Pollut Bull **74:**19–31. doi:10.1016/j.marpolbul.2013.05.042.

Bruin G, Russell B, Bogusch A. 1995. The marine fishery resources of Sri Lanka. FAO species identification field guide for fishery purposes. Rome: FAO Library.

Burke L, Selig L. 2002. Reefs at risk in Southeast Asia. World Resources Institute.

Carpenter KE, Niem VH. 1999a. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 3. Batoid fishes, chimaeras and bony fishes part 1 (Elopidae to Linophrynidae). Rome: FAO Library.

Carpenter KE, Niem VH. 1999b. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 4. Bony fishes part 2 (Mugilidae to Carangidae). Rome: FAO Library.

Carpenter KE, Niem VH. 2001a. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 5. Bony fishes part 3 (Menidae to Pomacentridae). Rome: FAO Library.

Carpenter KE, Niem VH. 2001b. 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. Rome: FAO Library.

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:**796–805. doi:10.1111/1755-0998.12601.

Chin TC, Adibah A, Hariz ZD, Azizah MS. 2016. Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. Food Control **64:**247–256. doi:10.1016/j.foodcont.2015.11.042.

Chong V, Jamizan A, Rizman–Idid M, SH MA, Natin P. 2010. Diversity and abundance of fish and invertebrates of Semerak Estuary and adjacent inshore waters, Kelantan. Malaysian Journal of Science **29:**95–110. doi:10.22452/mjs.vol29nosp.11.

Collins R, Cruickshank R. 2013. The seven deadly sins of DNA barcoding. Mol Ecol Resour **13:**969–975. doi:10.1111/1755-0998.12046.

DeSalle R, Egan MG, Siddall M. 2005. The unholy trinity: Taxonomy, species delimitation and DNA barcoding. Philosophical Transactions of the Royal Society B: Biological Sciences **360:**1905–1916. doi:10.1098/rstb.2005.1722.

DiBattista JD, Roberts MB, Bouwmeester J, Bowen BW, Coker DJ, Lozano-Cortés DF, Howard Choat J, Gaither MR, Hobbs JPA, Khalil MT. 2016. A review of contemporary patterns of

endemism for shallow water reef fauna in the Red Sea. J Biogeogr **43:**423–439. doi:10.1111/jbi.12649.

DoF. 2000. Fisheries resources survey in the Exclusive Economic Zone (EEZ) of Malaysia. *In*: Department of Fisheries, Editor. Putrajaya Malaysia: Ministry of Agriculture Malaysia, p. 374.

DoF. 2018. Fisheries resources surveys in malaysian waters 2013–2016. *In*: Department of Fisheries, Editor. Putrajaya Malaysia: Ministry of Agriculture & Agro-Based Industry, pp. 809.

Du J, Loh K–H, Hu W, Zheng X, Affendi YA, Ooi JLS, Ma Z, Rizman-Idid M, Chan AA. 2019. An updated checklist of the marine fish fauna of Redang islands, Malaysia. Biodiversity Data Journal. **7:** e47537. doi:10.3897/bdj.7.e47537.

Dupuis JR, Roe AD, Sperling FA. 2012. Multi-locus species delimitation in closely related animals and fungi: One marker is not enough. Mol Ecol **21:**4422–4436. doi:10.1111/j.1365-294x.2012.05642.x

Edler D, Klein J, Antonelli A, Silvestro D. 2021. Raxmlgui 2.0: A graphical interface and toolkit for phylogenetic analyses using Raxml. Methods Ecol Evol **12:**373–377. doi:10.1111/2041-210x.13512.

Fang G, Susanto RD, Wirasantosa S, Qiao F, Supangat A, Fan B, Wei Z, Sulistiyo B, Li S. 2010. Volume, heat, and freshwater transports from the South China Sea to Indonesian Seas in the boreal winter of 2007–2008. J Geophys Res: Oceans **115:C12**. doi:10.1029/2010jc006225.

Fricke R, Eschmeyer W, van der Laan R. 2023. Eschmeyer's catalog of fishes: Genera, species, references. Available at: http://researcharchive.calacademy.org/research/ichthyology/ Accessed 1 Oct. 2023.

Froese R, Pauly D. 2023. Fishbase. www.fishbase.org. Accessed 1 Oct. 2023.

Fujisawa T, Barraclough TG. 2013. Delimiting species using single-ocus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. Syst Biol **62:**707–724. doi:10.1093/sysbio/syt033.

GBIF.org. 2023. Gbif home page. GBIF Secretariat Universitetsparken 15. Accessed 1 Oct. 2023

Hebert PD, Cywinska A, Ball SL, DeWaard JR. 2003a. Biological identifications through DNA barcodes. P Roy Soc Lond B-Bio **270:**313–321. doi:10.5772/49967.

Hebert PD, Ratnasingham S, De Waard JR. 2003b. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species. P Roy Soc Lond B-Bio **270:**S96–S99. doi:10.1098/rsbl.2003.0025.

Hou G, Chen WT, Lu HS, Cheng F, Xie SG. 2018. Developing a DNA barcode library for perciform fishes in the south china sea: Species identification, accuracy and cryptic diversity. Mol Ecol Resour **18:**137–146. doi:10.1111/1755-0998.12718.

Huang D, Hoeksema BW, Affendi YA, Ang PO, Chen CA, Huang H, Lane DJ, Licuanan WY, Vibol O, Vo ST. 2016. Conservation of reef corals in the South China Sea based on species and evolutionary diversity. Biodivers Conserv **25:**331–344. doi:10.1007/s10531-016-1052-7.

Hubert N, Dettai A, Pruvost P, Cruaud C, Kulbicki M, Myers RF, Borsa P. 2017. Geography and life history traits account for the accumulation of cryptic diversity among indo-west pacific coral reef fishes. Marine Ecology Progress Series. **583:**179–193. doi:10.3354/meps12316.

Hubert N, Hanner R. 2015. DNA barcoding, species delineation and taxonomy: A historical perspective. DNA Barcodes **3:**doi:10.1515/dna-2015-0006.

Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burridge M, Watkinson D, Dumont P, Curry A, Bentzen P. 2008. Identifying Canadian freshwater fishes through DNA barcodes. PLoS ONE **3:**e2490. doi:10.1371/journal.pone.0002490.

Hubert N, Kadarusman PD, Wibowo A, Busson F, Caruso D, Sulandari S, Nafiqoh N, Pouyaud L, Rüber L, Avarre J-C. 2015. DNA barcoding Indonesian freshwater fishes: Challenges and prospects. DNA Barcodes **3:**144–169. doi:10.1515/dna-2015-0018.

Hubert N, Meyer CP, Bruggemann HJ, Guerin F, Komeno RJ, Espiau B, Causse R, Williams JT, Planes S. 2012. Cryptic diversity in Indo-pacific coral reef fishes revealed by DNA barcoding provides new support to the centre-of-overlap hypothesis. PLoS ONE **7(3)**:e28987. doi:10.1371/journal.pone.0028987.

Ghazali SZ, Lavoué S, Mohd Abu Hassan Alshari NF, Zainal Abidin DH, Jamaluddin JAF, Tan MP, Mohd Nor SA. 2023. Cracking the glass-perchlet code: Integrative taxonomy uncovers high species-level diversity within the glass-perchlet genus *Ambassis* (Ambassidae) in tropical Asia. Zool Scr **00**:1–18. doi:10.1111/zsc.12640.

Imamura H. 2007. *Rogadius mcgroutheri*, a new species of flathead (Teleostei: Platycephalidae) collected from eastern australia and new caledonia. Ichthyol Res **54:**303–307. doi:10.1007/s10228-007-0405-3.

Imamura H, Gomon MF. 2010. Taxonomic revision of the genus *Ratabulus* (Teleostei: Platycephalidae), with descriptions of two new species from Australia. Memoirs of Museum Victoria. **67:**19–33. doi:10.24199/j.mmv.2010.67.03.

Imamura H, Knapp LW. 2009. *Platycephalus orbitalis*, a new species of flathead (teleostei: Platycephalidae) collected from western Australia. Zootaxa **2271:**57–63. doi:10.11646/zootaxa.2271.1.5.

Kalyaanamoorthy S, Minh B, Wong T, Von Haeseler A, Jermiin L. 2017. Modelfinder: Fast model selection for accurate phylogenetic estimates. Nat Methods **14:**587–589. doi:10.1038/nmeth.4285.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol **16:**111–120.

Knapp LW, Imamura H. 2004. *Sunagocia sainsburyi*, a new flathead fish (Scorpaeniformes: Platycephalidae) from northwestern Australia. P Biol Soc Wash **117:**545–550.

Knebelsberger T, Landi M, Neumann H, Kloppmann M, Sell AF, Campbell PD, Laakmann S, Raupach MJ, Carvalho GR, Costa FO. 2014. A reliable DNA barcode reference library for the identification of the North European shelf fish fauna. Mol Ecol Resour **14:**1060–1071. doi:10.1111/1755-0998.12238.

Lakra W, Verma M, Goswami M, Lal KK, Mohindra V, Punia P, Gopalakrishnan A, Singh K, Ward RD, Hebert P. 2011. DNA barcoding Indian marine fishes. Mol Ecol Resour **11:**60–71.

Landi M, Dimech M, Arculeo M, Biondo G, Martins R, Carneiro M, Carvalho GR, Brutto SL, Costa FO. 2014. DNA barcoding for species assignment: The case of mediterranean marine fishes. PLoS ONE **9:**:e106135. doi:10.1371/journal.pone.0106135.

Lara A, Ponce de León JL, Rodriguez R, Casane D, Cote G, Bernatchez L, García-Machado E. 2010. DNA barcoding of cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. Mol Ecol Resour **10:**421–430. doi:10.1111/j.1755-0998.2009.02785.x.

Luo A, Ling C, Ho SY, Zhu C-D. 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst Biol **67:**830–846. doi:10.1093/sysbio/syy011.

Mat Jaafar TNA, Taylor MI, Mohd Nor SA, de Bruyn M, Carvalho GR. 2012. DNA barcoding reveals cryptic diversity within commercially exploited Indo-malay Carangidae (Teleostei: Perciformes). PLoS ONE **7:**e49623. doi:10.1371/journal.pone.0049623.

Matsunuma M, Motomura K, Matsuura, Noor Azhar MS, 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, ix + 251 pages.

Mecklenburg CW, Møller PR, Steinke D. 2011. Biodiversity of arctic marine fishes: Taxonomy and zoogeography. Mar Biodivers **41:**109–140. doi:10.1007/s12526-010-0070-z.

Miller MA, Pfeiffer W, Schwartz T. 2011. The cipres science gateway: A community resource for phylogenetic analyses. Paper presented at: Proceedings of the 2011 TeraGrid Conference: extreme digital discovery.

Miralles A, Vences M. 2013. New metrics for comparison of taxonomies reveal striking discrepancies among species delimitation methods in madascincus lizards. PLoS ONE **8:**e68242. doi:10.1371/journal.pone.0068242.

Mora C, Tittensor DP, Adl S, Simpson AG, Worm B. 2011. How many species are there on earth and in the ocean? PLoS Biology **9:**e1001127. doi:10.1371/journal.pbio.1001127.

Motomura H, Kimura S, Seah YG, Sheikh Abdul Kadir ST, Ghaffar MA. 2021. Reef and Shore Fishes of Bidong Island, Off East Coast of Malay Peninsula. Kagoshima: The Kagoshima University Museum, p. 80.

Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature **403:**853–858. doi:10.1038/35002501.

Nadira YK, Mustapha M, Ghaffar M. 2019. The indian mackerel aggregation areas in relation to their oceanographic conditions. Sains Malaysiana **48:**2575–2581. doi:10.17576/jsm-2019-4811-27.

Pauly D, Liang C. 2020. The fisheries of the South China Sea: Major trends since 1950. Mar Policy **121:**103584. doi:10.1016/j.marpol.2019.103584.

Pecl GT, Ward TM, Doubleday ZA, Clarke S, Day J, Dixon C, Frusher S, Gibbs P, Hobday AJ, Hutchinson N. 2014. Rapid assessment of fisheries species sensitivity to climate change. Climatic Change **127:**505–520. doi:10.1007/s10584-014-1284-z.

Poling GB. 2013. The South China Sea in focus: Clarifying the limits of maritime dispute. Rowman & Littlefield.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst Biol **55:**595–609. doi:10.1080/10635150600852011.

Puckridge M. 2006. DNA barcoding of Australia's marine fish. Honours Thesis, University of Tasmania, Hobart.

Puckridge M, Andreakis N, Appleyard SA, Ward RD. 2013. Cryptic diversity in flathead fishes (Scorpaeniformes: Platycephalidae) across the Indo-west Pacific uncovered by DNA barcoding. Mol Ecol Resour **13:**32–42. doi:10.1111/1755-0998.12022.

Puillandre N, Brouillet S, Achaz G. 2021. ASAP: Assemble species by automatic partitioning. Mol Ecol Resour **21:**609–620. doi:10.1111/1755-0998.13281

Puillandre N, Lambert A, Brouillet S, Achaz G. 2012. ABGD, automatic barcode gap discovery for primary species delimitation. Mol Ecol **21:**1864–1877. doi:10.1111/j.1365-294x.2011.05239.x.

Rabosky DL, Chang J, Title PO, Cowman PF, Sallan L, Friedman M, Kaschner K, Garilao C, Near TJ, Coll M. 2018. An inverse latitudinal gradient in speciation rate for marine fishes. Nature. **559:**392–395. doi:10.1038/s41586-018-0273-1.

Rambaut A, Drummond AJ. 2013. Treeannotator v1.7.0.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in bayesian phylogenetics using tracer 1.7. Syst Biol **67:**901–904. doi:10.1093/sysbio/syy032.

Ramirez JL, Birindelli JL, Carvalho DC, Affonso PR, Venere PC, Ortega H, Carrillo-Avila M, Rodriguez-Pulido JA, Galetti Jr PM. 2017. Revealing hidden diversity of the underestimated neotropical ichthyofauna: DNA barcoding in the recently described genus megaleporinus (Characiformes: Anostomidae). Front Genet **8:**149. doi:10.3389/fgene.2017.00149.

Ratnasingham S, Hebert PD. 2007. BOLD: The barcode of life data system. Mol Ecol Notes **7:**355–364. doi:10.1111/j.1471-8286.2007.01678.x.

Ratnasingham S, Hebert PD. 2013. A DNA-based registry for all animal species: The barcode index number (BIN) system. PLoS ONE **8:**e66213. doi:10.1371/journal.pone.0066213.

Sanbrook J, Fritsch E, Maniatis T. 2001. Molecular cloning: A laboratory manual. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory **11:**31.

Seah YG, Ali MS, Ghaffar MA, Mat Jaafar TNA. 2021. Marine Fishes of Kuantan, Malaysia Biodiversity Information System (MYBIS). Terengganu: Penerbit Universiti Malaysia Terengganu, p. 172.

Seah YG, Ariffin AF, Mat Jaafar TNA. 2017. Levels of *COI* divergence in Family Leiognathidae using sequences available in GenBank and BOLD Systems: A review on the accuracy of public databases. AACL Bioflux **10:**391–401.

Seah YG, Mat Jaafar TNA, Ali MS. 2020. Field Guide to Trawl Fishes Near Bidong Island. Terengganu: Penerbit Universiti Malaysia Terengganu, p. 144.

Sholihah A, Delrieu-Trottin E, Sukmono T, Dahruddin H, Risdawati R, Elvyra R, Wibowo A, Kustiati K, Busson F, Sauri S. 2020. Disentangling the taxonomy of the subfamily rasborinae (Cypriniformes, Danionidae) in sundaland using DNA barcodes. Sci Rep-UK **10:**2818. doi:10.1038/s41598-020-59544-9.

Steinke D, Zemlak TS, Hebert PD. 2009. Barcoding nemo: DNA–based identifications for the ornamental fish trade. PLoS ONE **4:**e6300. doi:10.1371/journal.pone.0006300.

Teh LC, Pauly D. 2018. Who brings in the fish? The relative contribution of small–scale and industrial fisheries to food security in Southeast Asia. Frontiers in Marine Science **5:**44. doi:10.3389/fmars.2018.00044.

van der Laan R, Fricke R, Eschmeyer W. 2023. Eschmeyer's catalog of fishes: Classification. Available at: http://www.calacademy.org/scientists/catalog-of-fishes-classification/. Accessed 1 Oct. 2023.

Villesen P. 2007. Fabox: An online toolbox for fasta sequences. Mol Ecol Notes **7:**965–968. doi:10.1111/j.1471-8286.2007.01821.x.

Ward RD, Costa FO, Holmes BH, Steinke D. 2008. DNA barcoding of shared fish species from the North Atlantic and Australasia: Minimal divergence for most taxa, but *Zeus faber* and *Lepidopus caudatus* each probably constitute two species. Aquat Biol **3:**71–78. doi:10.3354/ab00068.

Ward RD, Hanner R, Hebert PD. 2009. The campaign to DNA barcode all fishes, FISH-BOL. J Fish Biol **74:**329–356. doi:10.1111/j.1095-8649.2008.02080.x.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society B: Biological Sciences **360:**1847–1857.

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. doi:10.1016/j.fishres.2020.105771.

Zainal Abidin DH, Mohd. Nor SA, Lavoue S, A. Rahim M, Jamaludin NA, Mohammed Akib NA. 2021. DNA-based taxonomy of a mangrove–associated community of fishes in Southeast Asia. Sci Rep-UK **11:**17800. doi:10.1038/s41598-021-97324-1.

Zainal Abidin DH, Mohd. Nor SA, Lavoué S, A. Rahim M, Mohammed Akib NA. 2022. Assessing a megadiverse but poorly known community of fishes in a tropical mangrove estuary through environmental DNA (eDNA) metabarcoding. Sci Rep-UK **12:**16346. doi:10.1038/s41598-022-19954-3.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PD. 2009. DNA barcoding reveals overlooked marine fishes. Mol Ecol Res **9:**237–242. doi:10.1111/j.1755-998.2009.02649.x.

Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics **29:**2869–2876. doi:10.1093/bioinformatics/btt499.

Zhang J, Zhang K, Jiang Y-e, Li Y-j, Fan J-t, Yu W-m, Chen Z-z. 2022. Diversity and structure of demersal fish community over the northern slope in the South China Sea. Frontiers in Marine Science **9:**809636. doi:10.3389/fmars.2022.809636.

Supplementary materials

Table S1. Comparison of K2P intraspecific (bold values) and interspecific genetic divergences of OTUs separated by lineages. The italicised values are the standard error. (download)

Table S2. List of GenBank accession numbers with associated information including sample/museum ID, Barcode of Life. (download)