DNA BARCODING OF THE FAMILY SCOMBRIDAE IN MALAYSIA AND PHYLOGEOGRAPHY AND POPULATION STRUCTURE OF THE INDIAN MACKEREL, *Rastrelliger kanagurta* FOR SUSTAINABLE FISHERIES

NOOR ADELYNA MOHAMMED AKIB

UNIVERSITI SAINS MALAYSIA
2013

DNA BARCODING OF THE FAMILY SCOMBRIDAE IN MALAYSIA AND PHYLOGEOGRAPHY AND POPULATION STRUCTURE OF THE INDIAN MACKEREL, *Rastrelliger kanagurta* FOR SUSTAINABLE FISHERIES

by

NOOR ADELYNA MOHAMMED AKIB

Thesis submitted in fulfillment of the requirements
for the degree of
Doctor of Philosophy

November 2013

ACKNOWLEDGMENTS

My sincere, utmost delight and profound gratitude goes to my supervisor Prof. Siti Azizah Mohd. Nor, thank you so much for your continuous guidance, encouragement, absolute patience, invaluable advice, (*occasional nagging*) and especially for believing in me, in which I am able to complete this thesis. May Allah repay you in abundance and bless you with all his grace. I am profoundly grateful to Prof. Peter Mather, my co-supervisor from Queensland University of Technology (QUT). Thank you so much for having me in your lab for two months and also for your guidance, invaluable advice and patience especially while looking through my thesis. I am forever grateful to Dr. Hyungtaek Jung and Mr. Vincent Chand from QUT for helping me with the Ion Torrent experiment and the analyses, thank you so much from the bottom of my heart. My special thanks to Prof. Geoffrey Chambers, Victoria University of Wellington, for sharing your knowledge and expertise during your many visits here in USM. Many thanks to Fisheries Research Institutes, Penang and Department of Fisheries throughout Malaysia for helping me with my sampling.

To my mates of lab 308, thank you so much for teaching me to do the lab work, helping me with my sampling and thesis analyses, companionship over so many lunches and occasional dinner. I will miss our 'intellectual' and sometimes silly discussions. From the bottom of my heart I am thankful to Jamsari, Su Yin, Danial, Faisal, Lutfi, Fong, Nazia, Hong Chiun, Min Pau, Adibah, Ana, Balkhis, Hasmah, Kak Masa, Elham, Layla, Amirah, Zalina, Suzana, Katie and everyone that might have cross my path.

My profound gratitude goes to all my family and especially to my dear beloved husband, Hizam. Thank you so much for supporting me in my study and especially for putting up with me during the whole 3 years of my study. My deepest thanks to the lovely ladies - Iza, Selvi and Vivian, thank you for always being there for me and two decades of friendship!

Thank you to Universiti Sains Malaysia for funding my study under the Academic Staff Higher Education Scholarship (ASHES) and also for funding my research under the Research University Grant (1001/PBIOLOGI/815051) and Postgraduate Research Grant Scheme (1001/PPANTAI/844103).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

AMOVA – Analysis of molecular variance

BOLD – Barcode of life database

COI – Cytochrome c oxidase subunit I

CNI – Close-neighbor-interchange

Cyt *b* – Cytochrome *b*

EST – Express sequence tag

FDR – False discovery rate

GO – Gene ontology

GSS – Genome survey sequence

HWE – Hardy-Weinberg equilibrium

IAM – Infinite allele mutation

KEGG – Kyoto encyclopedia of genes and genomes

NGS – Next generation sequencing

NJ – Neighbor-joining

MSN – Minimum spanning network

MP – Maximum parsimony

RAPD – Random amplified polymorphic DNA

SAMOVA – Spatial analysis of molecular variance

SSM – Single stepwise mutation

SSR – Single sequence repeat

LIST OF SYMBOLS

#V – number of variable sites

π – nucleotide diversity

θs – number of segregating sites

A_R – allelic richness

F_{IS} – inbreeding coefficient

F_{CT} – variance among groups

F_{ST} – variance within population

F_{SC} – variance among populations within group

H – number of haplotypes

Hd – haplotype diversity

H_O – observed heterozygosity

H_E – expected heterozygosity

H*ri* – Harpending's raggedness index

k – number of groups

N – sample size

N_A – number of alleles

Nm – gene flow estimates

MP – maximum parsimony

R_2 – population growth estimate

PENGKODAN DNA FAMILI SCOMBRIDAE DI MALAYSIA DAN FILOGEOGRAFI DAN STRUKTUR POPULASI IKAN KEMBONG BOREK, *Rastrelliger kanagurta* UNTUK PERIKANAN LESTARI.

ABSTRAK

Kajian ini dijalankan untuk membangunkan pengecaman sistematik molekul bagi famili Scombridae yang berkepentingan komersil dan juga kefahaman filogeografi dan filogenetik satu spesies ahli dari kumpulan ini, iaitu, kembong borek, *Rastrelliger kanagurta,* telah dijalankan bagi memastikan perikanan lestari di perairan Malaysia. Pengkodan DNA berasaskan gen sitokrom oksidase subunit I mitokondria atau secara meluas dikenali sebagai gen pengkodan telah berjaya mengenalpasti dan memisahkan 14 spesies daripada famili Scombridae yang disampel di perairan Malaysia. Filogeografi dan struktur populasi kembong borek, *R. kanagurta* juga telah dikaji di kalangan 19 populasi dari Malaysia dan setiap satu populasi dari Thailand, Vietnam, Indonesia (kawasan selatan timur Asia) dan satu populasi dari Iran (lautan India Barat – WIO) juga menggunakan gen sitokrom *b* mitokondria. Satu kajian selari telah dijalankan ke atas 11 populasi dari Malaysia dan satu populasi dari Iran berdasarkan lapan penanda mikrosatelit yang baru dibangunkan. Dalam kajian ini, pelantar jujukan generasi hadapan Ion Torrent PGM^{TM} telah digunakan untuk menjanakan set data jujukan 'survey' genom (GSS) separa bagi pembangunan penanda mikrosatelit daripada DNA genom *R. kanagurta.* Data yang terjana termasuk bacaan jujukan sebanyak 399,794 (81.29 Mbp) di mana 16,209 bacaan jujukan telah berjaya dikumpul dan menghasilkan 327 kontig dengan purata panjang 677 bp di samping pembangunan penanda ulangan jujukan tertunggal (SSR). Berdasarkan keputusan GSS-Blastx, 18.3% kontig mempunyai persamaan

signifikan (nilai $E < 10^{-6}$) dengan data yang sedia ada, dengan majoritinya sepadan dengan jujukan ikan yang telah dilaporkan. Analisis KEGG telah mengenalpasti dua laluan metabolisme yang memberi pemerhatian khusus ke arah potensi peranan spesifik dan fungsi jujukan yang terlibat di dalam proses molekul dalam *R. kanagurta.* Domain protein utama telah dikenalpasti termasuk imunoglobin dan transkriptase terbalik. Sejumlah 7891 motif berulang yang mengandungi SSR di mana 1688 daripadanya layak untuk merekabentuk primer. Selepas pengoptimum dan ujian bagi memastikan kebolehgunaan semula dan polimorfisme, lapan penanda mikrosatelit telah dikenalpasti sesuai untuk analisis genetik populasi. Penanda mitokondria (23 populasi) dan penanda mikrosatelit (12 populasi) secara konsisten telah membuktikan variasi genetik intrapopulasi yang tinggi untuk semua populasi *R. kanagurta* yang telah dikaji. Secara am, data kajian ini telah membuktikan spesies di perairan Malaysia terdiri daripada saiz bancian yang baik dengan tahap variasi genetik yang sihat yang diperlukan untuk adaptasi evolusi terhadap persekitaran yang sering berubah. Penanda mikrosatelit juga menunjukkan populasi telah menyisih daripada keseimbangan Hardy-Weinberg, yang disebabkan oleh kehadiran alel nol. Pembezaan populasi yang rendah dalam kalangan populasi *R. kanagurta* di sepanjang empat lautan sekitar Malaysia dan kawasan berdekatan (Thailand, Vietnam, Indonesia) adalah bersesuaian dengan tabiat ikan marin yang sentiasa berhijrah dan mempunyai penyebaran larva yang tinggi berserta aliran gen yang tinggi, dan juga ketiadaan sempadan fizikal di persekitaran marin. Walaubagaimana pun, dua stok berasingan telah dikenalpasti apabila populasi Iran dimasukkan, pertama terdiri daripada populasi Asia Tenggara dan kedua terdiri daripada populasi WIO tunggal. Ini adalah kerana WIO dan kawasan Asia Tenggara terlibat dengan pengasingan berkurun yang diakibatkan oleh turun naik aras laut semasa lewat

tempoh Pleistocene, seterusnya telah mengakibatkan permisahan filogeografi untuk spesies ini di arah utara laut Andaman. Walaubagaimana pun, tiada bukti corak penstrukturan geografi yang jelas di rantau Asia Tenggara, maka ini menunjukkan populasi di rantau ini ialah homogen dan pengasingan berkurun tidak mencukupi untuk *R. kanagurta* mencapai keseimbangan peralihan migrasi. Berdasarkan kajian ini, *R. kanagurta* dianggap sebagai satu unit pengurusan di Malaysia, dan perlu diuruskan sewajarnya. Walaubagaimana pun, oleh kerana spesies ini mempunyai kepentingan komersil, kemungkinan risiko eksploitasi yang berlebihan boleh berlaku. Oleh itu langkah-langkah pengawalan perlu dikuatkuasa. Sehingga kini, tiada dokumentasi tangkapan berlebihan dan eksploitasi berlebihan *R. kanagurta* di negara ini.

DNA BARCODING OF THE FAMILY SCOMBRIDAE IN MALAYSIA AND PHYLOGEOGRAPHY AND POPULATION STRUCTURE OF THE INDIAN MACKEREL, *Rastrelliger kanagurta* FOR SUSTAINABLE FISHERIES

ABSTRACT

This research was conducted to develop a molecular systematic identification of the commercially important family Scombridae as well as to understand the phylogeography and phylogenetics of a member species, the Indian mackerel, *Rastrelliger kanagurta* for sustainable fisheries in Malaysian waters. DNA barcoding based on the mitochondrial cytochrome oxidase subunit I gene, widely known as the barcoding gene successfully identified and delineated 14 species of the family Scombridae sampled in Malaysian waters. The phylogeography and population structure of the pelagic Indian mackerel, *R. kanagurta* was also investigated among 19 Malaysian populations and one population each from Thailand, Vietnam, Indonesia (Southeast Asian region) and an Iranian population (West Indian Ocean-WIO) using mitochondrial cytochrome *b* gene analysis. A parallel investigation was conducted on 11 Malaysian populations and an Iranian population based on eight newly developed microsatellite markers. In this study, the next generation sequencing platform, Ion Torrent PGM^{TM} was used to generate a partial genome survey sequence (GSS) dataset to develop microsatellite markers from *R. kanagurta* genomic DNA. The data generated included a total of 399,794 sequence reads (81.29 Mbp) of which 16,209 sequence reads were successfully assembled, producing 327 contigs averaging 677 bp in length in addition to the single sequence repeats (SSR) markers. Based on GSS-BLASTx results, 18.3% of the contigs possessed significant similarities (E value < 10^{-6}) to the available data,

with the majority of them matching well to reported fish sequences. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis identified two metabolic pathways that provided insights into the specific potential roles and functions of the sequences involved in molecular processes in *R. kanagurta*. The top protein domains detected included immunoglobin and reverse transcriptase. A total of 7891 SSR-containing motif repeats of which were found 1688 qualified for primer design. After optimization and testing for reproducibility and polymorphism, eight microsatellite markers were deemed suitable for use in population genetics analysis. Mitochondrial (23 populations) and microsatellite markers (12 populations) consistently revealed high intrapopulation genetic variations in all the *R. kanagurta* populations investigated. In general the data revealed that the species in these waters are composed of a good census size with healthy genetic variability levels essential for evolutionary adaptation to the rapidly changing environment. The microsatellite markers also revealed that the populations deviated from Hardy Weinberg equilibrium, attributable to the high occurrence of null alleles. Shallow population differentiation among *R. kanagurta* populations across the four surrounding seas of Malaysia and the neighbouring areas (Thailand, Vietnam, Indonesia) were observed typifying a migratory marine fish species with potentially high larval dispersal associated with high levels of gene flow, and absence of physical boundaries. However, two distinct stocks were revealed when the Iranian population (WIO) was included; the first comprising of the Southeast Asian populations and the second, the sole WIO population. Plausible explanations for the genetic differentiation observed between the WIO and Southeast Asian regions suggested historical isolation as a result of fluctuations in sea levels during the late Pleistocene which resulted in a phylogeograhic break for this species to the north of the Andaman Sea. However,

there was no clear geographical structuring pattern evident within the Southeast Asian region, thus indicating homogeneity of the *R. kanagurta* populations in this region and that historical isolation was insufficient for *R. kanagurta* to attain migration drift equilibrium. Based on these findings, this species is considered as a single management unit in Malaysia, and should be managed accordingly. However, due to its commercial importance, there is a real risk of overexploitation and therefore regulatory measures should be enforced. There is no documentation of overfishing and overexploitation of *R. kanagurta* populations in this country at present.

CHAPTER 1.0

INTRODUCTION

1.1 Introduction

The Scombridae is a family of fast-swimming marine epipelagic fishes. They are among the more highly migratory marine teleosts with cosmopolitan distributions in tropical and temperate regions. This family is composed of 51 presumed species in 15 genera within two subfamilies (Scombrinae and Gasterocismatinae). Members possess many morphological and physiological characteristics for adaptations in various marine habitats that are of great interests to physiologists and evolutionary biologists (Collete, *et al.*, 2001). In Malaysian waters, 20 presumed species within 11 genera have been recorded in this family (Froese & Pauly, 2013). The most dominant representative genera caught in fish landing sites in Malaysia are *Thunnus, Euthynnus, Auxis, Katsuwonus, Scomberomous* and *Rastrelliger* (Department of Fisheries Malaysia, 2006). All are of commercial importance. Traditionally, identification of this family has been based on morphological keys and generally straight forward (Collette & Nauen, 1983). Species identification of early larval stages, however is difficult since morphological characters are not well characterised in them (Paine *et al.*, 2008). Such information is vital for formulating conservation strategies such as in the determination of spawning sites. A molecular approach that has gained wide application in taxonomy, DNA barcoding, can be applied to overcome this limitation as well as for the assessment of cryptic species.

In brief, DNA barcoding is a universal bio-identification system based on a short, standardized gene region of the mitochondrial DNA cytochrome oxidase I

(COI) gene (Hebert & Gregory, 2005). It has become a popular tool for species discrimination by highlighting genetically distinct groups exhibiting levels of sequence divergence suggestive of species status. In this study, a DNA barcoding project for the family Scombridae was carried out to develop a molecular species identification of this family in Malaysia. For fishery managers and aquaculturists, DNA barcoding can serve many important functions including for elucidation of systematics and to permit phylogenetic insights, marketing, substitutions, quota and bycatch management, identification of fragmentary and processed products as well as conservation and management of commercially important and endangered species (Ward *et al.,* 2005; Dawnay *et al.,* 2007; Armstrong & Ball, 2009; Ogden *et al.,* 2009; Ardura *et al.,* 2010; Botti & Giuffra, 2010; Miller & Mariani, 2010; Nagoshi *et al.,* 2011).

This project is an important contribution to the Consortium for the Barcode of Life database (CBOL - http://www.barcoding.si.edu/) established in 2003, aimed at barcoding all groups of life forms. One of the major global project in CBOL is the Fish Barcode of Life (FISH-BOL www.fishbol.org/), initiated to barcode each of the more than 20,000 marine and 15,000 freshwater species in the world. The DNA barcoding study will be detailed in Chapter 3.

Since the advancement of genetic approaches in fisheries in the 1950's, the use of genetic markers has had major impacts on three fisheries areas; stock structure analysis, taxonomic analysis and aquaculture (Ward & Grewe, 1994). This study focused on two of these; stock structure analysis and taxonomic analysis. However, the data accumulated would also assist in planning future aquaculture programme. In the management and conservation of commercially important group(s), just as critical as the systematics and precise taxonomic identification, is the population

genetics data. In this study, a population genetics investigation focused on the pelagic and highly migratory Indian mackerel, *Rastrelliger kanagurta* from the family Scombridae inferred from mitochondrial DNA cytochrome *b*. This would be described in Chapter 4.

Rastrelliger kanagurta, known locally as 'kembung', 'kembung borek', 'mabong' and 'rumahan' is widely distributed across the tropical Indo-West Pacific region, from South Africa to the Red Sea, east through Indonesia and off northern Australia to Samoa, Malaysia, China, Ryukyu Islands and the east Mediterranean (Mohsin & Mohd. Azmi, 1996). *Rastrelliger kanagurta* and its two congeners, *R. brachysoma*, *R. faughni* are three of the most important commercial fish species in Malaysia providing cheap protein and as fish baits (FAO, 1987; Froese & Pauly, 2013). *Rastrelliger* also constitutes the most dominant marine food fish on the west coast of Peninsular Malaysia (Chee, 2000). They also contribute to the small pelagic fishery that provides one of the important marine resources (Mansor *et al.*, 1996). Total *Rastrelliger* landings increased from 101,003 tonnes in 2000 (Chee, 2000) to 185,463 tonnes in 2009 (Department of Fisheries Malaysia, 2009). Several molecular studies on *R. kanagurta* have been reported based on mitochondrial DNA and RAPD markers in Malaysia (Jayasankar *et al.,* 2004; Darlina *et al.,* 2011; Ahmad Faisal *et al.*, 2012). Earlier studies had been reported by Menezes *et al.* (1993) and Jayasankar *et al.,* (2004) using allozymes and RAPD markers respectively in the Indian Peninsular.

Several methods for stock identification have been used, such as parasite distribution, morphometrics and meristics, alloenzymes and DNA analysis (Menezes *et al.* 1993; Elliott & Ward, 1995; Cadrin & Friedland, 1999; Begg & Waldman, 1999; Sun *et al.*, 2012). Since the advent of molecular genetics and fishery

management in the 1950s, the use of molecular genetics markers have addressed the issues of stock identification and population structure, analysis of mixed-stock harvests, and assessments of levels of genetic variations within populations (Waples & Naish, 2009) with much success (Reiss *et al.*, 2009; Wang *et al.*, 2011; Sun *et al.*, 2012).

Mitochondrial DNA has proven to be a very useful marker and has been extensively used as a marker for evolutionary and population genetics studies in fisheries management (Garber *et al.*, 2005; Bakke & Johansen, 2005; Hoolihan *et al.*, 2006; Boustany *et al*., 2008) due to its higher sequence variability when compared with most single copy nuclear genes (Brown *et al.*, 1979). Since mtDNA genes are maternally inherited, the whole genome behaves as a single, non-recombining locus where all sites share a single genealogical history (Galtier *et al.*, 2009). However, to have a complete understanding, nuclear markers should also be investigated to assess the paternal lineage as well.

One such marker, microsatellite or simple sequence repeat (SSR) marker consists of tandemly repeated mono- to hexanucleotide motifs dispersed throughout the genome. They are usually characterized by high degrees of polymorphisms when compared with those of other molecular markers. Since SSRs contains high variability, they are very powerful genetic markers with applications that span over a wide area from forensic DNA studies to population genetics, aquaculture and conservation (Zane *et al.*, 2002). Bearing this information in mind, this technique was used to complement mtDNA sequencing in this project.

Until fairly recently, the development of microsatellites markers had been costly and time consuming. The conventional approaches had been based on DNA

sequence information deposited in databases, or on the screening of genomic DNA libraries. With the advent of next-generation sequencing (NGS), the development of SSR has become more efficient and cost effective. Once developed, they can be used for a plethora of conservation and aquaculture objectives at comparable costs to other less efficient markers (Liu & Cordes, 2004). It could also potentially be used to cross- amplify other closely related and important species saving costs of development. There are already several NGS platforms available in the market; 454 (Roche Applied Science), Illumina (Illumina), SOLiD (Life Technologies), HeliCope (Helicos) and Ion Torrent (Life Technologies) (Glenn, 2011). In this study the SSR markers were developed using Ion Torrent Technologies. For a holistic approach to the study, novel nuclear microsatellite markers would be developed using Next Generation Sequencing based on the Ion Torrent platform and utilised in complement to the mtDNA data (Chapter 5 & 6).

Rastrelliger kanagurta is now one of the flagship species of the Fisheries Department, Malaysia. Work is on-going between the Department in collaboration with international bodies and national institutions including Universiti Sains Malaysia, Southeast Asian Fisheries Development Center (SEAFDEC) and Bay of Bengal Large Marine Ecosystem Programme (BOBP-LME) on various aspects of this species. The data generated from the present study would contribute towards this important programme. Furthermore, the study of population genetics of any member of the family Scombridae has never been recorded in Malaysia.

Although, as aforementioned, Malaysia is recognized as one of the biodiversity hotspots of the world yet it is also ranked as one of the highest in the IUCN red list of nations with many threatened and endangered species presumably due to its dependence on capture fisheries and lack of effective conservation

management. It is hoped that this study could address some of the issues listed before through the use of molecular techniques.

1.2 Objectives

Therefore to address the issues described above, this study was aimed to achieve the following objectives:

1. To barcode the Family Scombridae of Malaysia
2. To analyse the population genetics and phylogeograpy of *Rastrelliger kanagurta* based on mitochondrial cytochrome *b*
3. To develop novel microsatellite primers and to assess the population genetic variation of *Rastrelliger kanagurta* (to complement mtDNA as data as in 2)

The first objective is detailed in Chapter 3, followed by the second objective which is detailed in Chapter 4. Objective 3 is divided into two sections – development of novel microsatellite markers of *R. kanagurta* as described in Chapter 5 and population genetic variation of this species inferred from novel microsatellite markers in Chapter 6.

CHAPTER 2.0

LITERATURE REVIEW

2.1 Family Scombridae

The family Scombridae is composed of mostly epipelagic marine fishes known collectively as the mackerels; Spanish mackerels, bonitos, and tunas. There are 15 genera and 51 presumed species worldwide (Collette *et al.,* 2001). These are divided into two subfamilies, the Gasterochismatinae with a single species, *Gasterochisma melampus* and the Scombrinae which consists of four tribes – Sardini (bonitos), Scombrini (mackerels), Scomberomorini (Spanish mackerels) and Thunnini **(**tunas) (Figure 2.1). There are to be reported 11 genera and 20 species of the family Scombridae (Froese & Pauly, 2013) (Table 2.1) in Malaysia. The taxonomic hierarchy of Scombridae according to ITIS (Integrated Taxonomic Information System) is as below:

Kingdom	Animalia
Phylum	Chordata
Subphylum	Vertebrata
Superclass	Osteichthyes
Class	Actinopterygii
Subclass	Neopterygii
Infraclass	Teleostei
Superorder	Acanthopterygii
Order	Perciformes
Suborder	Scombroidei
Family	Scombridae
	Direct Children:
Subfamily	Gasterochismatinae Lahille, 1903
Subfamily	Scombrinae Bonaparte, 1831

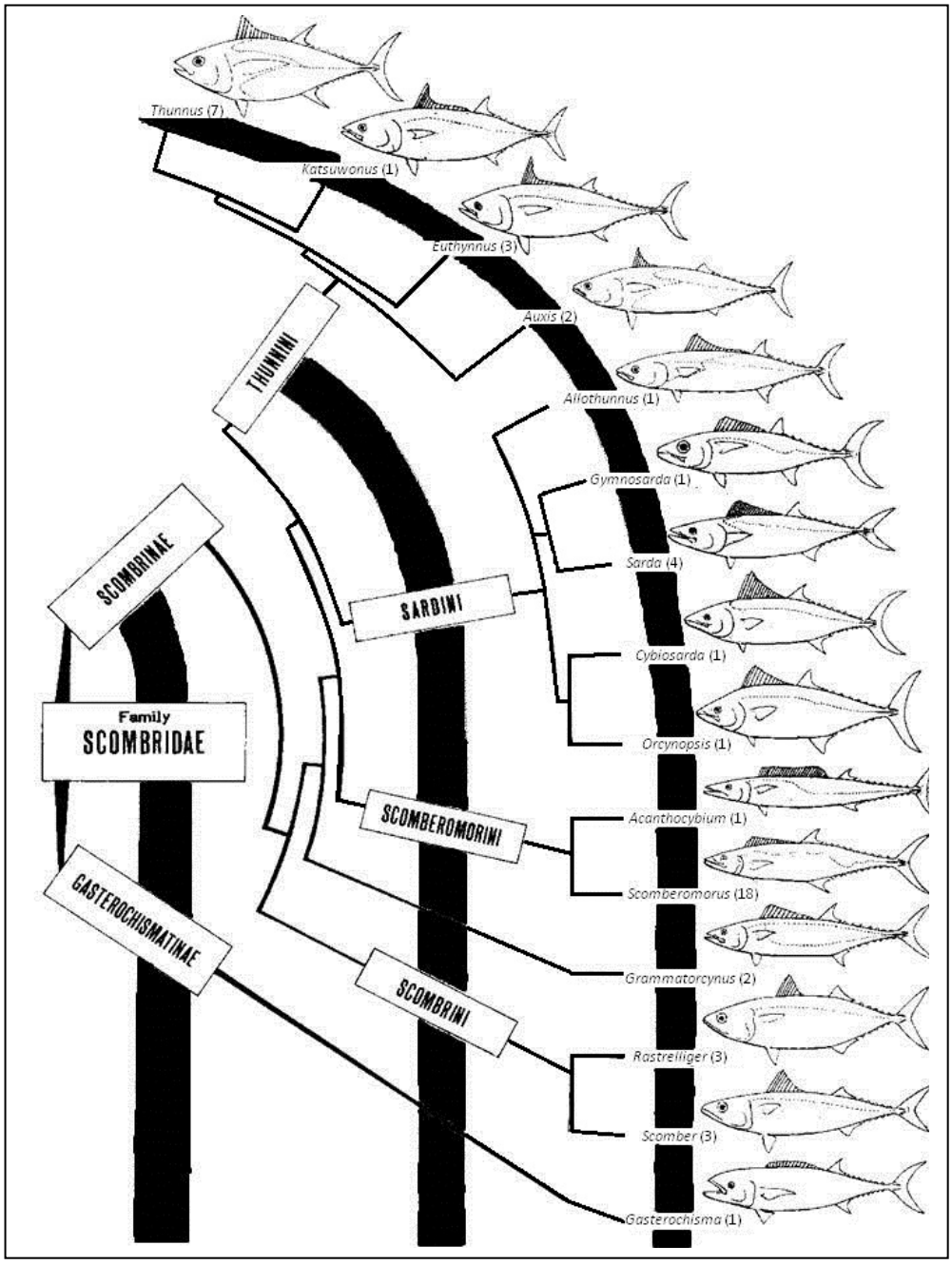

Figure 2.1: Classification of the Family Scombridae (Collette & Nauen, 1983)

Table 2.1: List of Scombridae species in Malaysian waters.

Tribes	Species name	Common name	Distribution
Scomberomorini	1. *Acanthocybium solandri* (Cuvier, 1832)	Wahoo/Tenggiri wahoo/Barracuda	Atlantic, Indian and Pacific Oceans
	2. *Grammatorcynus bilineatus* (Rüppell, 1836)	Double-lined mackerel / Aya	Indian Ocean, Western Pacific
	3. Scomberomorus commerson (Lacepède, 1800)	Narrow-barred Spanish mackerel / Tenggiri	Indo-West Pacific
	4. *Scomberomorus guttatus* (Bloch & Schneider, 1801)	Indo-Pacific king mackerel / Tenggiri	Indo-West Pacific
	5.*Scomberomorus koreanus* (Kishinouye, 1915)	Korean seerfish / Tenggiri	Indo-West Pacific
	6. *Scomberomorus lineolatus* (Cuvier, 1829)	Streaked seerfish / Tenggiri	Indo-West Pacific
Thunnini	7. *Auxis rochei* (Risso, 1810)	Bullet tuna/Aya/Aya peluru/Aya selaseh/Kayu/Tongkol	Atlantic, Indian and Pacific Oceans
	8. *Auxis thazard* (Lacepède, 1800)	Frigate tuna/Aya Kurik/Aya selaseh/Bakulan/Tongkol kurik/selaseh	Atlantic, Indian and Pacific Oceans
	9. Euthynnus affinis (Cantor, 1849)	Kawakawa / Aya / Aya Kurik/ Tongkol kurik/ Kayu	Indo-West Pacific
	10. *Katsuwonus pelamis* (Linnaeus, 1758)	Skipjack tuna / Oceanic bonito / Aya / Aya jalur / Aya Jepun / Tongkol	Cosmopolitan in tropical and warm-temperate waters
	11.*Thunnus alalunga* (Bonnaterre, 1788)	Albacore / Aya / Kayu / Tongkol	Cosmopolitan in tropical and temperate waters of all oceans
	12. *Thunnus albacares* (Bonnaterre, 1788)	Yellowfin tuna / Aya / Aya tuna / Tongkol	Tropical and subtropical seas, but absent from the Mediterranean Sea
	13. *Thunnus obesus* (Lowe, 1839)	Bigeye tuna / Aya / Aya Hitam / Kayu / Tongkol	Atlantic, Indian and Pacific
	14. *Thunnus tonggol* (Bleeker, 1851)	Longtail tuna /	Indo-West Pacific
Sardini	15. *Gymnosarda unicolor* (Rüppell, 1836)	Dogtooth tuna / Tuna / Tuna Tenggiri	Indo-Pacific
	16. *Sarda orientalis* (Temminck & Schlegel, 1844)	Striped bonito / Aya / Aya bonito / Kayu / Tenggiri	Southwest Pacific
Scombrini	17. *Rastrelliger brachysoma* (Bleeker, 1851)	Short mackerel / kembong / kembong pelaling / kembong perempuan	Pacific Ocean: Andaman Sea to Thailand, Indonesia, Papua New Guinea, Philippines, Solomon Islands and Fiji.
	18 *Rastrelliger faughni* Matsui, 1967	Island mackerel / Kembong / Kembong Lampai / Mabong / Rumahan	Indo-West Pacific
	19. *Rastrelliger kanagurta* (Cuvier, 1816)	Indian mackerel / Kembong / Kembong Borek / Kembong Jantan	Indo-Pacific, Eastern Pacific
	20. *Scomber australasicus* (Cuvier, 1832)	Blue mackerel / Aya / Tenggiri biru	Indo-West Pacific

(Froese & Pauly, 2013)

2.2 Nomenclature and Taxonomy

The taxonomy of the family Scombridae has been well documented (Collette & Nauen, 1983; Collette *et al.*, 2001; Collette, 2002; Collette, 2003) based on traditional approaches. Generally individuals have an elongated, fusiform body with two dorsal fins, and a series of finlets behind the rear dorsal fin and anal fin. The caudal fin is deeply forked with supporting caudal rays completely covering the hypural plate. There are at least two small keels on each side of the caudal fin base but in more advanced species, a larger keel is found in between the caudal peduncle, in addition to the two small keels.

Species length varies from 20 cm in the island mackerel, *Rastrelliger faughni* to 458 cm recorded for the very large Atlantic bluefin tuna, *Thunnus thynnus*. The only species in the subfamily Gasterochismatinae, *Gasterochisma melampus* is not present in Malaysia but members of the subfamily Scombrinae that consists of four tribes (with presumably 20 species) – Scomberomorini (Spanish mackerels), Thunnini **(**tunas), Sardini (bonitos) and Scombrini (mackerels) can be found in Malaysian waters.

2.2.1 Subfamily Scombrinae

2.2.1.1 Tribe Scomberomorini

There are three genera that included this tribe; *Acanthocybium* Gill, 1862**,** *Scomberomorus* Lacepède, 1801 and *Grammatorcynus* Gill, 1862. The genus *Acanthocybium* is monotypic, and is represented by *A. solandri (*Figure 2.2) which is commonly known as wahoo, a large species reaching over 1500 mm (SL) (Collette *et al.*, 2001) that is found in Malaysian waters. The genus *Scomberomorus* consists of

18 species (Collette & Russo, 1984), of which four can be found in Malaysia; namely *S. commerson*, *S. guttatus*, *S. koreanus* and *S. lineolatus* (Figure 2.3). The genus *Grammatorcynus* has two representative species, of which only *G. bilineatus* (Plate 2.1) has been reported in Malaysia (Froese & Pauly, 2013).

2.2.1.2 Tribe Thunnini

There are five genera under this tribe of which four can be found in Malaysia, namely *Auxis* Cuvier, 1829; *Euthynnus* Lütken, 1882; *Katsuwonus* Kishinouye, 1923 and *Thunnus* South, 1845. According to Collette *et al.*, (2001), the genus *Auxis* is the next most primitive genus after *Allothunnus.* Two of the four species included in this genus namely *A. thazard* (Plate 2.2) and *A. rochei* are also found in Malaysia.

Of the three members of the genus *Euthynnus*, only *E. affinis* has been recorded in Malaysia (Plate 2.3). The genus *Katsuwonus* is a monotypic genus that is closely related to *Euthynnus* and *Thunnus* (Collette *et al.*, 2001). Its sole member, *Katsuwonus pelamis* also inhabits Malaysian waters (Plate 2.4). The genus *Thunnus* that is considered to be the most advanced genus in the family Scombridae, consists of eight species, of which four can be found locally in Malaysia; *T. albacares*, *T. obesus* (Plate 2.5), *T. alalunga* and *T. tonggol.*

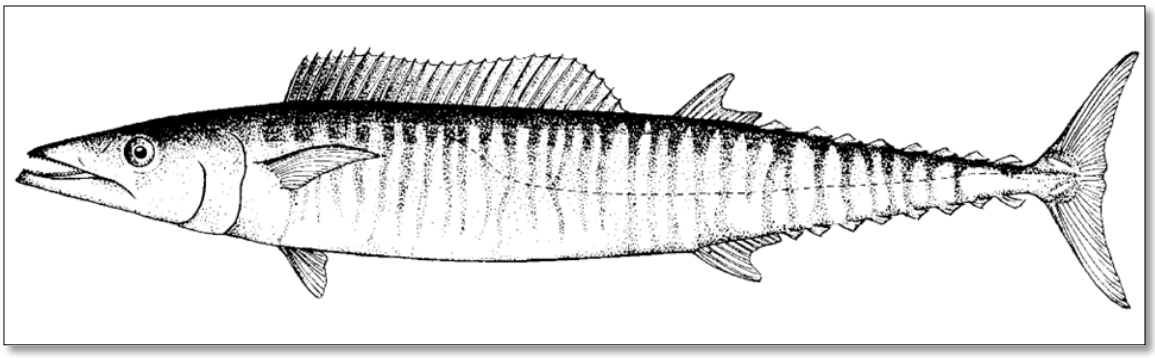

Figure 2.2: *Acanthocybium solandri* (from Collete & Nauen, 1983).

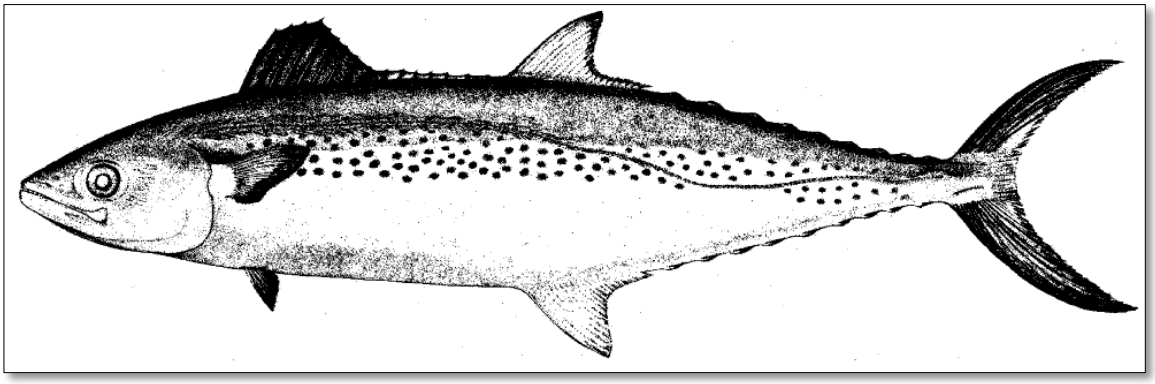

Figure 2.3: *Scomberomorus guttatus* (from Collete & Nauen, 1983)

Plate 2.1: *Grammatorcynus bilineatus*

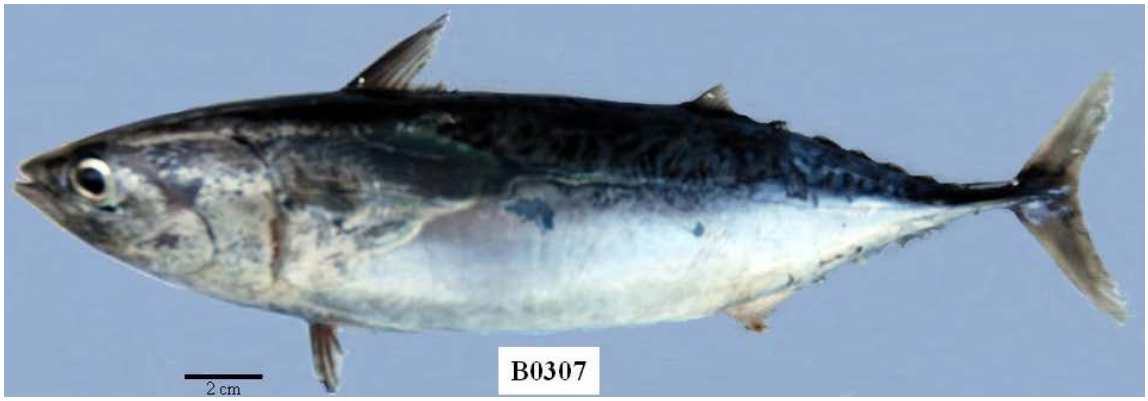

Plate 2.2: *Auxis thazard*

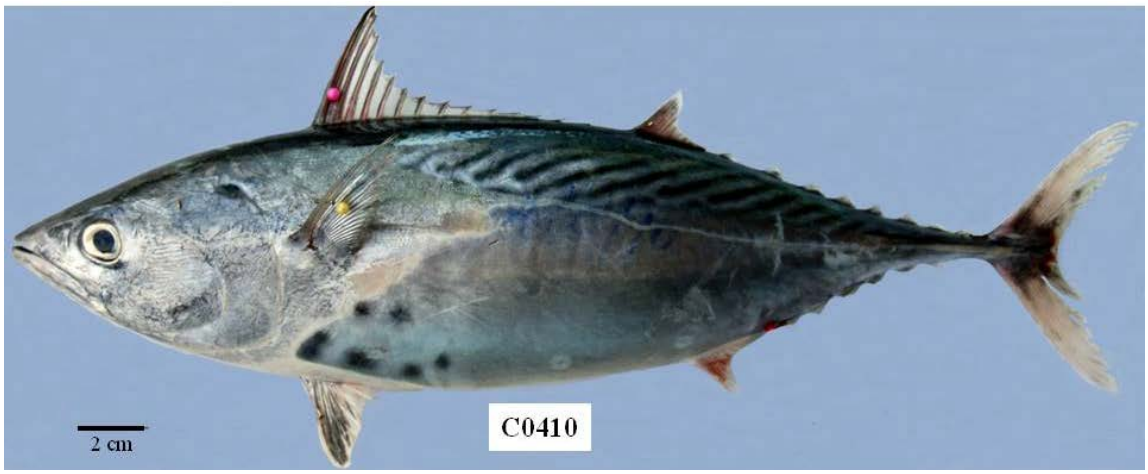

Plate 2.3: *Euthynnus affinis*

Plate 2.4: *Katsuwonus pelamis*

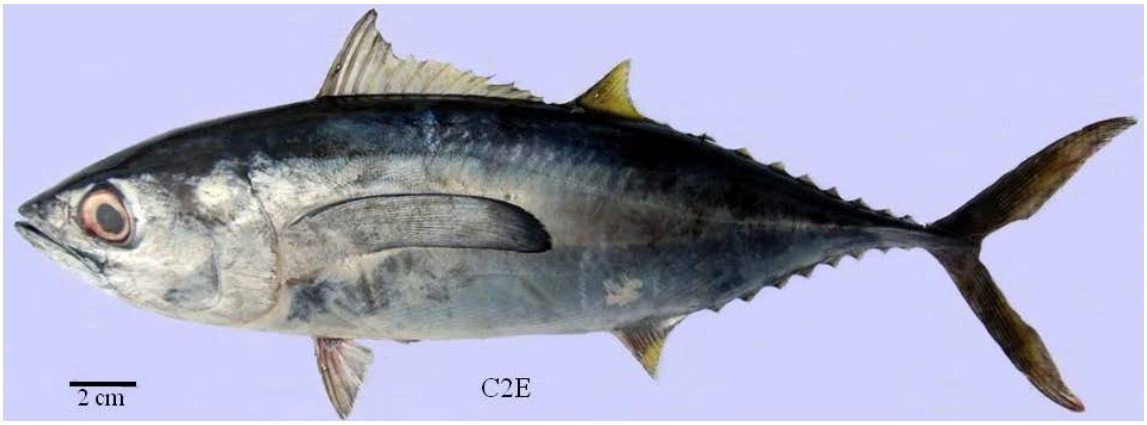

Plate 2.5: *Thunnus obesus*

2.2.1.3 Tribe Sardini

There are four genera in this tribe and of this, two can be found in Malaysia, that belong to *Gymnosarda* Gill, 1862 and *Sarda* Cuvier, 1829. Only a single species is represented in the genus *Gymnosarda, G. unicolor* and this species is found in Malaysia (Plate 2.6). Five species are found in the genus *Sarda* (Collete *et al.*, 2001) of which only *Sarda orientalis* (Plate 2.7) has been recorded in Malaysia.

2.2.1.4 Tribe Scombrini

Two genera represent this tribe; *Rastrelliger* Jordan and Starks, 1908 and *Scomber* Linnaeus, 1758. Morphogically these two genera possess an elongated body that is rounded and slightly compressed. All three species included in the genus *Rastrelliger* are found in Malaysia, namely *R. brachysoma, R. faughni* (Plate 2.8) and lastly *R. kanagurta* which is the focus of the current study. In contrast, within the genus *Scomber* that consists of four species only *S. australasicus* (Figure 2.4) occurs in Malaysian waters.

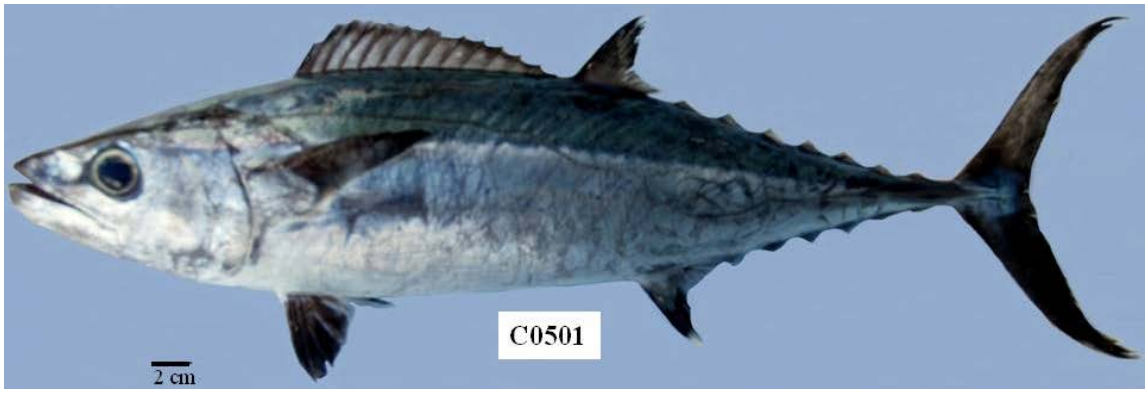

Plate 2.6: *Gymnosarda unicolor*

Plate 2.7: *Sarda orientalis*

Plate 2.8: *Rastrelliger faughni*

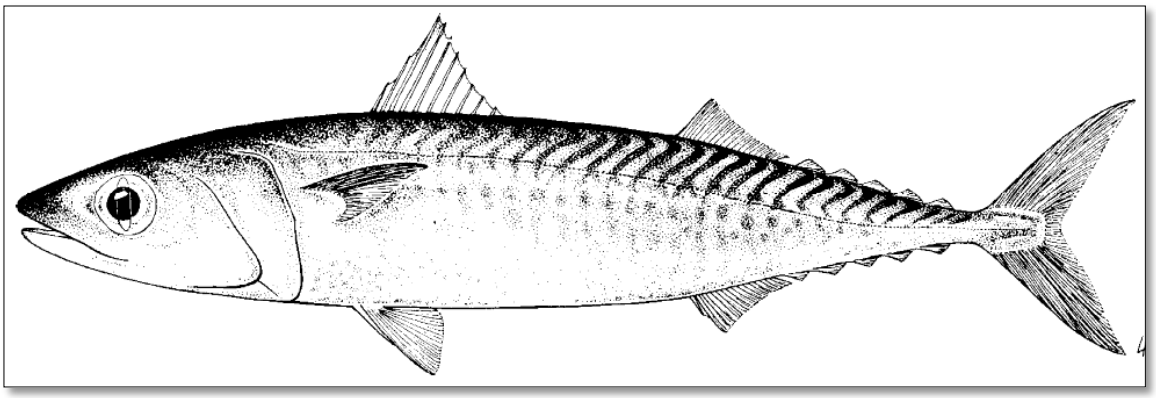

Figure 2.4: *Scomber australasicus* (from Collette & Nauen, 1983)

2.2.2 Habitat and Distributions

The ubiquitous family Scombridae is widely distributed from temperate to tropical waters. Individuals can be found in coastal waters occurring from the surface to a depth of 250 – 350 m (e.g *Scomber japonicus*) and in oceanic waters from epi-pelagic to midwater (the tunas) to a depth of more than 500m (Collette & Nauen, 1983). Members of genus *Thunnus* belonging to the advanced tribe, Thunnini can be found in both tropical and temperate waters. The temperate *T. orientalis* (Pacific bluefin tuna) occurs in the northern Pacific, *T. maccoyii* (Southern bluefin tuna) occurs in the Southern ocean and *T. atlanticus* (Blackfin tuna) in the western Atlantic. Tropical species include *T. albacares* that occurs in tropical and subtropical waters and *T. tonggol* that is found in the Indo-West Pacific. Several species are also known to inhabit both tropical and temperate waters including the albacore, *T. alalunga*.

The genus *Scomber* in the tribe Scombrini, is epipelagic and neritic with individuals occurring in temperate to subtropical waters, a group that is replaced by the genus *Rastrelliger* that are restricted to tropical waters in the Indo-West Pacific region (Collette & Nauen, 1983). Although *S. japonicus, S. australasicus* and *S. scombrus* are found widely in temperate to subtropical waters, *S. japonicus* and *S.*

australasicus display antitropical distributions (Scoles *et al.*, 1998) and as such, are not found in tropical waters.

2.2.3 Molecular Taxonomy of the Family Scombridae

Effective management of sustainable fisheries require the underlying taxonomy of target species to be well-defined, thus allowing for catch statistics to be accurately monitored. Molecular taxonomic and phylogenetic studies of the Family Scombridae were first initiated in the early 1990's. Early investigations were mainly conducted on the genus *Thunnus* because of the commercial importance of these fishes with the primary focus on stock delineation and species identification (Elliott & Ward, 1995). These early studies employed mtDNA markers for examples the cytochrome *b,* ATPase and cytochrome oxidase subunit I (COI) genes (Bartlett & Davidson, 1991; Block *et al.*, 1993; Chow & Kishino, 1995; Elliott & Ward, 1995; Ward *et al.*, 2005). Precise species identification is crucial for detecting illegal fishing and trading, especially with regard to northern and southern bluefin tunas, but may be difficult when diagnostic external and internal morphological characters are removed when fish are filleted (Takeyama *et al.*, 2001). The eight nominal species included in the genus *Thunnus* (*T. alalunga, T. albacares, T. atlanticus, T. maccoyii, T. obesus, T. thynnus* and *T. tonggol*) have been clearly identified using mtDNA cytochrome *b* and ATPase gene sequences (Bartlett & Davidson, 1991; Finnerty & Block, 1995; Chow & Kishino, 1995) and recently the eight species were also validated successfully based on the COI, also referred to as the barcoding approach by Ward *et al.* (2005). Identification of different life history stages including; egg, larvae and small juvenile is also important for clarifying species distributions and their reproductive activities (Takeyama *et al.*, 2001) as reported in Scombrid larval

studies by Chow *et al.,* (2003) (cytochrome *b* gene and flanking region between ATPase and CO III genes) and Ko *et al.,* (2013) who employed DNA barcoding.

While the genus *Scomber* is also commercially important, there have been a number of taxonomic ambiguities associated with *S. colias, S. scombrus, S. australasicus* and *S. japonicus* based solely on external morphological characters (Matsui, 1967). The monophyly of this genus however, with regard to other scombrid fish and the taxonomic status of *S. colias* and *S. japonicus* specifically have been resolved via phylogenetic analyses of their complete mitogenomes (Catanese *et al.,* 2010) in addition to Cyt *b*, COI, Control region and 5S rDNA (Cheng *et al,* 2011).

2.2.4 Economic Importance of Scombrid Species

According to an FAO report in 2010, the total global catch of tuna and tuna-like species approached 6.3 million tones. The principal market tuna species – albacore, bigeye, bluefin (three species), skipjack and yellowfin – together contributed 4.2 million tonnes, a level that had been declined approximately 0.2 million tonnes from a peak in 2005. About 70 % of this catch occurred in the Pacific. In Malaysia, in 2008, a total of 3,437 tonnes of oceanic tuna valued at RM28.84 million were taken, an increase of 25.83% from 2007. The species landed included yellow fin tuna, big eye tuna and albacore. By 2009 however oceanic tuna landings in Penang, Malaysia had showed a 33.60% declined as compared with 2008 (Department of Fisheries Malaysia, 2009).

Recognising the importance of this fishery and the various threats to it, global and intergovernmental efforts have been initiated to better manage the resource. One such effort is organized by the Indian Ocean Tuna Commission (Indian Ocean Tuna

Commision, 2009) to ensure the conservation and sustainable utilisation of wild tuna stocks. The IOTC mandate is to manage tuna and tuna-like species in the Indian Ocean and adjacent seas. Its objective is to promote cooperation among its member countries and encourage sustainable development of fisheries via appropriate management. Malaysia along with countries bordering the Indian Ocean is a member of this organization.

2.3 The Biology of *Rastrelliger kanagurta*

2.3.1 Nomenclature and Taxonomy

The focus of the current study is the Indian mackerel, *Rastrelliger kanagurta*. Locally this species is known as 'kembung', 'kembung borek', 'mabong' and 'rumahan' (Froese & Pauly, 2013). The taxonomic classification of *R. kanagurta* according to the Integrated Taxonomy Information System (ITIS) is as follows:

Kingdom	Animalia
Phylum	Chordata
Subphylum	Vertebrata
Superclass	Osteichthyes
Class	Actinopterygii
Subclass	Neopterygii
Infraclass	Teleostei
Superorder	Acanthopterygii
Order	Perciformes
Suborder	Scombroidei
Family	Scombridae
Genus	*Rastrelliger*
Species	*Rastrelliger kanagurta* (Cuvier, 1816)

2.3.2 Diagnostic Morphological Characters of *Rastrelliger kanagurta*

Morphologically, the body is moderately deep, with the head longer than the body depth. The species has long gill rakers that are visible when the mouth is open. It has narrow dark longitudinal bands on the upper part of the body (golden in fresh

specimens and a black spot on the body near the lower margin of the pectoral fin, dorsal fins that are yellowish with black tips, caudal and pectoral fins yellowish while other fins are dusky (Plate 2.9) as described by Collette & Nauen (1983). Species of the genus *Scomber* are often mistakenly identified as *Rastrelliger* due to their similar morphological appearance. Each possess an adipose eye-lid, poorly developed corselet and with only two small keels on each side of the caudal peduncle.

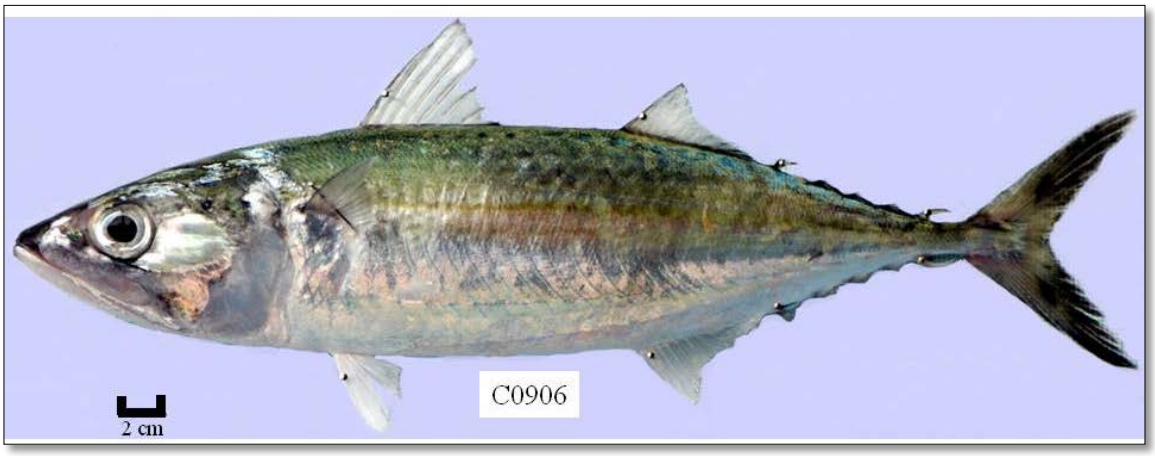

Plate 2.9: *Rastrelliger kanagurta* in lateral view

2.3.3 Habitat and Distribution

Rastrelliger kanagurta is widely distributed across the tropical Indo-West Pacific region, approximately from longitudes 300° E to 1600° W and latitudes 300°S to 300°N. The species has been recorded from almost the entire east coast of Africa, from Madagascar, Mauritius, Reunion Islands, Seychelles, the countries bordering the Red Sea and the Persian Gulf, from the coasts of Pakistan, India, Ceylon, Burma, Thailand, Malaysia, Cambodia, Indonesia, northern Australia, New Guinea, the Micronesian, Melanesian, Polynesian and Solomon Islands, the New Hebrides, Fiji and Samoa Islands, the Philippine Islands, People's Republic of China and Hong Kong, Taiwan and Ryukyu Islands and some of the central group of

Pacific Islands including those of Hawaii (Figure 2.5) (Collette & Nauen, 1983). It is believed to have entered the eastern Mediterranean Sea through the Suez Canal (Froese & Pauly, 2013**).**

2.3.4 Economic Importance

Rastrelliger kanagurta and its two congeners, *R. brachysoma*, *R. faughni* are three of the most important commercial fish species in Malaysia. According to Chee (2000), *Rastrelliger* constitute the most dominant marine food fish on the west coast of Peninsular Malaysia. In 2007, the Indian Mackerel (*Rastrelliger* spp.) was the dominant fish harvested with ~ 156 tonnes caught by purse seiners and drift/gill netters in Malaysia (FAO, 2013).

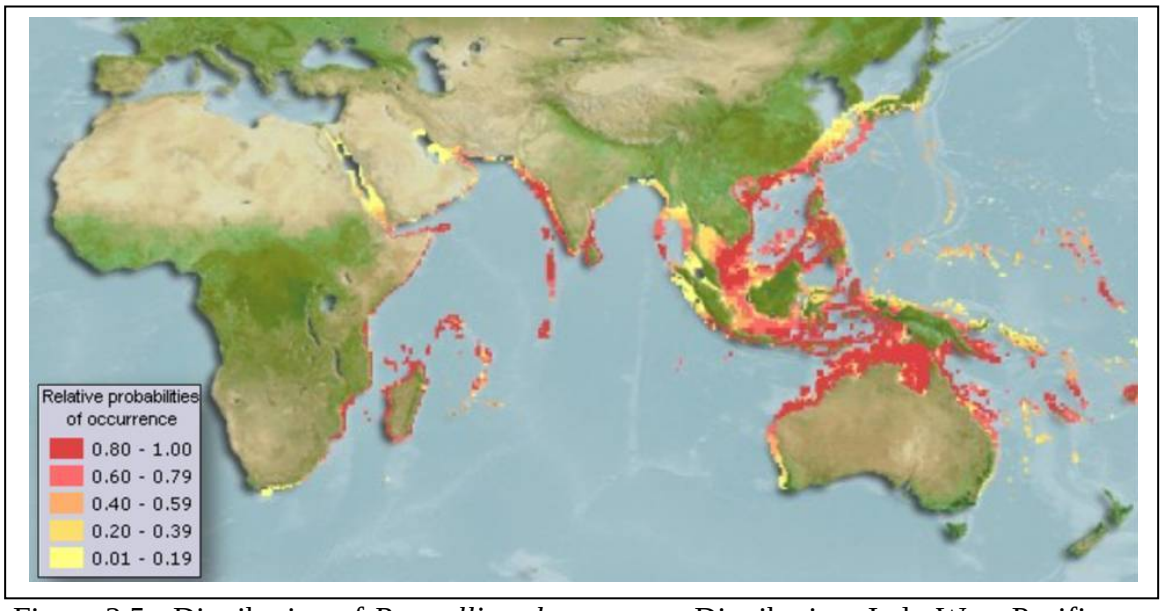

Figure 2.5: Distribution of *Rastrelliger kanagurta* : Distribution: Indo-West Pacific: Red Sea and East Africa to Indonesia, north to the Ryukyu Islands and China, south to Australia, Melanesia and Samoa. Retrieved from http://www.aquamaps.org/receive.php (29 May 2012).

Total *Rastrelliger* landings have increased 45.5% since 1995 from 101,003 tonnes (Chee, 2000) to 185,463 tonnes in 2009 (Department of Fisheries Malaysia, 2009). This species also contributed approximately 13.1% of the total marine resources caught in 2009, with *R. kanagurta* landings contributing a substantial quantity to the total landings i.e. about 4.1% (56,250 tonnes), while *R. brachysoma* was higher at 9.2% (128, 970 tonnes). *Rastrelliger kanagurta* captures increased slighly in 2010 (186, 225 tonnes) accounting for 10% of total fish production in Malaysia (Department of Fisheries Malaysia, 2010). Although landings of this species remains robust, there is considerable concern about the fishery which is considered to be over-fished across much of this region (BOBLME, 2011).

Rastrelliger kanagurta is one of the key species in the Bay of Bengal Large Marine Ecosystem (BOBLME) project aimed at ensuring best fishery practices that makes recommendations to ensure sustainable fisheries of this species in BOB nations. The BOB is an area of the Indian Ocean, between India in the west and the Malay Peninsula in the east, covering approximate area of 2,090 km long and 1,600 km wide. The specific aim of this project is to provide strategies for development of regional Fisheries Management Plan (FMPs).

2.4. DNA Barcoding

DNA barcoding has become an increasingly important taxonomic tool for species identification. The approach is based on a short section from a standardized region of the genome, typically the mitochondrial cytochrome oxidase subunit I gene is targetted (with some variants). The approach was initiated by Hebert *et al.* (2003a, 2003b) who used a system that employs a DNA sequence as a taxon ‘barcode’ with the standard barcode of 648 bp at the 5’ end of the COI gene to identify all forms of organisms globally. This system also maintains all morphological information

associated with whole specimens (Tautz *et al.*, 2003). While, there have been several debates however about the relative utility of traditional taxonomy vs DNA barcoding for species discrimination (Hebert & Gregory, 2005), in general barcoding is now widely accepted as a useful complement to conventional taxonomy.

Given that taxonomic expertise is limited around the globe and millions of species on either remain to be described or required appraisal, DNA barcoding through genomic approach could potentially facilitate identification of the vast biodiversity that exists. While morphology will remain the cornerstone of taxonomy, Hebert *et al.* (2003a) highlighted four limitations; 1) both phenotypic plasticity and genetic variability in the characters employed for species recognition which can lead to incorrect identifications; 2) can overlook cryptic taxa that are common in many groups; 3) morphological keys are often effective only for a particular life stage or gender; and 4) the use of keys often demands a very high level of expertise so that misdiagnoses are common.

According to Kress & Erickson (2008), the gene region selected for use as a DNA barcode must satisfy three criteria; it should contain significant species-level genetic variation and divergence, possess conserved flanking sites and consist of only a short sequence length. In animals, the mtDNA cytochrome oxidase subunit I (COI) has been accepted as the barcoding gene (Hebert *et al.*, 2003a), while for plant, a two locus barcode (*rbcl* and *matK*) is the accepted DNA barcode (Hollingsworth *et al.*, 2011; Cowan & Fay, 2012)

Despite possessing conserved amino acid sequences, the COI has one of the fastest mean rates of nucleotide substitution and the greatest variation in rates in animal species (Mueller, 2006). Hebert *et al.* (2003a) pointed out that there are two

advantages of using the mitochondrial COI as a barcoding gene; (i) the universal primers are very robust and possess a greater range of phylogenetic signal than any other mitochondrial gene, and (ii) its high evolutionary rate allows the discrimination of closely allied species and phylogeographic inferences within a single species.

The COI has been used successfully as a systematic marker in many animal taxa; for example - discrimination among species in Lepidoptera families (Hajibabaei *et al.*, 2005; Wilson, 2010), marine invertebrates in the Antarctic (Grant & Linse, 2009); Arctic springtails – Hexapoda: Collembola (Hogg & Hebert, 2004); Birds (Hebert *et al*, 2004); Golden Silkmoth – Saturniidae (Suriana *et al.*, 2012); marine and freshwater fishes (Ward *et al.*, 2005; Steinke *et al.*, 2009).

Development of Next Generation Sequencing (NGS) platforms has seen further refinement of the barcoding method. An example is the mini-barcode fragments (shorter fragment between ~ 100bp cf 650bp in standard barcodes) of the mtDNA COI that can be readily obtained and are robust after development on their 454 pyrosequencing platform (Hajibabaei *et al.*, 2011). Its resolution efficiency is comparable (at 90% species resolution) (Hajibabaei *et al.*, 2005; Hajibabaei *et al.*, 2007; Meusnier *et al.*, 2008) with the 97% species resolution using full length DNA barcode sequences (~650 bp) (Meusnier *et al.*, 2008). Furthermore, the mini-barcode is particularly useful for archival specimens and for processed biological materials such as canned food which generally do not have the full sequence DNA barcode intact (Hajibabaei *et al.*, 2006; Meusnier *et al.*, 2008).