UNIVERSITI SAINS MALAYSIA

UNIVERSITI SAINS MALAYSIA

Molecular characterization of *Selar crumenophthalmus* (ikan Lolong), *Decapterus maruadsi* (ikan Sadin) and *Rastrelliger kanagurta* (ikan Kembong hitam) fish species using mtDNA COI gene (DNA barcoding)

Dissertation submitted in partial fulfillment for the Degree of Bachelor of Science in Forensic Science

Loo Suet Kee

School of Health Sciences
Universiti Sains Malaysia
16150 Kubang Kerian Kelantan
Malaysia

2011

ACKNOWLEDGEMENTS

I would like to express my greatest gratitude to both my project supervisor Mr. S. Panneerchelvam and also my project co-supervisor, Prof. Dr. Norazmi Mohd. Nor for their continuous guidance and support throughout the whole project. I would also like to thank Mdm. Nor Haslindawaty, Mdm. Hafizah Harun, Mdm. Rosniah Yusoff and Ms. Nor Amila Amri for their kind assistance throughout the project. My family has been my biggest motivator through the completion of this project. Thank you to my parents Mr. Loo Seng Yee and Mdm. Liang Siew Ying and my siblings as well for providing all forms of support throughout the completion of this project. Lastly, my thanks also express to my colleague Mr. Che Alhafizol Che Ismail for his cooperation in the project.

TABLE OF CONTENTS

No.	Content	Page no.
1.	Abstract	1
2.	Introduction	2 - 5
3.	Literature review	6 - 9
4.	Objectives	10
5.	Materials and methods	11 - 25
6.	Result	26 - 47
7.	Discussion	48 - 51
8.	Limitation	52
9.	Future work	52
10.	Conclusion	52
11.	References	53 - 60
12.	Appendices	Appendix 1 and Appendix 2

LIST OF TABLES AND FIGURES

Table

No.	Captions	Page no.
1.	The sequences and characteristics of ST F1 (forward primer) and ST R3 (reverse primer).	19
2.	Compositions of PCR master mix.	20
3.	PCR amplification parameters.	21
4.	Components of termination reaction mixture together with their respective volumes.	23
5.	PCR parameters for cycle termination process.	24
6.	Nucleotide compositions for the three species of fish for the partial mtDNA COI gene. A: Adenine; C: Cytosine; G: Guanine; T: Thymine.	36
7.	Percentage of similarities between species based on pair wise analysis using BioEdit sequence alignment editor.	39
8.	Multiple alignment shows the nucleotide variations present in the 3 species namely *Selar crumenophthalmus, Decapterus maruadsi,* and *Rastrelliger kanagurta* compared with *Gobio*	40

gobio as reference (GenBank accession no.: AB239596).

9. Polymorphisms present in *Selar crumenophthalmus, Decapterus maruadsi,* and *Rastrelliger kanagurta* with *Gobio gobio* as reference species. 42

10. Polymorphisms present among *Selar crumenophthalmus* and *Decapterus maruadsi.* 43

11. Polymorphisms present among *Selar crumenophthalmus* and *Rastrelliger kanagurta.* 44

12. Polymorphisms present among *Decapterus maruadsi* and *Rastrelliger kanagurta.* 45

Figure

No.	Caption	Page no.
1.	Electropherogram showing extracted genomic DNA. Lane M: 100bp DNA ladder; lane 1: *Selar crumenophthalmus*; lane 2: *Decapterus maruadsi*; lane 3: *Rastrelliger kanagurta.*	26
2.	Electropherogram showing the result of screening of PCR amplified products of 3 fish species, successfully amplified using the ST F1forward and ST R3 reverse primers. Lane M: 100bp DNA ladder; lane 1: *Selar crumenophthalmus*; lane 2: *Decapterus maruadsi*; lane 3: *Rastrelliger kanagurta.*	27
3.	Electopherogram showing the result of screening of purified PCR products for three species of fish. Lane M: 100bp DNA ladder; lane 1: *Selar crumenophthalmus*; lane 2: *Decapterus maruadsi*; lane 3: *Rastrelliger kanagurta.*	28
4.	Chromatogram showing the forward sequence of the amplicon of *Selar crumenophthalmus*.	29
5.	Chromatogram showing the reverse sequence of the amplicon of *Selar crumenophthalmus*.	30
6.	Chromatogram showing the forward sequence of the amplicon of *Decapterus maruadsi*.	31

7.	Chromatogram showing the reverse sequence of the amplicon of *Decapterus maruadsi*	32
8.	Chromatogram showing the forward sequence of the amplicon of *Rastrelliger kanagurta*.	33
9.	Chromatogram showing the reverse sequence of the amplicon of *Rastrelliger kanagurta*.	34
10.	Chart showing the relative nucleotide composition of the 645bp length partial mtDNA COI gene of *Selar crumenophthalmus*.	37
11.	Chart showing the relative nucleotide composition of the 557bp length partial mtDNA COI gene of *Decapterus maruadsi*.	37
12.	Chart showing the relative nucleotide composition of the 586bp length partial mtDNA COI gene of *Rastrelliger kanagurta*.	38
13.	Summary representing results shown in table 9, 10, 11, and 12. Group A: Polymorphisms present in *Selar crumenophthalmus, Decapterus maruadsi,* and *Rastrelliger kanagurta* with *Gobio gobio* as reference species; group B: Polymorphisms present among *Selar crumenophthalmus* and *Decapterus maruadsi;* group C: Polymorphisms present among *Selar crumenophthalmus* and *Rastrelliger kanagurta;* group D: Polymorphisms present among *Decapterus maruadsi* and	46

Rastrelliger kanagurta.

14. Phylogenetic tree of the eight species namely *Decapterus maraudsi, Rastrelliger kanagurta, Atule mate, Scomberomorus maculatus, Selar crumenophthalmus, Decapteus russelli, Gobio gobio* and *Amphioxus.* 47

ABSTRACT

This study was undertaken to design a set of common primers to amplify the mitochondria DNA (mtDNA) cytochrome c oxidase subunit I (COI) barcoding region of marine fish species. The primer set namely: forward primer (ST F1) "CAC AAA GAC ATC GGC ACCC" and the reverse primer (ST R3) "AAG AAT CAA/G AAT AG/AG TGT TG" were able to amplify the three species chosen for this study. The sequence data generated were compared and analysed using the sequence of mtDNA COI gene of *Gobio gobio* (GenBank accession no.: AB239596) as reference. On analysis, about 20% polymorphic positional variations were detected across the three species. The sequence data generated were used to study their phylogenetic relationship by constructing a rooted neighbor joining (NJ) dendrogram.

INTRODUCTION

DNA barcoding is the use of a short sequence of DNA at a standard location that enables species identification, recognition, and discovery in a particular domain of life (Hebert *et al.*, 2003; Bucklin *et al.*, 2010). Recent explosion of DNA barcoding interest is credited to Paul D. N. Hebert who proposed the compilation of public library of DNA barcodes that contains DNA barcodes linked to corresponding species, together with other relevant information (Bucklin *et al.*, 2010).

Ever since the introduction of Linnaeus system of species classification, the methodology of species identification has not undergone any major transformation. This was true until the introduction of DNA barcoding in the year of 2003 by Paul D. N. Hebert. The traditional Linnaeus system strongly depends on morphological characters. DNA barcoding employs sequence diversity in short, standardized gene regions to generate 'species barcode' for the purpose of species identification and discovery (Ratnasingham and Hebert, 2007).

Morphological diagnosis demands a high level of expertise especially in diagnosing closely resembling species and therefore misidentifications are common (Hebert *et al.*, 2003). It is estimated that 10 to 15 million living species present on the earth. Only a few taxonomists can critically identify more than 0.01% of them based on morphological diagnosis (Hammond 1992; Hawksworth and Kalin-Arroyo, 1995). If morphological taxonomy is to be sustained, 15,000 taxonomists will be required in perpetuity to identify all the species on earth (Hebert *et al.*, 2003). However, the reality is that after centuries of acquiring knowledge, today, morphological taxonomy had started to lose popularity to other fields and results in a global shortage of taxonomists (Radulovici *et al.*, 2010).

When morphology is compromised, genetic species identification provides a good means of complimentary for it. Genetic species identification by means of DNA barcodes uses 4 alternate nucleotides at each position of the chosen gene to generate 'species barcode' for species identification. Taking into consideration of just 15 nucleotides positions, a total number of 4^{15} (1 billion) possible codes can be created. This number is 100 times the number that is needed to discriminate life if each taxon is uniquely branded (Hebert *et al.*, 2003).

A 648 base pairs (bp) region of the mitochondria DNA cytochrome c oxidase subunit I (mtDNA COI) gene has proved to be useful for the purpose of DNA barcoding (Hebert *et al.*, 2003; Dawnay *et al.*, 2007). A typical mtDNA is supercoiled, double stranded, circular molecule found within the cytoplasm of mitochondrion. It has approximately 15,000 to 19,000 bp across various species. It consists of two parts: the control region that does not code for any gene products and also the coding region which codes for the gene products. Most of the mtDNA genome codes for 37 gene products used in the oxidative phosphorylation process (Butler, 2010).

With a few exceptions, an identical sets of genes present in the mitochondria of animal: 13 protein-coding gene, 2 ribosomal RNA, and 22 transfer RNA genes. The order of these genes as well as their polarity varies markedly between different animal phyla. In order to compare sequences from diverse organism, the barcode locality must be limited to one gene (Stoeckle *et al.*, 2005). It also eases the recovery of DNA barcode from diverse organisms. One potential gene for the purpose of DNA barcoding will be the protein coding COI gene. The total length of COI gene varies between species and is approximately 1500 bp in length (Vilaca *et al.*, 2006). The COI barcode region is located at the 5' end of the gene with an

approximate length of 648 bp, flanked by conserved regions which are the usual basis for design of conserved PCR (polymerase chain reaction) primers (Bucklin *et al.*, 2010).

DNA barcoding identifies a questioned specimen by matching the DNA barcode of the specimen to a reference sequence in barcode library. Using suitable primer sets, the DNA barcoding region of fresh specimen or voucher specimen identified by means of morphological taxonomy will be amplified. Barcoding protocol includes standardized procedure of extraction of DNA, PCR amplification, sequencing and storage of data. DNA barcodes generated from unknown specimen will then be compared to the DNA barcodes stored in the barcode library. Any match between the DNA barcodes of the unknown specimen and named specimen in the library will leads to identification of the unknown specimen. DNA barcoding functions to complement morphological taxonomy but not replacing it (DeSalle *et al.*, 2005). The named specimen in the barcode library would have to be generated in advance by means of morphological taxonomy before DNA barcoding can be used for species identification.

Standardization of species identification method will enable comparison between datasets generated by different researchers, revealing cases of synonymy, potential cryptic species or genetically distinct population (Radulovici *et al.*, 2010). Apart from that, standardization will also aid the construction of a comprehensive and consistent barcode library. Once a complete database is available, it will enable accurate, quick and cost effective species identification. Example of such library is the developing BoLD (The barcode of life data system) which assembles molecular, morphological and distributional data of many species on earth (Ratnasingham and Hebert, 2007).

DNA barcoding requires only a minimum of 10 ng to 20 ng of template DNA from specimen. So, it is particularly useful when dealing with forensic specimen which is usually present in small amount and morphologically distorted. Apart from identifying whole adult organism which is morphologically distinct, DNA barcoding also works with fragments of specimen and works for all stages of life (Bucklin *et al.*, 2003; Stoeckle *et al.*, 2005).

Incidentally, DNA barcodes can also be used for the study of phylogenetic relationship among different species. According to Butler (2010), phylogenetics is a method of grouping genetic information together based on similarities in order to infer descent from a common ancestor; this study of genetic diversity enables construction of evolutionary trees suggesting evolutionary distance and relationships among species. In short, DNA barcode library will also help evolutionary study.

Fishery is an important field to be explored on DNA barcoding. According to Food and agriculture organization of the United Nations, FAO (2002), in year 2000 alone, fisheries provided more than 15% of total animal protein to the global food supply. This shows that fish and fish products are important contributors to human food security (Ward *et al.*, 2005). Accurate and unambiguous identification of fish and fish products, from eggs to adults is crucial in addressing issues such as market substitution, quota management and products from regulated species (Bucklin *et al.*, 2010). In year 2007, there are an estimated of 15,700 marine and 13,700 freshwater fish species present on earth (Ivanova *et al.*, 2007). To date, there are only an estimated of 7,437 fish species, most of which are marine that had been barcoded (Bucklin *et al.*, 2010). These data indicate that there are still much works to be done on DNA barcoding of fish.

LITERATURE REVIEW

Considering the many shortcomings of morphological taxonomy, scientists took the opportunity provided by the development of molecular methods to clarify many uncertainties in morphological taxonomy (Radulovici, 2010). Allozymes as the first molecular markers were extensively used in population genetics (Avise, 1975). Medeiros-Bergen *et al.* (1995) successfully identified 3 holothurian species with 16S mitochondria gene sequence. Heist and Gold (1999) identified species by using restriction fragment length polymorphism (RFLPs) of sections of the mitochondria genes cytochrome b (cyt b). Hoelzel (2001) then designed species specific primers in the cyt b gene for *Cetorhinus maximus* and examined short segments of mitochondria $NADH_2$ and cyt b gene for several Lamniform species.

Despite the great potential of genetics in assisting species identification, there has been little consensus on which gene region would be the most suitable (Holmes *et al.*, 2009). This situation persisted until Hebert *et al.* (2003) promoted DNA barcodes as the global standard for species identification of animals. They suggested a 648 bp region of the mtDNA COI gene as the standard region for DNA barcoding. mtDNA has several advantages over nuclear DNA as the target of analysis which include its lack of introns, has limited exposure to recombination, haploid mode of inheritance (Saccone *et al.*, 1999), high copy numbers in every cell, and a generally strict maternal mode of inheritance (Bucklin *et al.*, 2010).

Past phylogeny work has focused on mitochondria genes encoding ribosomal DNA (12S rRNA, 16S rRNA). However, they share a disadvantage in the prevalence of insertions and deletions (indels) that greatly complicate sequence alignments (Doyle and Gaut, 2000). The 13 protein-coding genes in the mitochondria genome serve a better target as indels are rare

since most lead to a shift in the reading frame (Hebert *et al.*, 2003). mtDNA COI gene is chosen despite the many protein-coding genes in the mitochondria genome due to two advantages: first, the universal primers for this gene is very robust enabling recovery of its 5' end for species from a broad range of animal phyla (Folmer *et al.*, 1994). Second, COI, compared with other mitochondria genes, appears to possess a greater range of phylogenetic signal. Similar to other protein coding genes, high incident of base substitutions is shown in its third-position nucleotides, leading to a rate of molecular evolution that is nearly three times greater than that of 12S or 16S mitochondria rRNA genes (Knowlton and Weigt, 1998).

Furthermore, the mtDNA COI gene evolution is sufficiently rapid to allow the discrimination of not only closely allied species, but also phylogeographic groups within a single species (Cox and Hebert, 2001; Wares and Cunningham, 2001). For instance, mtDNA COI gene is more likely to provide deeper phylogenetics insights than alternatives such as cyt b gene (Simmons and Weller, 2001). There are two particularly important findings from the past research on COI divergence: first, intraspecific divergences is usually lower than 2% and most likely less than 1% (Avise, 2000); and second, instances of higher divergences usually occur as geographical isolates of variants, reflecting their origin in past episodes of gene pool fragmentation. Moreover, most of the high divergences involve cases of taxonomic uncertainty in which lineages share a species epithet, but their actual status is unclear (Avise and Walker, 1999).

In sum, the above stated phylogeographic studies have shown that intraspecific divergences are ordinarily well below those that separate congeneric species pairs. Hence, with few exceptions, COI divergences is an effective tool in species recognition (Hebert *et al.*, 2003). These few exceptions include the phyla Porifera, Ctenophora, Cnidarians class Anthozoa

(Bucklin *et al.*, 2010), and also the plant Kingdom (Palmer, 1992). In these phyla and kingdom, the evolutionary rate of mtDNA was found to be too low to allow reliable discrimination between closely related species (Bucklin *et al.*, 2010).

Series of validation studies have also been conducted on mtDNA COI gene to determine its reproducibility and its limitation for accuracy in result reporting under a variety of conditions. Validation experiments include examination of reproducibility and heteroplasmy, mixed DNA, DNA template concentration, chemical treatments, substrate variation, environmental conditions and thermocycling parameters. The results of research studies show that species-diagnostic COI sequences were successfully obtained in all validation experiments (Dawnay *et al.*, 2007). It indicates that COI gene enables accurate animal species identification provided that adequate reference sequence data exists and it is suitable for use in forensic species determination (Dawnay *et al.*, 2007).

DNA barcoding has its limitations as well. In some cases of hybrids, recently diverged species, species complexes or slow evolving group, DNA barcoding shows low resolution (Radulovici *et al.*, 2010). DNA barcoding also fails to reveal the presence of pseudogenes (Song *et al.*, 2008), contaminants amplified with primers (Siddall *et al.*, 2009), cases of mitochondria introgression (Kemppainen *et al.*, 2009), and also fails to identify functional group of many organism (Radulovici *et al.*, 2010). One technical disadvantage of DNA barcoding at present is that it requires access to sequencing technology. However, sequencing process is becoming faster, more commonplace and increasingly less expensive (Holmes *et al.*, 2009).

Fishes currently represent the most comprehensive group of marine metazoans been researched upon (Bucklin *et al.*, 2010). One of the early DNA barcode study on fish was done

by Ward *et al.* (2005). They examined a number of taxa comprising 207 species of fish species from Australia and had shown that all could be discriminated based on their COI sequence. DNA barcoding has also been used to identify shark fins confiscated from illegal fishers in northern Australian waters (Holmes *et al.*, 2009). DNA barcoding has also proved to be successful in identifying fish larvae taken from the Great Barrier Reef (Pegg *et al.*, 2006). Research in DNA barcoding of fish also gained significance by the launching of Fish barcode of life (FISH-BOL) in 2005. FISH-BOL is gathering DNA barcode records for all 30,000 known species of fishes (Bucklin *et al.*, 2010).

Regarding DNA barcoding other than fish, Hebert *et al.* (2003) had successfully identified all the lepidopteran species involved in their study. DNA barcodes also able to identify all the 191 species of echinoderms in the study done by Ward *et al.* (2008), confirming the usefulness of sequence variation for species discrimination across the phylum. DNA barcoding was found to be useful for conservation and wildlife forensics. It helped the identification of turtle meats and eggs illegally traded or carcasses on beaches (Vargas *et al.*, 2009).

Other examples of DNA barcoding in forensics and seafood safety include that DNA barcoding can be used on cooked or processed seafood. Smith *et al.* (2008) succeeded in identifying species of origin for smoked fish fillets of several genera. Recently, Barbuto *et al.* (2010) had shown that at least 25% of the seafood sampled at commercial markets and restaurants was mislabeled. Lastly, the Food and Drug Administration in the United States is also considering the replacement of protein isoelectric focusing technique for fish and fish product identification with DNA barcoding (Cohen *et al.* 2009).

OBJECTIVES

1. To understand and apply the concept of DNA barcoding on fish species characterization.

2. To understand and apply techniques used in DNA barcoding.

3. To compile DNA barcoding data for the selected fish species namely *Selar crumenophthalmus, Decapterus maruadsi, Rastrelliger kanagurta.*

4. To study phylogenetic relationship of the selected fish species.

MATERIALS AND METHODS

A) Materials

1.0 Reagents

1.1 Common reagent

95% ethanol

1.2 Specific reagents

- **Sample preparation and preservation**

  1X Phosphate Buffer Saline (PBS)

- **DNA extraction using DNeasy Blood and Tissue Kit (Qiagen)**

  ATL buffer

  Proteinase K (Promega)

  AL buffer

  AW1 buffer

  AW2 buffer

  AE buffer

- **Agarose gel electrophoresis**

  0.5X Tris Borate EDTA (TBE) buffer

  100 bp DNA ladder

- **PCR**

  ReddyMix PCR master mix (Thermo Scientific)

  A set of primer (Sigma Aldrich)

- **PCR product purification using GeneAll Expin PCR SV kit (General Biosystem)**

  PB buffer

  NW buffer

  EB buffer

- **Termination cycle of purified PCR product using BigDye Terminator Cycle Sequencing Kit (Applied Biosystem)**

  5X buffer

  BigDye

- **Ethanol precipitation of purified PCR product and preparation for sequencing**

  2 M Sodium acetate

  70% chilled ethanol

  Hi-Di formamide (Applied Biosystem)

2.0 Chemicals

2.1 Specific chemicals

- **10X TBE**

  Tris base (Promega)

  Ethylene Diamine Tetraacetic Acid (EDTA) (Promega)

  boric acid (Promega)

- **Agarose gel electrophoresis**

  Agarose (Promega)

Ethidium bromide (Syne)

Orange-G dye (Syne)

- **Ethanol precipitation of purified PCR product and sequencing preparation**

Sodium acetate (Merck)

3.0 Instruments and apparatus

3.1 Common instruments and apparatus

Conform XT powder-free latex exam gloves (Ansell)

Dragon 204 Analytical Balance (Mettler Toledo)

0.5 mL and 1.5 ml reaction tubes (Greiner Bio-One GmBH)

Spectrafuge 24D microcentrifuge (Labnet)

Pipetman Ultra pipette 10 µl, 20 µl, 100 µl, 200 µl, 1000 µl (Gilson)

0.5 µL – 10 µL clear pipette tips, 1 µL – 200 µL yellow pipette tips (Axygen Scientific)

1000 µL blue pipette tips (Greiner Bio-One GmBH)

Vertical kelvinator freezer KN 308 (Fisher and Paykel)

ERLA EVM 6000 Vortex mixer

Hybrid guard system refrigerator (Toshiba)

AF 100 ice maker (Scotsman)

Combi-spin centrifuge PCV-2400 (Grant-bio)

GeneAmp PCR System 9700 (Applied Biosystems)

ERLA laminar air flow cabinet CFM-4

3.2 Specific instruments and apparatus

- **DNA extraction using DNeasy Blood and Tissue kit**

  Hot air oven sterilizer UE600 (Memmert)

  Laboratory waterbath WB29 (Memmert)

  DNeasy Mini spin column and 2 mL collection tubes (Qiagen)

- **10X TBE**

  ERLA hotplate and magnetic stirrer

  pH meter pH211 (Hanna)

- **Gel electrophoresis**

  MiniRun GE-100 gel electrophoresis system (Bioer Technology)

  Power-pac 3000V (Bio-Rad)

  OWL electrophoresis system B1, B1A (OWL Seperation system)

  ELBA microwave oven EMO-1706

  UV transilluminator (UVITEC Cambridge)

  Epichemi^3 Darkroom (UVP, Bioimaging System)

- **PCR product purification using GeneAll PCR SV kit (General Biosystem)**

  SV columns and collection tubes

- **DNA bidirectional sequencing**

  96 well plate

  Applied Biosystems Genetic Analyzer 3130xl

4.0 Softwares

BioEdit sequence alignment editor v 7.0.5.3

MEGA 5 Beta #7

B) Reagent preparation

All reagents were prepared using distilled water or deionized water

10X Tris Borate EDTA (TBE) buffer

107.8 g of Tris base and 7.44 g of EDTA were diluted in 500 mL of deionized water and made dissolved using the magnetic stirrer. pH was then adjusted to 8.3 using boric acid. The solution was then made up to 1000 mL using deionized water and autoclaved.

0.5X TBE buffer

50 mL of 10X TBE buffer was added into 950 mL of deionized water to make up a total volume of 1000 mL.

Proteinase K

20 mg of Proteinase K was mixed with 1 mL of deionized water.

3M Sodium Acetate

102.025 g of sodium acetate was added to 200 mL of deionized water. The pH was adjusted to 5.2 with glacial acetic acid. The solution was made up to 250 mL using deionized water and autoclaved.

2 M Sodium Acetate

16 mL of 3 M sodium acetate was added to 8 mL of deionized water.

70% Ethanol

350 mL of absolute ethanol was mixed with 150 mL of deionized water.

Ethidium Bromide stock solution

1 g of ethidium bromide was dissolved in 100 mL of deionized water and stored in amber bottle.

1% Agarose gel preparation (50 mL gel)

0.5 g of agarose powder was dissolved in 50 mL of TBE buffer and then was added with 1 μL of ethidium bromide stock solution.

1.2% Agarose gel preparation (50 mL gel)

0.6 g of agarose powder was dissolved in 50 mL of TBE buffer and then was added with 1 μL of ethidium bromide stock solution.

Orange-G dye

0.125 g of Orange-G and 20 g of sucrose was dissolved in 30 mL distilled water. The solution was then made up to 50 mL by the addition of distilled water. It was then aliquoted into 1.5 mL microfuge tube and kept under 4°C for further use.

PBS

0.08 g of sodium chloride (NaCl), 0.2 g of potassium chloride (KCl), 14.4 g of hydrated disodium hydrogen phosphate ($Na_2HPO_4.2H_2O$), and 2.4 g of potassium dihydrogen phosphate (KH_2PO_4) are dissolved in 800 mL of distilled water and then made up to 1 L.

C) Methods

Sample collection

Three species of commercially available fish from the local market were chosen for this study. The three species are namely *Selar crumenophthalmus* (ikan Lolong), *Decapterus maruadsi* (ikan Sadin) and *Rastrelliger kanagurta* (ikan Kembong Hitam).

DNA extraction using DNeasy Blood and Tissue kit

DNA was extracted from each of the sample by following the animal tissue spin-column protocol provided along with the kit. Firstly, 25 mg of fish tissue was cut into small pieces and then placed in a 1.5 mL microcentrifuge tube. A total of 180 µL of ATL buffer was added into the 1.5 mL microcentrifuge tube containing the fish tissue. This was followed by the addition of 20 µL of proteinase K into the same 1.5 mL microcentrifuge tube. The mixture was mixed thoroughly by vortexing. After that, the mixture was incubated for 3 to 4 hours at 56°C until the fish tissues completely lysed. The mixture was vortexed occasionally during the incubation period in order to disperse the tissues.

After the tissues had completely lysed, which was indicated by the clear solution of the mixture, the mixture was removed from the hot air oven and immediately vortexed for 15 seconds. An amount of 200 µL of AL buffer was then added into the vortexed mixture. This was followed by the addition of 200 µL of 95% ethanol into the vortexed mixture. After vortexing the contents, it was transferred into the DNeasy Mini spin column placed in a 2 mL collection tube. Then, the transferred mixture was centrifuged for 1 minute at 8000 rpm. The collection tube containing the flow-through was discarded.

The DNeasy Mini spin column was placed in a new 2 mL collection tube and 500 µL of AW1 buffer was added into the DNeasy Mini spin column and then centrifuged for 1 minute at 8000 rpm. The collection tube containing the flow-through was discarded. The DNeasy Mini spin column was then transferred in a new 2 mL collection tube and 500 µL of AW2 buffer was added into it and centrifuged for 3 minutes at 13,000 rpm to dry the DNeasy membrane. The collection tube containing the flow-through was discarded. The DNeasy Mini spin column was transferred in a clean 1.5 mL microcentrifuge tube and 200 µL of AE buffer was pipitted directly into the DNeasy membrane and incubated for 1 minute at room temperature. After incubation, it was centrifuged for 1 minute at 8000 rpm to elute. The eluted genomic DNA was stored at -20°C for future use.

Screening of extracted DNA by 1% agarose gel electrophoresis

The extracted genomic DNA was screened using 1% agarose gel in TBE buffer. A total of 3 µL of genomic DNA was loaded into individual well. Electrophoresis was performed for 30 minutes at 100 V. On completion of electrophoresis, the gel was visualized on UV transilluminator. High molecular weight genomic DNA was extracted from each of the 3 species namely *Selar crumenophthalmus* (ikan Lolong), *Decapterus maruadsi* (ikan Sadin) and *Rastrelliger kanagurta* (ikan Kembong hitam). The result of screening is shown in fig. 1.

Quantification of genomic DNA

The extracted genomic DNA samples from the three species of fish were quantified using single beam UV spectrophotometer and optical density (O.D) value at 260 nm were recorded. The genomic DNA was then diluted to 20 ng/μL.

Primer designing for PCR amplification

The published GenBank database for mtDNA COI gene for various species of vertebrates including fish were downloaded and aligned in BioEdit software. Primers were designed for the barcoding region of mtDNA COI gene from the conserved region flanking the barcoding region of mtDNA COI gene. During primer designing, percentage of GC content, the temperature of melting and the non-complementary between the forward and reverse primers were all taken into consideration. The primers designed are shown in Table 1.

Table 1. The sequences and characteristics of ST F1 (forward primer) and ST R3 (reverse primer).

Primer	Sequences	T_m (°C)	GC contents (%)
ST F1	CAC AAA GAC ATC GGC ACCC	60°C	57.9%
ST R3	AAG AAT CAA/G AAT AG/AG TGT TG	52°C	30.0%

PCR amplification with the designed oligoprimers

PCR amplification was carried out for the genomic DNA extracted from the three fish species. For each fish species, 20 ng of genomic DNA was used as template for PCR amplification using primer set ST F1 and ST R3. The total volume of master mix for PCR amplification is 25 μL. The components of PCR master mix is shown in table 2.

Table 2. Compositions of PCR master mix.

Components of master mix	*Selar crumenophthalmus* (μL)	*Decapterus maruadsi* (μL)	*Rastrelliger kanagurta* (μL)
ReddyMix	21.0	21.0	21.0
Primer set			
ST F1	1.5	1.5	1.5
ST R3	1.5	1.5	1.5
Template DNA	1.0	1.0	1.0
Total volume	**25.0**	**25.0**	**25.0**

The optimized PCR cycling parameter used for amplification process is shown in table 3. Amplification was performed using GeneAmp PCR system 9700 (Applied Biosystem).

Table 3. PCR amplification parameters.

Temperature (°C)	Time	Cycle	
95.0	3 minutes	Enzyme activation	
95.0	30 seconds	Template denaturation	30 cycles
52.0	30 seconds	Primer annealing	
72.0	50 seconds	Primer extension	
72.0	5 minutes	Final extension	
4.0	Hold		

Screening of PCR products using 1.2% agarose gel electrophoresis

After PCR amplification by using the extracted DNA from 3 species namely *Selar crumenophthalmus*, *Decapterus maruadsi* and *Rastrelliger kanagurta* as template DNA, the PCR products were checked by 1.2% agarose gel electrophoresis. A 30 mL 1.2% agarose gel was prepared with TBE buffer. A total of 2 µL of PCR product were loaded into individual well and electrophoresis was performed for 30 minutes at 100 V. The agarose gel was then visualized under UV transillumination for the presence of PCR products. The screening result is shown in fig. 2.

Purification of PCR products using GeneAll PCR SV kit

A total of 20 µL of PCR product was mixed with 250 µL of PB buffer. The mixture was then transferred into the SV column in 2 mL collection tube and centrifuged at 12,000 rpm for 30 seconds. The pass-through inside the 2 mL collection tube was discarded and the SV column was reinserted in the same 2 mL collection tube. An amount of 350 µL of NW buffer was added into the SV column and then was centrifuged at 12,000 rpm for 30 seconds. Again, the pass-through was discarded and the SV column was reinserted back in the same collection tube and then was centrifuged for 1 minute at 12,000 rpm. The collection tube together with the pass-through was discarded and the SV column was inserted in a new 1.5 mL micrcentrifuge tube. A total of 30 µL of EB buffer was added to the center of the SV membrane. The SV column containing EB buffer was incubated at room temperature for 1 minute. This was followed by centrifugation at 12,000 rpm for 1 minute. The SV column was then discarded and the eluted product preserved and examined by 1.2% agarose gel electrophoresis.

Screening of purified PCR products using 1.2% agarose gel electrophoresis

Purified products of PCR were checked in 1.2% agarose gel electrophoresis and visualized under UV transillumination. A 30 mL 1.2% agarose gel was prepared with TBE buffer. An amount of 2 µL of purified PCR product was loaded into individual well and electrophoresis was performed for 30 minutes at 100 V. Fig. 3 shows the purified PCR products.

Termination cycle using BigDye Terminator Cycle Sequencing Kit (Applied Biosystem)

Each component with their respective volume was added into a 0.5 mL microcentrifuge tube which made up a total volume of 7 µL. Finally, each of the 3 µL of purified PCR product was added into 0.5 mL microcentrifuge tube containing the primer set.

Table 4. Components of termination reaction mixture together with their respective volumes.

Components of termination reaction mixture	Volume per component (µL)
BigDye	0.5
5X buffer	1.75
*Primer	1.0
Distilled water	3.75
Total volume	**7.0**

* The volume of primer for termination reaction mixture was calculated by adding 6.7 µL of distilled water to 3.3 µL of each primer which was then equal to 10 pmol/µL.

PCR for the cycle termination reaction was performed using parameters as shown in table 5:

Table 5. PCR parameters for cycle termination process.

Temperature (°C)	Time	Cycle	
96.0	1 minute	Enzyme activation	
96.0	10 seconds	Template denaturation	35 cycles
50.0	5 seconds	Primer annealing	
60.0	4 minutes	Primer extension	
4.0	Hold		

Ethanol precipitation for purification of PCR products of cycle termination process

An amount of 10 µL of PCR product of cycle termination process was transferred into 1.5 mL microcentrifuge tube. A total of 3 µL of 2 M sodium acetate (pH 5.2) was added into the microcentrifuge tube. This was followed by the addition of 50 µL chilled 70% ethanol. The mixture was vortexed and left at room temperature for 30 minutes. The mixture was then centrifuged at 6000 rpm at 4°C for 40 minutes. The supernatant was discarded and 100 µL of