

**MITOCHONDRIAL GENOME ANALYSIS AND
GENETIC CHARACTERIZATION OF
ENDANGERED HUMPHEAD WRASSE,
CHEILINUS UNDULATUS (RÜPPELL, 1835)
FROM SABAH, MALAYSIA**



UMS

PRESSLY ANAK MATTHEW SABAH

**BORNEO MARINE RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH**

2019

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS

JUDUL: **MITOCHONDRIAL GENOME ANALYSIS AND GENETIC CHARACTERIZATION OF ENDANGERED HUMPHEAD WRASSE, CHEILINUS UNDULATUS (RÜPPELL, 1835) FROM SABAH, MALAYSIA**

IJAZAH: **SARJANA SAINS**

Saya **PRESSLY ANAK MATTHEW**, Sesi **2015-2019**, mengaku membenarkan tesis Sarjana Sains ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:-

1. Tesis ini adalah hak milik Universiti Malaysia Sabah
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (/):

SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA 1972)

TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

TIDAK TERHAD

Disahkan Oleh,

PRESSLY ANAK MATTHEW
MY1511004T

(Tandatangan Pustakawan)

Tarikh : 11 Julai 2019

(Prof. Madya Dr.Mabel Manjaji Matsumoto)
Penyelia Utama

(Prof. Madya Dr. Kenneth Francis Rodrigues)
Penyelia Bersama

**MITOCHONDRIAL GENOME ANALYSIS AND
GENETIC CHARACTERIZATION OF
ENDANGERED HUMPHEAD WRASSE,
CHEILINUS UNDULATUS (RÜPPELL, 1835)
FROM SABAH, MALAYSIA**



PRESSLY ANAK MATTHEW

UMMS
UNIVERSITI MALAYSIA SABAH

**THIS THESIS IS SUBMITTED IN
FULLFILLMENT FOR THE DEGREE OF
MASTER OF SCIENCE**

**BORNEO MARINE RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH
2019**

DECLARATION

I hereby declare that the material in this thesis is my own except of quotation, expects, equations, summaries and references, which have been duly acknowledged.

12th JULY 2019

.....
PRESSLY ANAK MATTHEW

MY1511004T



UMS
UNIVERSITI MALAYSIA SABAH

CERTIFICATION

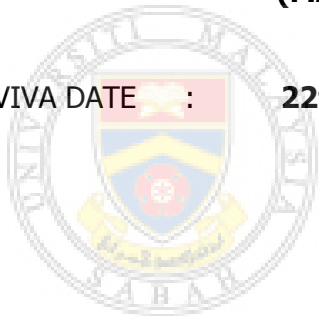
NAME : **PRESSLY ANAK MATTHEW**

MATRIC NO. : **MY1511004T**

TITLE : **MITOCHONDRIAL GENOME ANALYSIS
AND GENETIC CHARACTERIZATION OF
ENDANGERED HUMPHEAD WRASSE, *CHEILINUS
UNDULATUS* (RÜPPELL, 1835) FROM SABAH,
MALAYSIA**

DEGREE : **MASTER OF SCIENCE
(MARINE BIOTECHNOLOGY)**

VIVA DATE : **22th JANUARY 2019**



UMS
CERTIFIED BY,
UNIVERSITI MALAYSIA SABAH

1. MAIN SUPERVISOR

Assoc. Prof. Dr. Mabel Manjaji Matsumoto

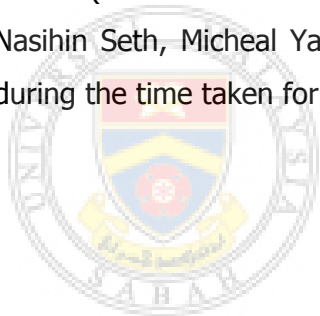
SIGNATURE

2. CO-SUPERVISOR

Assoc. Prof. Dr. Kenneth Francis Rodrigues

ACKNOWLEDGEMENT

First, I would like to express my deepest gratitude and appreciation to Universiti Malaysia Sabah for providing all the resources to accomplish this master thesis. I would also like to give my appreciation to main supervisor Dr. B. Mabel Manjaji Matsumoto for her patience, advice, guidance, and support during the completion of this study. In addition, I am grateful that my co-supervisor A.P. Dr. Kenneth Francis Rodrigues has lend a tremendous support in this study. This research was financially supported by the Ministry of Higher Education Malaysia (Research grant code: FRGS-0401-STWN-2/2014) and a matching grant from Universiti Malaysia Sabah (Research grant code: SBK0182-STWN-2015). Lastly, I wish to appreciate my family (Matthew Anak Umam, Abi Anak Rampeng, Rinie Anak Matthew, Cassey Anak Matthew) and friends (Carlmond Goh Kah Wun, Shahniza Tusin, Rolando Robert, Jovi Silvester, Siti Nasihin Seth, Micheal Yap) support as they are pillars of motivation and comfort during the time taken for this thesis completion.



UTMS
UNIVERSITI MALAYSIA SABAH

PRESSLY ANAK MATTHEW

11th JULY 2019

ABSTRACT

Mitochondrial (mt) Deoxyribonucleic acid (DNA) loci have been employed in various marine taxa, including reef fishes with the objective of resolving phylogeny, assigning parentage and delineating haplotypes. In this study, complete mtDNA genome was sequenced and the D-Loop locus was applied to estimate the population status of the Humphead Wrasse (HHW) *Cheilinus undulatus*. The Humphead Wrasse (HHW) is a long-lived reef-associated fish and the largest member of wrasses (family Labridae). However, it is endangered throughout its natural distributional range, primarily due to being a target in the Live Reef Fish for Food Trade. Specimens tissue were collected from three localities in Sabah, i.e. Kota Kinabalu, Sandakan and Semporna. The objectives of this study are 1) to sequence the mtDNA of the HHW, and 2) to identify population structure using D-LOOP of the HHW. The result of the sequencing data assemblies yielded six mitochondrial genomes and deposited into the National Center for Biotechnology Information (NCBI) with the accession number MH675879, MH675880, MH688049, MH688050, MH688051, MH688052. The result from AMOVA and overall F_{st} (-0.02914) showed that all three population absent from population structuring despite geographical distinct. The AMOVA showed a variation (100%) arise within the population instead among population. The neutrality test from Tajima's D (Kota Kinabalu: -1.825 and Sandakan: -1.009), and more sensitive Fu's F_s Statistic (Kota Kinabalu: -3.9844 and Sandakan-1.149), and mismatch distribution showed unimodal graph pattern suggested as *C. undulatus* undergoes recent population expansion. Haplotype frequency analysis produce a total 25 haplotypes whereas unique 11 haplotype is detected from three local population. High haplotypic diversity encountered in the Kota Kinabalu (0.682), Sandakan (0.902) and Semporna (1.000) and low nucleotide diversity in Kota Kinabalu (0.003), Sandakan (0.004), and Semporna (0.002), and starburst pattern of common haplotypes median joining network support *C. undulatus* population expansion. The first genetic characterization data from mtDNA and all three populations structure analysis of *C. undulatus* from Sabah study provide a new insight for conservation attention.

ABSTRAK

ANALISIS GENOM MITOKONDRIA DAN PENCIRIAN GENETIK, HUMPHEAD WRASSE TERANCAM CHEILINUS UNDULATUS (RÜPPELL, 1835) DARI SABAH, MALAYSIA

Loci nukleik asid mitokondria telah digunakan dalam beberapa kajian ikan terumbu karang dengan tujuan menyelesaikan filogeni serta asal usul dan menggambarkan haplotip. Dalam kajian ini, genom mtDNA lengkap telah diujukan serta locus d-loop digunakan untuk mengetahui status populasi spesis ikan Cheilinus undulatus (CU). CU adalah spesis ikan terumbu karang yang dapat hidup lama dan ahli terbesar dalam keluarga Labridae di bawah turutan Perciformes. Namun, CU terancam di seluruh habitat semulajadi kerana menjadi sasaran dalam Perdagangan Ikan Terumbu Karang Hidup. Tisu spesimen dikumpulkan dari tiga kawasan di Sabah, iaitu Kota Kinabalu, Sandakan dan Semporna. Objektif kajian ini adalah 1) untuk menjujuk urutan lengkap mtDNA genom dari HHW, dan 2) untuk mengenal pasti struktur populasi CU menggunakan D-LOOP. Hasil data penjujukan menghasilkan enam gen mitokondria dan disimpan ke Pusat Kebangsaan Maklumat Bioteknologi (NCBI) dengan nombor assesi MH675879, MH675880, MH688049, MH688050, MH688051, MH688052. Hasil daripada AMOVA dan keseluruhan Fst (-0.02914) menunjukkan bahawa ketiga-tiga populasi tidak mengalami penstrukturan walaupun berbeza secara geografi. AMOVA menunjukkan variasi (100%) timbul dalam populasi bukannya di antara populasi. Ujian neutral dari Tajima `D (Kota Kinabalu: -1.825 dan Sandakan: -1,009), dan statistik Fu's Fs yang lebih sensitif (Kota Kinabalu: -3.9844 dan Sandakan-1.149), dan mismatch distribution menunjukkan corak grafik unimodal Justeru dicadangkan bahawa C. undulatus mengalami perkembangan populasi yang terdekat. Analisis frekuensi haplotype menghasilkan sejumlah 25 haplotip manakala 11 haplotip yang unik dikesan hasil dari tiga populasi CU tempatan. Kepelbagaian haplotip tinggi yang terdapat di Kota Kinabalu (0.682), Sandakan (0.902) dan Sempona (1.000) dan kepelbagaian nukleotida yang rendah di Kota Kinabalu (0.003), Sandakan (0.004) dan Semporna (0.002) turut menyokong bahawa CU mengalami perkembangan populasi yang terdekat. Data pencirian genetik pertama dari mtDNA dan ketiga-tiga analisis struktur populasi CU dari Sabah memberikan pandangan baru untuk perhatian pemuliharaan yang segera.

TABLE OF CONTENTS

	Page
TITLE	i
DECLARATION	ii
CERTIFICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
ABSTRAK	vi
TABLE OF CONTENT	vii
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS AND SYMBOLS	xii
LIST OF APPENDICES	xiv
CHAPTER 1: INTRODUCTION	
1.1 Introduction	1
1.2 Research Problem	3
1.3 Objectives	3
1.4 Research Approach	4
CHAPTER 2: LITERATURE REVIEW	
2.1 General Physical, Appearance and Sex Change of <i>Cheilinus undulatus</i> (Rüppell, 1835).	5
2.1.1 Global Distribution, Habitat and Roles.	7
2.2 Live Reef Fish For Food Trade	8
2.2.1 Illegal, Unregulated, and Unreported Activities (IUU)	10
2.3 Conservation and Protection Status	11
2.4 Geography Of Sabah	12
2.5 Origin of Mitochondrial, Function and Inheritance	13
2.5.1 Gene order in Mitochondrial genome	15
2.5.2 Evolution in Mitochondrial	16
2.5.3 Mitochondrial as a Genetic Marker	17
2.6 Population Genetics and Conservation	18
2.7 Phylogenetic Tree	20
2.7.1 Method of Analysis for Phylogenetic Tree	22

2.8 Population Structuring Analysis	23
2.8.1 Fixation index (Fst)	23
2.8.2 AMOVA	24
2.8.3 MISMATCH DISTRIBUTION	25
2.8.4 NEUTRALITY TEST	26
CHAPTER 3: METHODOLOGY	
3.1 Specimen Collection	27
3.1.1 Market Sampling	27
3.1.2 Specimen Voucher Preparation	28
3.1.3 Tissue Muscle and Fin Clip Collection	28
3.2 DNA Isolation and Quantification	28
3.2.1 Whole DNA Extraction from White Tissue Muscle	28
3.2.2 Gel Electrophoresis and Spectrophotometry	29
3.3 Target Enrichment Strategy	31
3.3.1 Reference Mitochondrial Genome	31
3.3.2 Primer Design	31
3.3.3 Amplification	33
3.3.4 Purification	33
3.4 DNA Sequencing	34
3.4.1 Library Preparations	35
3.4.2 Tagment Genomic DNA	35
3.4.3 Library Clean Up	37
3.4.4 Libraries Normalisation	37
3.4.5 Pool Libraries	38
3.5 Bioinformatics Analysis	38
3.5.1 Genome Assembly	38
3.5.2 Genome Annotation	39
3.6 Phylogenetic Analysis of Labridae	40
3.7 Population Analysis using D-Loop Gene	41
3.8 Demographic History and Population Mismatch Analysis	41
3.9 Haplotype Frequency and Median-Joining Network	41
Chapter 4 : Result	
4.1 Specimen Collection	43

4.2 Target Enrichment Amplification	43
4.3 Mitochondrial Genome Assembly	44
4.4 Mitochondrial Genome Annotation	47
4.5 Gene Arrangement.	52
4.6 Labridae Phylogenies Comparative Analysis	52
4.6.1 Neighbour -Joining Phylogenetic Tree	53
4.6.2 Maximum Evolution Phylogenetic Tree	54
4.6.3 Maximum Likelihood Phylogenetic Tree	55
4.6.4 Genetic Differences among Labridae Species	56
4.7 Molecular Diversity of D-loop Region among Populations	57
4.8 Population Comparative Analysis	58
4.9 Population Demographic Analysis	59
4.10 Population Mismatch Analysis	60
4.11 Mismatch Distribution	61
4.12 Relative Frequencies of <i>Cheilinus undulatus</i> Haplotypes	62
4.13 Median-Joining Network Analysis	63
4.14 Conserved Region in D-Loop Site	64
Chapter 5: Discussion	
5.1 Mitochondrial DNA	65
5.1.1 Genome assemblies	65
5.1.2 Gene composition and gene arrangement	67
5.2 Phylogenetic tree	68
5.3 Population comparative analysis	69
5.3.1 AMOVA results	71
5.3.2 Demographic data	71
5.3.3 Median Joining Network	71
5.4 Conservation of <i>Cheilinus undulatus</i> based on the genetic makeup	73
5.5 Recommendation	76
Chapter 6: Summary	
REFERENCES	78
APPENDICES	86

LIST OF TABLES

		Pages
Table 2.1	Example of inferring methods	22
Table 3.1	The primer for this study	32
Table 3.2	The primer design using Primer3 software	32
Table 3.3	PCR cocktails	33
Table 3.4	PCR protocol	33
Table 3.5	Library amplification	35
Table 3.6	Sequencing library index	36
Table 3.7	Complete mtDNA from Labridae species	40
Table 3.8	Source of D-loop of <i>Cheilinus undulatus</i> used in this comparative study	42
Table 4.1	Mitochondrial genome assembly result.	44
Table 4.2	MH675879 Mitochondrial genome annotation	45
Table 4.3	MH675880 Mitochondrial genome annotation	46
Table 4.4	MH688049 Mitochondrial genome annotation	47
Table 4.5	MH688050 Mitochondrial genome annotation	48
Table 4.6	MH688051 Mitochondrial genome annotation	49
Table 4.7	MH688052 Mitochondrial genome annotation	50
Table 4.8	Nucleotide composition of Labridae species	52
Table 4.9	Molecular diversity	57
Table 4.10	Population pairwise FST using Kimura 2-parameter in Arlequin	58
Table 4.11	Population average pairwise differences.	58
Table 4.12	AMOVA	58
Table 4.13	Tajima's D selective neutrality test	59
Table 4.14	Fu's Fs neutrality test	59
Table 4.15	Demographic expansion data	60
Table 4.16	Spatial expansion data	60
Table 4.17	Relative frequencies of <i>Cheilinus undulatus</i> haplotypes with corresponding localities.	62

LIST OF FIGURES

		Pages
Figure 2.1	<i>Cheilinus undulatus</i> voucher specimen	6
Figure 2.2	Global distribution of <i>C. undulatus</i>	7
Figure 2.3	Sabah map	13
Figure 2.4	A circular structure of mitochondrial genome	15
Figure 2.5	A typical component of phylogenetic tree	20
Figure 2.6	Mismatch distribution graph	25
Figure 3.1	The fish market collection site	27
Figure 3.2	Library preparation of fragments DNA	34
Figure 4.1	The complete circular mitochondrial genome of <i>Cheilinus undulatus</i> illustrated using MITOfish online pipeline annotation web interface	51
Figure 4.2	Neighbour-Joining Phylogenetic Tree.	53
Figure 4.3	Maximum Evolution Phylogenetic Tree	54
Figure 4.4	Maximum Likelihood Phylogenetic Tree	55
Figure 4.5	Genetic Differences among Labridae Species	56
Figure 4.6	The mismatch distribution	61
Figure 4.7	Median-joining networks	63
Figure 4.20	The Conserved region in the D-loop region of <i>Cheilinus undulatus</i> .	64

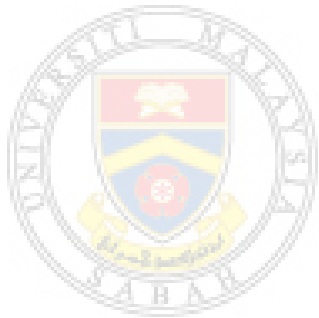
LIST OF ABBREVIATIONS AND SYMBOL

ATM	Amplicon Tagment Mix
ATP	Adenosine Trisphosphate
CE	Critically Endangered
CITES	Convention on International Trade of Endangered Species
COI	Cytochrome oxidase 1
COII	Cytochrome oxidase 2
COIII	Cytochrome oxidase 3
CTI	Coral Triangle Initiative
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
EN	Endangered
ETBR	Ethidium Bromide
ETOH	Ethanol
GBR	Great Barrier Reef
IUCN	International Union for Conservation Nature
IUU	Illegal, Unreported, Unregulated
LNA1	Library Normalization Additives 1
LNB1	Library Normalization Beads 1
LNS1	Library Normalization Storage Buffer 1
LNW1	Library Normalization Wash 1
LRFFT	Live Reef Fish for Food Trade
MEGA	Molecular Evolutionary Genetic Analysis
ML	Maximum Likelihood
MPA	Marine Protected Area
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NJ	Neighbor Joining
nm	Nano meter

NPM	Nextra PCR Master Mix
NT	Neutralize Tagment Buffer
OTU	Operational Taxonomic Units
PCR	Polymerase Chain Reaction
PPE	Personal Protection Equipment
REGO	Registration Number
RPM	Rotation Per Minute
rRNA	Ribosomal Ribonucleic Acid
RSB	Resuspension Buffer
SEA	South-East Asia
SL	Standard Length
SNP	Single Nucleotide Polymorphism
TBE	Tris Borate EDTA
TD	Tagment DNA Buffer
TL	Total Length
tRNA	Transfer Ribonucleic acid
UPGMA	Unweighted Pair Group Method with Arithmetic
UV	Ultra Violet
km²	Kilometer Square
cm	Centimeter
m	Meter
ml	Millilitre
mm	Millimeter
x g	Force unit
µl	Microlitre

LIST OF APPENDICES

	Pages
Appendix A Specimen Collection	86
Appendix B PCR results	87
Appendix C Published Journal	93
Appendix D Presentation & Conferences	94



UMS
UNIVERSITI MALAYSIA SABAH

CHAPTER 1

INTRODUCTION

1.1 Introduction

Sabah is known for its outstanding diversity in both marine and terrestrial species. Geographically, Sabah is bounded by three seas, namely, South China Sea, Sulu Sea and Celebes Sea. These Sea combinations produced today's marine fish biodiversity. In addition, Sabah is located in the Coral Triangle which is a marine biodiversity hotspot with the largest density of coral fishes ever recorded. Sabah waters are natural habitat of the giant coral reef fish, *Cheilinus undulatus*. For the last few decades, *C. undulatus* populations have been shown widely distributed in tropical waters. This species was mostly reported in Oceania, Great Barrier Reef (GBR), North-Western Indian Ocean, Red Sea and Western Pacific Ocean. Taxonomy-wise, *C. undulatus* is a member of the wrasse family (Labridae) and a natural cleaner fish in coral reef ecosystems. The local community calls it "Mameng", "Ikan Licin", and "Ikan Sumei" while the international community recognise it as Humphead wrasse, Napoleon wrasse and Maori wrasse. In Malaysia, wild *C. undulatus* populations are mostly distributed throughout western, northern and southern coasts of Sabah (Arai, 2015; Chen & Ng, 2009).

The wild population of *C. undulatus* is under threat from Live Reef Fish for Food Trade (LRFFT) and natural habitat degradation. Together with other cousin species such as groupers (Serranidae), *C. undulatus* is recognised as a famous seafood item for the South-East Asian market. The "economic value" of *C. undulatus* in LRFFT has contributed to the rapid shrinking of its wild population. In addition, the illegal, unreported, and unregulated fishing (IUU) activities quickly drive *C. undulatus* populations into a more vulnerable state.

In the year 2003, early-stage conservation efforts introduced *C. undulatus* as a threatened species under the International Union for Conservation of Nature (IUCN) Red List. As of 2019, *C. undulatus* is declared as endangered by the IUCN. Secondly, The Convention on International Trade of Endangered Species (CITES) has listed *C. undulatus* under Appendix Two which states that any trade of *C. undulatus* should be strictly controlled in the global market. The decrease in wild *C. undulatus* population numbers within a human's life span has attracted the attention of marine conservationists around the globe.

Physical and biological scientific data are needed to mitigate the extinction risk of *C. undulatus*. Molecular insights such as genetic structure, genetic variation and gene flow provide helpful input in conservation means rather than depending solely on annual fisheries landing statistics. Therefore, to understand the wild population of *C. undulatus*, investigations based on molecular studies using genetic markers is required. Using the Next-Generation Sequencing (NGS) Technology Platform, the molecular markers, namely, mitochondrial DNA (mtDNA) and the control region (D-loop) are proposed in this study to highlight the genetic insight on *C. undulatus* from Sabah. To date, this study revealed the first complete mitochondrial genome of *C. undulatus* from Sabah. The first part of this study is the genetic characterization of the mtDNA which exhibits a total of 37 regions and the second part is the investigation of the population using D-loop regions. This study addresses the magnitude of the genetic diversity level within and between *C. undulatus* populations as reflected in their geographical differences. By comparing the differences, a selected population could be suggested for further conservation efforts such as artificial breeding programs.

This genetic marker allows a comparison between populations based on their different geographical locations to be made. The selected population may reveal new opportunities for the species to survive overexploitation. It is crucial that *C. undulatus* populations are monitored closely both biologically and physically. Hence, a suitable conservation management strategy could be improved using both means to reduce the extinction risk of *C. undulatus*.

1.2 Research Problem

The strategies for implementing effective conservation and management practices in marine resources are justified by the current population stocks of *Cheilinus undulatus*. Firstly, there is a research knowledge gap in the available information on *C. undulatus* based on its genetic structures especially from Sabah waters. The investigations in this thesis therefore aims to contribute molecular insights of Sabah's giant coral reef fish. Secondly, such information will be crucial and will also be a cornerstone for the establishment and evaluation of the existing marine protected areas (MPAs) in Sabah. The findings from this study facilitates the selection of populations that are suitable for further conservation needs. This is done through investigations which highlight the significant genetic diversity in a given population area. The population with higher diversity could signify a higher fitness level. Thus, the said population may have better chances in population recovery programs such as artificial breeding as well as aquaculture-related programs.

In addition, selecting the better breeds also has economic importance such as preventing unnecessary actions in management which may incur more costs. Focusing only on the targeted population in the known marine conservation area rather than covering the entire marine conservation area would require less manpower, time, and financial resources. Likewise, the establishment of marine protected areas based mainly on historical or biological data might need the input from molecular genetic data. Reliance on historical and oceanographic data might limit the actual understanding of this species in the wild. Therefore, this study proposed the D-loop population analysis of *C. undulatus* will fill the research knowledge gap to benefit the future population recovery programs by selecting the most genetically diverse population.

1.3 Objectives

The objectives of this study were;

- (a) To produce the mitochondrial DNA sequence of *Cheilinus undulatus* individuals from (1) Kota Kinabalu and (2) Sandakan.
- b) To identify polymorphic site in mitochondrion genome that represents *Cheilinus undulatus* locality.

1.4 Research Approach

There were three approaches to this study. The first strategy was the random sampling technique (Green, 1979). This was conducted by collecting the *C. undulatus* specimens from the fish traders in different localities from local fish markets in Sabah.

The second strategy was to use Next-Generation Sequencing (NGS) technology to sequence the complete mitochondrial genome of *C. undulatus*. The mitochondrial genome template with accession number KM 461717 was retrieved from NCBI depository to synthesize a total of six sets of primers which target *C. undulatus*. We applied the long amplification of six fragments of mtDNA genome for this species in Sabah. The generated sequence data from NGS was assembled *de novo* using the SPAdE assembler (Bankevich *et al.*, 2012) before corrected using Geneious version 11 (Kearse *et al.*, 2011).

The third strategy was the genetic characterization of *C. undulatus* mtDNA and the population analysis of *C. undulatus*. The genetic characterization was done by annotating the correct gene position and gene arrangement using MitoFISH software. Next, a molecular phylogenetic tree analysis was used to construct the evolutionary relationships within the group of Labridae family. The suggested distance matrix method was Neighbour-Joining (Saitou and Nei, 1987) and p-distance was used in the nucleotide substitution model as it is versatile and less complex. The population analysis of *C. undulatus* in this study was conducted using the D-loop region. Population structuring was detected using AMOVA and F_{ST} (Reynolds *et al.* 1983; Slatkin, 1995) fixation index. Demographic history analysis was estimated using Tajima's D test (Tajima, 1989) and Fu's F_s test (Fu, 1997). In addition, a population mismatch graph was constructed together with demographic expansion (Harpending, 1994) and spatial expansion analysis. Next, the frequency of haplotype and the molecular diversity of the D-loop region among populations was included in end of this study and constructed using DNAsp (Rosas *et al.*, 2003) software. Lastly, the median joining network analysis was visualised using PopART software (Leigh and Bryant, 2015).

CHAPTER 2

LITERATURE REVIEW

2.1 General Physical Appearance and Sex Change of *Cheilinus undulatus* (Rüppell, 1835).

Cheilinus undulatus is a member of the Labridae family under the order of Perciformes which can attain over two meters in total length (TL) and is listed as one of the giants among reef fishes (Donaldson & Sadovy, 2001). This species is the longest living member of Labridae and has been reported to live up to 30 years (Sadovy *et al.*, 2003). Globally known as Maori, Napoleon, Humphead Wrasse in English, So-Mei in Hong Kong, Southern China and Taiwan (Donaldson & Sadovy, 2001). In Sabah, it is known as "Maming/Mameng", "Ikan Licin" and "Sumei" by the locals (Daw *et al.*, 2002). The famous naming of this species as Napoleon fish comes from symbolic trophy fish of Mr. Napoleon, a fisher and farmer in New Caledonia (Fourmanoir & Laboute, 1976).

Male *C. undulatus* usually grow larger than female members with the maximum recorded total length of two meters and weighing up to 190 kilograms (Donaldson & Sadovy, 2001). On the other hand, female *C. undulatus* attained sizes of no longer than one-meter in length and weigh up to 90 kilograms. Like most reef fishes, morphological changes of *C. undulatus* occur in its body colouration and physical form throughout its lifecycle. This accounts for its multiple names throughout the world. The external body colouration of *C. undulatus* changes according to its maturity status (Sadovy de Mitcheson *et al.*, 2010). The early juveniles of *C. undulatus* are black and white in external body colour. Meanwhile, the larger juveniles become pale green with a black marking on each scale forming two black lines from each eye.

Adult *C. undulatus* are often olive green to blueish green together with a spindle-shaped bar on each scale. A mature *C. undulatus* body colour is dull blue or green before turning into green vibrant shades and purplish blue large scales (Donaldson & Sadovy, 2001). The forehead only starts to form when the body reached 35.0 cm to 50.0 cm total length (Liu *et al.*). For example, in larger adults, the forehead may reach its anterior to the level of the eyes and comes together with its fleshy lips (Randall *et al.*, 1978). Meanwhile, females of *C. undulatus* reached the maturity stages at the ages of five years old with the total length (TL) between 35.00cm to 50.0 cm (Sadovy *et al.*, 2003). The males continue to grow rapidly than its female counterpart. Choat *et al.* (2006) reported that males reach a TL of 140 cm.

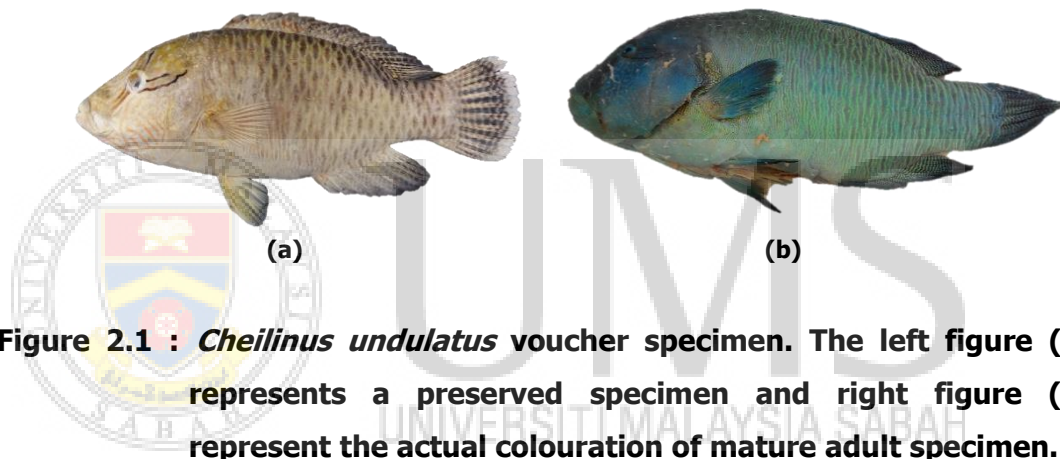
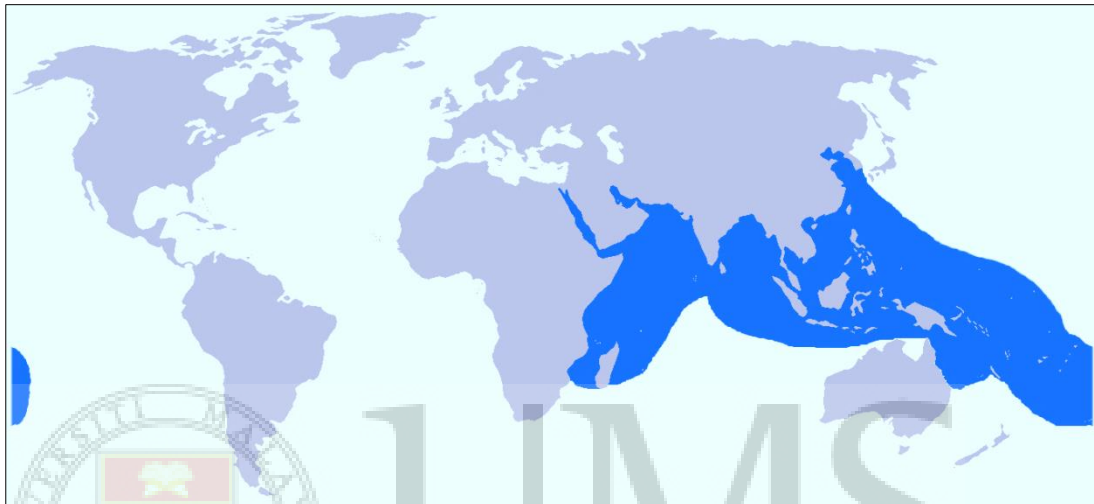


Figure 2.1 : *Cheilinus undulatus* voucher specimen. The left figure (a) represents a preserved specimen and right figure (b) represent the actual colouration of mature adult specimen.

Reef fishes are one of the remarkably diverse families among vertebrates due to their reproductive adaptations in their natural habitat. Reef fishes have the ability to change their sex especially when the male proportion is low. The Napoleon Wrasse is reported as a protogynous hermaphrodite species (Sadovy *et al.*, 2010). Protogynous hermaphroditism is the ability of a species to change its gender from female to male at a later stage of life (Warner, 1988). A study showed that *C. undulatus* females undergo sex change at the age of nine years old (Choat *et al.*, 2006). The lack of male opponents during mating season triggers female *C. undulatus* to undergo hormonal changes. The hormonal changes cause the ovaries to transform into male testes and they develop male morphological forms (forehead) before they finally mate with another female. This phenomenon is known as sequential hermaphroditism and is common in reef fish reproduction stages.

2.1.1 Global Distribution, Habitat and Roles.

Cheilinus undulatus is reported in warm water regions throughout Oceania from Indian Ocean to the Central Pacific Ocean, east coast of Africa and Red Sea (Donaldson & Sadovy, 2001). The global habitat distribution of *C. undulatus* is mainly in coral reefs and inshore throughout much of the tropical water of Indo-Pacific (Sadovy *et al.*, 2003).



Figures 2.2 : Global distribution of *C. undulatus*. The blue represents the natural occurrences of *C. undulatus* in the wild.

Source : Sadovy *et al.* (2003).

The habitat preferences of *C. undulatus* vary based on its maturity stages. Like coral reef fishes, *C. undulatus* are prone to select hard and soft branching corals community and sea grasses for habitat preferences. The juveniles tend to be more cryptic existences in areas of dense branching corals, bushy macroalgae or sea grasses, and shallow shady ranges that surround coral reef waters. Meanwhile, adults prefer habitats on the edge of the reefs, channels and reef passes which are mostly in shallow shady ranges. According to Sadovy *et al.* (2003), *C. undulatus* is one of the large teleost fish that are often observed at depths of 0.5 m to 100 m. Jones *et al.*'s (2002) study suggests that most large-sized coral reef fish preferred large geographic ranges.

Cheilinus undulatus is known to be diurnally active (Chateau & Wantiez, 2007), they rest at night in crevices in the reef within the normal home range that differ from spawning site. *Cheilinus undulatus* produce pelagic eggs and the larvae ultimately settle on or near coral reef habitats. The eggs are 0.65mm in diameter and are spherical with no pigments (Colin, 2010; Hutapea & Slamet, 2006). Most *Cheilinus* spp. larvae are typically small which ranges from 8.0mm to 11.00mm TL when they settle. The diets of *C. undulatus* include ocean invertebrates such as molluscs, crustaceans and echinoderms (Randall *et al.*, 1978).

Cheilinus undulatus plays crucial roles in ocean health by maintaining the coral ecosystem. Wrasses are well known as cleaners of the marine ecosystem (Sweatman, 1995). Like other cleaner wrasses, *C. undulatus* has a protruding jaw-like mouth for feeding purpose. It provides protection for the coral ecosystem by consuming sea urchins and crown of thorn starfish which are natural predators for coral reef species (Brodie *et al.*, 2005). The population of crown of thorn starfish is maintained at the lowest level possible with the presence of *C. undulatus*. By regulating the populations of poisonous organisms such as thorn starfish and sea urchins, the surrounding ecosystem is kept at safe levels. By removing this key predator from the wild coral reef ecosystem, the coral reef's resilience is at risk.

2.2 LIVE REEF FISH FOR FOOD TRADE

The live reef fish for food trade (LRFFT) refers to the trade of live reef fish for consumption and ornamental purposes. The main exporter countries of LRFFT are Indonesia, Philippines, Australia, Malaysia, Thailand, Vietnam, Taipei and China (Lee & Sadovy, 1998). In addition, pacific countries such as Fiji, Maldives, Marshall Islands, Papua New Guinea, Seychelles, Singapore and Solomon Islands are reported to be involved in LRFFT (McGilvray & Chan, 2003; Sadovy *et al.*, 2003). Recognised as one of luxury menu in seafood restaurants, *C. undulatus* wild population is increasingly threatened across the Indo-Pacific as a consequence of LRFFT activities which is centralised in Hong Kong (Sadovy *et al.*, 2003). The rise of LRFFT put this species population in danger more than any other factors.