Faculty of Resource Science and Technology

Genetic Diversity of False Gharial *Tomistoma schlegelii* based on Cytochrome *b*-Control Region (cyt *b*-CR) Gene Analysis

Muhamad Farhan Bin Badri (34975)

Bachelor of Science with Honours
(Aquatic Resource Science and Management)
2015

UNIVERSITI MALAYSIA SARAWAK

Grade: A-

Please tick (√)
Final Year Project Report ☑
Masters ☐
PhD ☐

DECLARATION OF ORIGINAL WORK

This declaration is made on the 2 day of July year 2015

Student's Declaration:

I MUHAMAD FARHAN BIN BADRI , 34975 , FSTS
(PLEASE INDICATE NAME, MATRIC NO. AND FACULTY) hereby declare that the work entitled, Genetic Diversity False Gharial T. Schlegelii based on cyt b-CR gene analysis is my original work. I have not copied from any other students' work or from any other sources with the exception where due reference or acknowledgement is made explicitly in the text, nor has any part of the work been written for me by another person.

2 / 7 / 2015
Date submitted

MUHAMAD FARHAN BIN BADRI (34975)
Name of the student (Matric No.)

Supervisor's Declaration:

I, DR. RUHANA BINTI HASSAN (SUPERVISOR'S NAME), hereby certify that the work entitled, Genetic Diversity False Gharial T. schlegelii based on cyt b-CR Gene Analysis (TITLE) was prepared by the aforementioned or above mentioned student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of Aquatic Resource Science and Management (PLEASE INDICATE THE DEGREE TITLE), and the aforementioned work, to the best of my knowledge, is the said student's work

Received for examination by: DR. RUHANA HASSAN
(Name of the supervisor)

Date: 2 / 7 / 2015

I declare this Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*
☐ **RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*
☑ **OPEN ACCESS**

I declare this Project/Thesis is to be submitted to the Centre for Academic Information Services (CAIS) and uploaded into UNIMAS Institutional Repository (UNIMAS IR) (Please tick (√)):

☑ **YES**
☐ **NO**

Validation of Project/Thesis

I hereby duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic and research purposes only and not for other purposes.
- The Centre for Academic Information Services has the lawful right to digitize the content to be uploaded into Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis if required for use by other parties for academic purposes or by other Higher Learning Institutes.
- No dispute or any claim shall arise from the student himself / herself neither a third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student himself/herself without first obtaining approval from UNIMAS.

Student's signature ______________________ Supervisor's signature: ______________________
(Date) (Date)

Current Address:
~~MUHAMAD FARHAN~~ NO 15, JALAN KELUBI, 4/7A, SEKSYEN 4, 40000, SHAH ALAM,
SELANGOR.

Notes: * If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the date of restriction indicated, and the reasons for the confidentiality and restriction.

[The instrument was prepared by The Centre for Academic Information Services]

DECLARATION

No portion of the work referred to in this dissertation has been submitted in support of an application for another degree qualification in this or any other university or institution of higher learning.

Muhamad Farhan Bin Badri

Aquatic Resource Science and Management Programme

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

Acknowledgement

Alhamdulillah and my utmost gratitude to Allah SWT, for giving me strength, wisdom and knowledge to complete this dissertation and to keep me through many obstacles experienced prior the completion of this project.

Hereby, I wish to take this opportunity to express my gratitude and appreciation to all those that have assisted, involved and guided me throughout the completion of this project. First of all, special thanks to Dr Ruhana Hassan, the Head of Department of Aquatic Science UNIMAS who is also my supervisor for trusting me to do this project and for giving me constructive advice, encouragement, guidance and assistance along the way until the completion of this project. Thank you very much for being so patient with me.

Next, my appreciation goes to all postgraduate students in Aquatic Molecular Lab especially Miss Nursyuhaida Md Shahid and Mr Amirul Arib for their kind assistance in lab work and useful discussion that help me a lot in this project. Both of them had been actively and always been available to advise me and gave their countless help. Thank you very much.

I would also like to extend my gratitude to my labmates and friends, especially to Miss Farhah Binti Dzarif for giving me continuous moral support. I could not finish this dissertation without their help. Last but not least, all my love and thanks goes to my family for always being there, giving me their support. Thank you.

Table of Contents

List of Abbreviations

Abbreviation	**Full Term**
bp	Base pairs
CIA	Chloroform-Isoamyl Alcohol
CITES	Convention on International Trade in Endangered Species
CTAB	Cetyl-trimethyl Ammonium Bromide
Cyt *b*	Cytochrome *b*
Cyt *b*-CR	Cytochrome *b*-control region
dNTPs	Deoxynucleotide triphosphate
DNA	Deoxyribonucleic acid
EDTA	Ethylene diaminetetra-acetic acid
EtOH	Ethanol
EtBr	Ethidium bromide
g	Gram
IUCN	International Union for Conservation of Nature
$MgCl_2$	Magnesium Chloride
min	Minutes
ml	Milliliter
mtDNA	Mitochondrial DNA
NaCl	Sodium chloride
rpm	Revolutions per minute
s	Seconds
VNTRs	Variable number tandem repeats
µl	Microliter
ºC	Degree Celcius

List of Tables

List of Figures

Genetic Diversity of False Gharial *Tomistoma schlegelii* based on Cytochrome b-Control Region (cyt b-CR) Gene Analysis

Muhamad Farhan Bin Badri

Aquatic Resource Science and Management

Faculty of Resource Science and Technology

Universiti Malaysia Sarawak

Abstract

Tomistoma schlegelii or known as false gharial, is a freshwater crocodilian endemic to Malaysia and Indonesia. This species is currently affected with severe habitat loss as their numbers in the wild have been decreasing due to legal and illegal logging, urbanization, and unsustainable fishing activities. The genetic studies on gharials have been carried out but the genetic information regarding *T. schlegelii* is still lacking. Therefore, this study is designed to sequence the cytochrome *b*-control region (cyt *b*-CR) gene and to determine genetic diversity of *T. schlegelii* individuals in Sarawak. A total of 647 bp of cyt *b*-CR gene has been sequenced and it matched 99% sequences data of *T. schlegelii* (Accession No: HM593977). The new cyt *b*-CR sequence data has been deposited in the GenBank (Accession No: KR674082). Genetic divergence value between *T. schlegelii* in Asajaya and those from GenBank originated from Sarawak showed 1.1% distance. Both Neigbour-Joining and Maximum Parsimony phylogenetic trees agreed on the monophyletic status of *T. schlegelii*.

Keywords: *Tomistoma schlegelii*, cytochrome *b*-control region

Abstrak

Tomistoma schlegelii atau dikenali sebagai buaya julung-julung adalah buaya air tawar yang hanya boleh dijumpai di Malaysia dan Indonesia. Spesies ini kini sedang mengalami kepupusan kerana aktiviti pembalakan, pembandaran dan penangkapan ikan secara tidak terkawal. Data kajian mengenai genetik *T. schlegelii* masih kurang walaupun kajian telah dilakukan. Kajian ini bertujuan untuk mendapatkan gen cytochrome *b*-control region (cyt *b*-CR) dan menentukan kepelbagaian genetik *T. schlegelii* di Sarawak. Penjujukan DNA menghasilkan 647 bp dan 99% padan dengan pangkalan data daripada GenBank (No. Akses: HM593977). Gen cyt *b*-CR terbaru telah dikemaskini di dalam Genbank (No. Akses: KR674082). Analisa variasi genetik menunjukkan 1.1% perbezaan antara *T. schlegelii* di Asajaya dan yang berasal dari Sarawak. Pokok filogeni membuktikan bahawa *T. schlegelii* ialah monofiletik.

Kata kunci: *Tomistoma schlegelii*, cytochrome b-control region

1.0 Introduction

The endangered *Tomistoma schlegelii*, also known as false gharial or Malayan gharial, is a freshwater crocodilian. This species is endemic to Malaysia and Indonesia (Auliya *et al*., 2006; Bezuijen *et al*., 2001) and is currently under Appendix I of Convention on International Trade in Endangered Species (CITES) (International Union for Conservation of Nature (IUCN), 2009) with an estimated population of less than 2500 individuals left in the wild (Bezuijen e*t al*., 2010). Studies had shown gradual decline of *Tomistoma* sightings in Kalimantan and Sumatra from the 1900s to 2000s (Auliya *et al*., 2006; Bezuijen *et al*., 1998; Bezuijen *et al*., 2001) while in Malaysia sightings have been rare since the late 1980s (Cox and Gombek, 1985; Sebastian, 1993; Simpson *et al*., 1998). The main threat to this freshwater crocodilian are habitat loss due to legal and illegal logging, urbanization, fishing activities and forest fires at peat swamps resulting in loss of suitable nesting sites and habitats for *Tomistoma* (Auliya *et al*., 2006; Bezuijen *et al*., 1998; Bezuijen *et al*., 2001).

Currently, most of these wild individuals are found in captivity in zoos and animal sanctuaries. Approximately 77 individuals were found in zoos and farms in Malaysia (Stuebing *et al*., 2004), while Kaur *et al.* (2012) only involved 7 individuals of *T. schlegelii* from Sarawak to complete their recent molecular phylogenetic research. Existing records have shown that *T. schlegelii* can still be found in the wild in Kalimantan and Sumatra, Indonesia (Bezuijen *et al*., 1998; Auliya *et al*., 2006).

Thus far, studies in both Malaysia and Indonesia have concentrated on surveys of *Tomistoma* in natural habitat and related issues involving conservation of habitat for the false gharial (Auliya *et al*., 2006; Bezuijen *et al*., 1998; Bezuijen *et al*., 2001; Bezuijen *et al*., 2002; Bezuijen *et al*., 2004; Stuebing *et al*., 2006). Previous studies on the use of mtDNA genes in genetic diversity for other crocodilians had been described. Most crocodilians show high genetic diversity, as seen in *Caiman crocodilus* (Vasconcelos *et al*., 2006; Venegas-Anaya *et al*., 2008) and *Melanosuchus niger* (Farais *et al*., 2004; Vasconcelos *et al*., 2008).

Mitochondrial DNA control region which is known to have repetitive regions containing variable number tandem repeats (VNTRs), had been used as the marker in crocodilian group were failed to be amplified in *Tomistoma* and *Gavialis* but were successfully described in other crocodilians (Ray and Densmore, 2002). As a result, new species-specific primers were developed to identify and characterize the repetitive region in *Tomistoma* (Kaur *et al*., 2012).

Information on the genetic diversity of *T. schlegelii* in Sarawak is still lacking. The only available genetic information in National Center for Biotechnology Information (NCBI) to date is on the cyt *b*-CR sequence of *T. schlegeli* (Kaur *et al*., 2012), although their study only involved 7 individuals from Sarawak in their analysis. In order to complement their findings, one sample of *T. schlegelii* was obtained from Asajaya in this study. The objective of this study is to determine the genetic variation among *T. schlegelii* in Sarawak using cytochrome *b*-control region (cyt *b*-CR) gene analysis.

2.0 Literature Review

2.1 Taxonomy

The taxonomy of *Tomistoma schlegelii* or known as false gharial is as shown as in Figure 2.1:

Kingdom: Animalia

Phylum: Chordata

Class: Reptilia

Order: Crocodylia

Family: Crocodylidae

Subfamily: Tomistominae

Genus: *Tomistoma*

Species: *Tomistoma schlegelii*

Local name:

a) False gharial
b) Malayan gharial

Figure 2.1: Taxonomy of *Tomistoma schlegelii* (adapted from www.iucnredlist.org).

2.2 *Tomistoma schlegelii*

2.2.1 Description

T. schlegelii is one of the largest crocodilian with adult males ranging from 4.5 – 5.0 m while adults females are about 2.5 – 3.0 m long (Bezuijen *et al.*, 1998; Bezuijen *et al.*, 2010). The number of eggs in a clutch is up to 35 eggs (Bezuijen *et al.*, 1998; Bezuijen *et al.*, 2001). The snout of this species resembles the slender snout of the Indian gharials, *Gavialis gangeticus* but slightly wider in *T. schlegelii* and therefore is commonly known as the false gharials (Figure 2.2).

(a)

(b)

Figure 2.2: The difference of snout between false gharial and Indian gharial. (a) indicates the snout of false gharial while (b) indicates the snout of Indian gharial (both figures were adapted from www.arkive.org).

The genus name means ‘cutting mouth’ in Greek while the species is named after its Dutch discoverer, H. Schlegel. Juveniles are dark chocolate brown in colour with black banding on the tail, body and dark blotches at the jaw area. Their underbelly is dark, turns lighter and finally white as they grow older. Adults retain most of their juvenile pattern with the dorsal colouration turning brown to olive-green with dark cross bands (Bezuijen *et al*., 2001). Its diet is reported to include a variety of invertebrates like shrimps to small vertebrates such as fish, monkeys and small deers (Bezuijen *et al*., 2001).

2.2.2 Classification based on Morphology

Based on the morphological data, outward similarities are found on the rostrum and jaws of the *T. schlegelii* and *G. gangeticus*. However, when older fossils were considered, the rostrum, jaw and dental similarities are highly similar to *Crocodylus* and therefore *Tomistoma* closer to *Crocodylus* while *Gavialis* forms a basal taxon (Brochu, 2003). The phylogenetic relationship within Tomistominae has changed in 2005 due to the discovery of newer fossils. Previously, *T. schlegelii* was placed together with *T. lusitanica* on the same clade and they formed a sister taxa to *T. cairense* but later *Gavialosuchus eggenburgensis* was found to be closer to *T. schlegelii* than *T. lusitanica* (Brochu, 2003). More recently, *Toyotamaphimeia machikanensis* found in Japan is considered the closest to the sole surviving Tomistominae (Kobayashi and Tomida, 2005; Piras *et al*., 2007).

2.2.3 Previous Studies related to *Tomistoma*

All previous studies on *T. schlegelii* showed that wild caught *Tomistoma* of Malaysia were caught within peat swamps restricted to Sarawak and Peninsular Malaysia with sightings being rare since the early 1990s (Sebastian, 1993; Simpson *et al*., 1998). The occurrence of *Tomistoma* in Thailand is unconfirmed. Stuebing *et al*. (2006) reported that a lack of supporting data either suggests this species did not occur in Thailand or occurred in the southern region now under Malaysian territory. There is no supporting data for *Tomistoma* occurrence in Vietnam and Singapore (Sebastian, 1993).

Existing records have shown that *T. schlegelii* in Kalimantan and Sumatra, Indonesia can still be found in the wild (Bezuijen *et al*., 1998; Auliya *et al*., 2006). This species was thought to face extinction in northern and western Sumatra, but their appearance in Java is unclear (Stuebing *et al*., 2006). Recent analysis of *Tomistoma* in South Sumatra, Jambi and West Kalimantan Province showed a continuously decline in numbers from the mid 1990s to 2004 (Bezuijen *et al*., 2001; 2004).

2.2.4 Ecology and Habitat

2.2.4.1 Historical Distribution

Historical distribution range of Tomistominae covered northern Africa, Europe, mainland Asia and Japan during the early Eocene of which some were distributed at estuarine and coastal regions (Brochu, 1997). *Tomistoma schlegelii* was found in freshwater or peat swamps that ranged across a land mass formerly known as Sundaland (Auliya *et al*., 2006; Bezuijen *et al*., 1998; 2002; Brochu, 2003; Stuebing *et al*., 2006).

Approximately 17000 years ago and at 120 meters below present level of shorelines, Sumatra, Jawa, Peninsular Malaysia and Borneo were thought to be connected by the exposed Sunda Shelf (Voris, 2000) with four major river systems; the Malacca Straits River System, the Siam River System, North and East Sunda River System. These systems contained peat deposits indicating the occurrence of Pleistocene peat swamps and freshwater lakes especially off the east coast of the Malay Peninsular (Voris, 2000). Stuebing *et al*. (2006) suggested that it had a wider distribution before the rising sea levels at the late Pleistocene altered land mass configuration of this region limiting the *Tomistoma* habitat to West Malaysia, Borneo and south-central Sumatra. The destruction of habitat at present time is also forcing this species to move farther inland from coastal areas (Stuebing *et al*., 2006).

2.2.4.2 Current Distribution

Field surveys within South East Asia so far had located *T. schlegelii* in the peat swamps of Indonesia and Malaysia (Auliya *et al*., 2006; Bezuijen *et al*., 2002; Cox and Gombek, 1985; Sebastian, 1993; Simpson *et al*., 1998) (Figure 2.3). This species is known to be associated with peat swamps and rivers in Sumatra especially the South Sumatra and Jambi Provinces (Bezuijen *et al*., 1998; Bezuijen et al., 2002), Central Kalimantan in the Tanjung Putting National Park rivers (Auliya *et al*., 2006) and Danau Sentarum in West Kalimantan (Bezuijen *et al*., 2004).

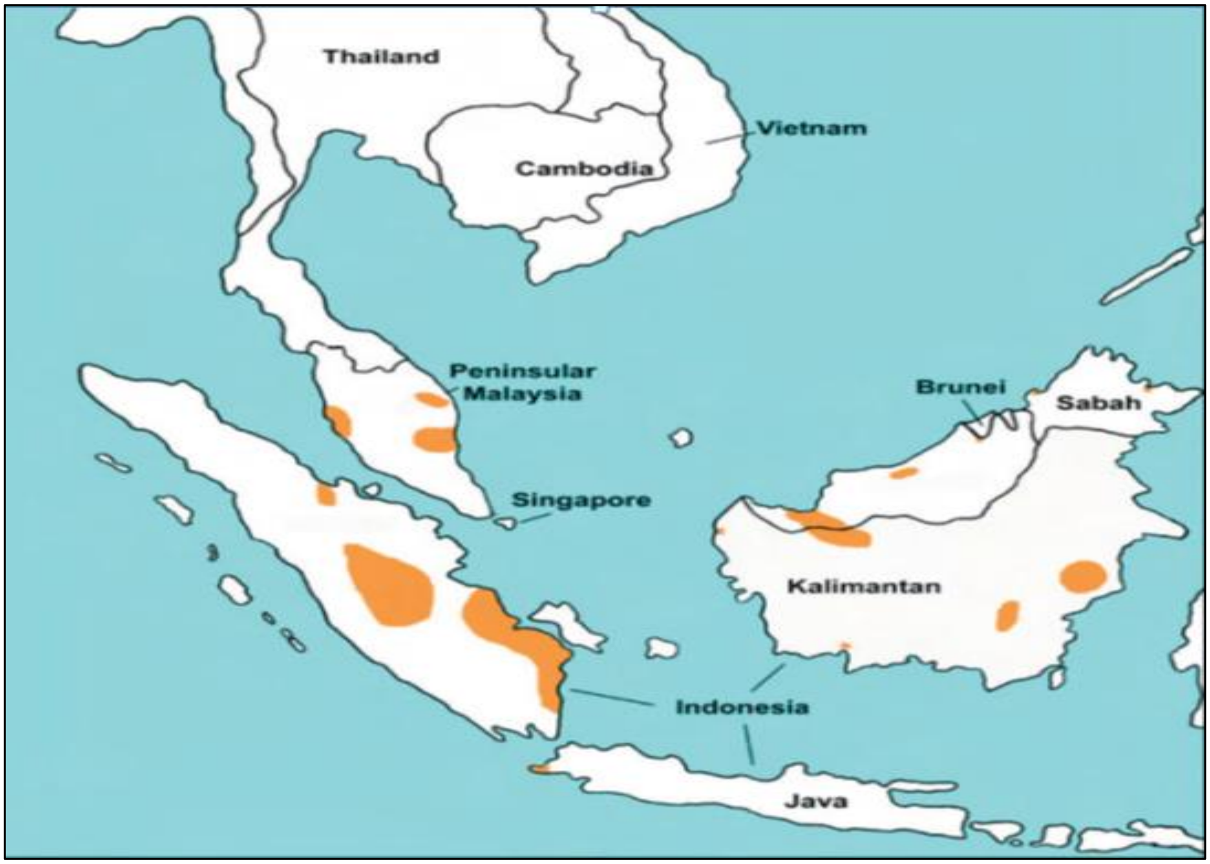

Figure 2.3: Distribution of *Tomistoma schlegelii* (Bezuijen *et al*., 2010).

2.3 Mitochondrial Genes

2.3.1 Mitochondrial DNA

Mitochondrial is a circular organelle DNA and is found in cells of higher organisms (Figure 2.4). It is responsible for oxidative phosphorylation and production of adenosine triphosphate, ATP (McBride *et al*., 2006). It contains DNA which is separated from the chromosomal or nuclear DNA and is maternally transmitted with low or no recombination (Avise, 2000). Mitochondria have been suggested for decades to be implicated in ageing (Harman, 1972) and it has been observed in a variety of organisms and tissues that mitochondria function declines with age (Navarro and Boveris, 2004; Marcineck *et al*., 2005; Hebert *et al*., 2010; Hartmann, 2011).

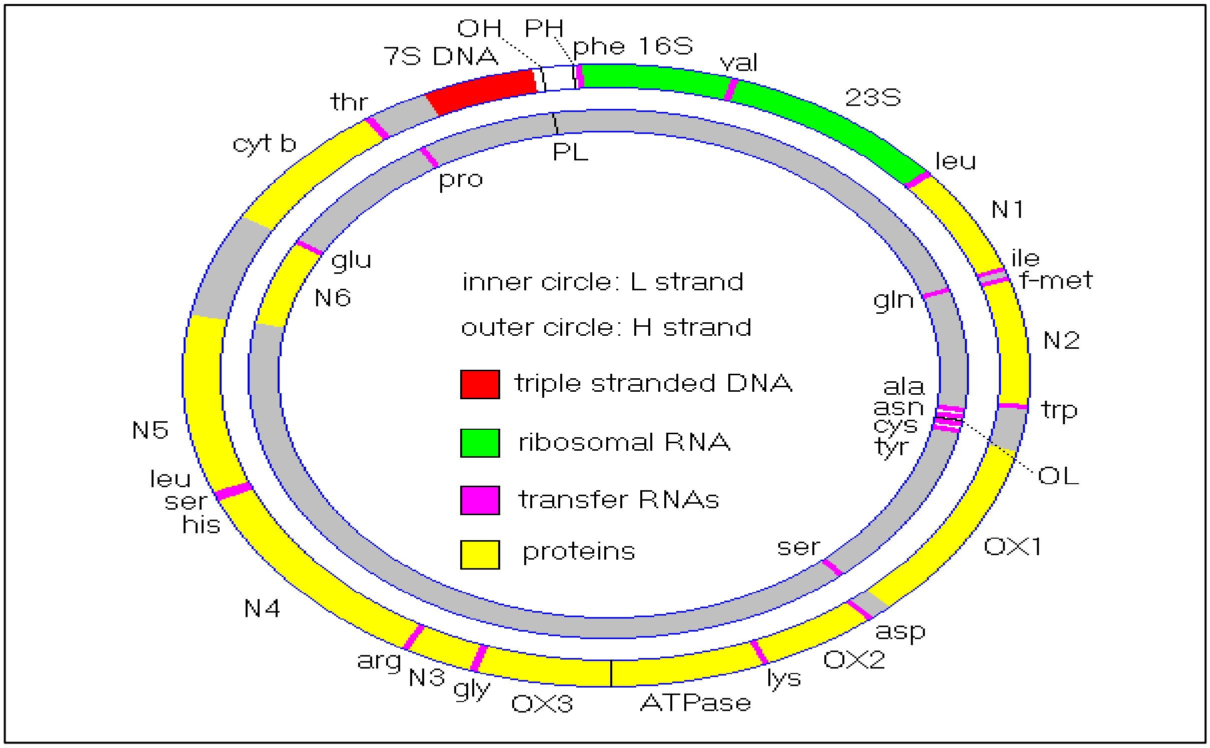

Figure 2.4: Mitochondrial DNA (adapted from www.bmb.leeds.ac.uk).

2.3.2 Cytochrome *b* Gene

The most common genome used in species identification is cytochrome *b* (cyt *b*) gene (Branicki *et al*., 2003) which had been used in phylogeny and forensic identification (Pancorbo *et al*., 2004). According to Johns and Avise (1998), cyt *b* gene is chosen because of several reasons: 1) the gene is most extensively sequenced to date, 2) the evolution of cyt *b* gene and the biochemistry of the protein products are better characterized than other molecular system, 3) cyt *b* gene is phylogenetically informative. The usefulness of cyt *b* gene analysis has been confirmed in identifying the biological origin of samples (Pancorbo *et al*., 2004). The location of cyt *b* gene on mtDNA is shown on the mtDNA gene map (Figure 2.5).

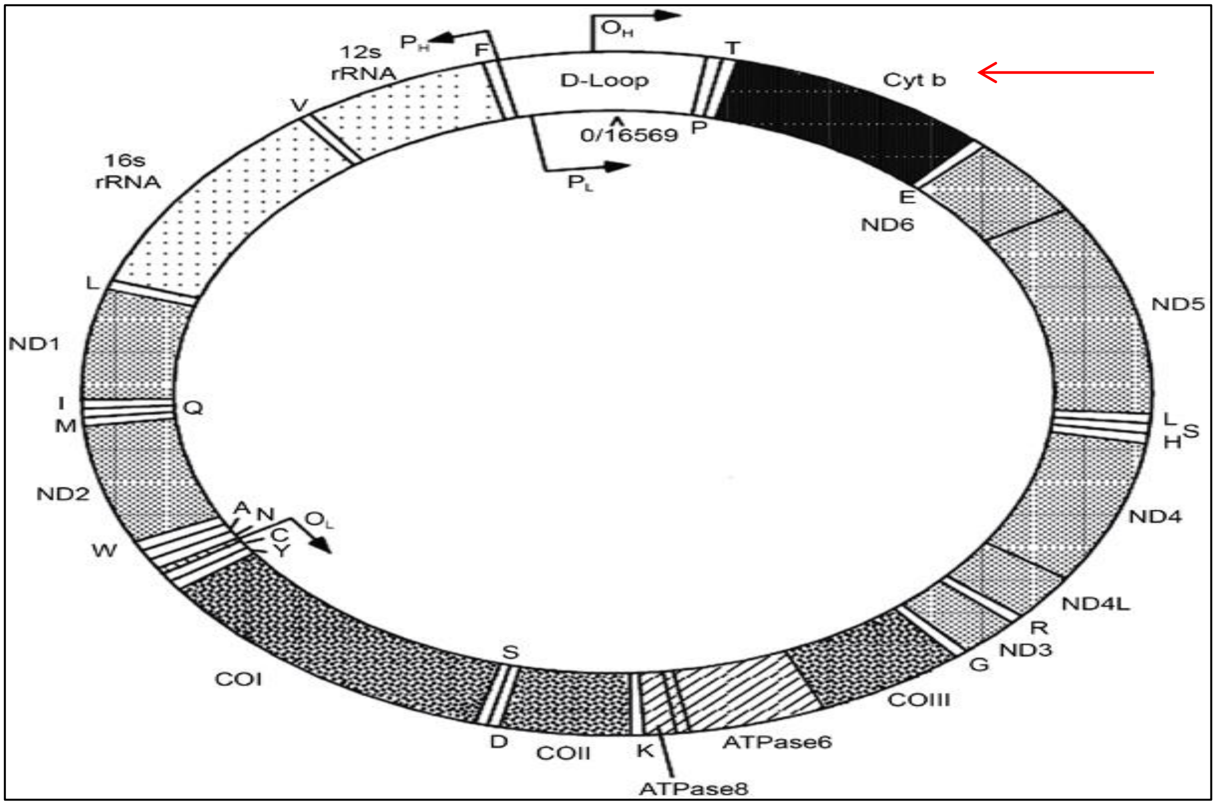

Figure 2.5: Location of cyt *b* gene on mtDNA (adapted from www.nature.com).

2.3.3 mtDNA Control Region

The control region is the only major non–coding segment in the vertebrate mitochondrial genome. The non–coding region known to be the most hypervariable region, evolving three to five times more rapidly compared with the rest of the mitochondrial genome and is not subjected to any selection making it suitable for population and species level analysis (Avise, 2000; Hwang and Kim, 1999). The non-coding region of most taxa consists of three parts known as Domain I, II and III with Domain I and III as the hypervariable domains with highly repetitive sequences (Sbisa *et al*., 1997). Although it is a hypervariable region, there are conserved sequences within this region found across taxa suggesting functional constraint in this region of the mtDNA genome.

The mitochondrial control region, which serves as the origin of replication for the mitochondrial genome and is usually more variable than other coding gene (Avise, 1994), has been proved to be an ideal marker for assessing genetic structure of recently diverged or closely related populations or species (Avise, 1994; Liu *et al*., 2009). The gene number and order is usually highly conserved among vertebrates and a number of mtDNA rearrangements have been described in birds, reptiles, amphibians and fishes (Boore, 1999; Hartmann, 2011).