

UNDERSTANDING DIVERGENCE IN *CHANNA STRIATA*:
A GENETIC AND MORPHOLOGICAL APPROACH

SONG LI MIN

DISSERTATION SUBMITTED IN FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR
2012

MEMAHAMI KECAPAHAN *CHANNA STRIATA*:
SATU PENDEKATAN GENETIK DAN MORFOLOGI

SONG LI MIN

DISERTASI YANG DIKEMUKAKAN UNTUK MEMENUHI SYARAT
MEMPEROLEHI IJAZAH
SARJANA SAINS

FAKULTI SAINS
UNIVERSITI MALAYA
KUALA LUMPUR
2012

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Song Li Min (I.C No: 861216-04-5174)
Registration/Matric No: SGR 100082
Name of Degree: Master of Science
Title of Dissertation: Understanding divergence in *Channa striata*: a genetic and morphological approach
Field of Study: Population Genetics

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature

Date: 18 August 2012

Subscribed and solemnly declared before,

Witness's Signature

Date:

Name:

Designation:

ABSTRACT

The snakehead murrel, *Channa striata*, or “Haruan” in Malaysia, is well known with its pharmaceutical purpose and hence is extensively studied for its biochemical compound involved in medication while its knowledge on population structure is poorly inventoried. This study was conducted to characterize the naturally occurring populations of *C. striata* using both molecular and morphological approaches. These specimens were extensively caught from wide geographical range across Malaysia in the aim to propose spatial conservation assessment on this species as well as to investigate the possible hidden cryptic species of *C. striata* in Malaysia. Our data evidence the presence of variation in *C. striata* with the molecular divergence pattern is not congruent to the morphological divergence pattern. In this study, for example, population from Negeri Sembilan and Johor were genetically alike; however, these two populations together with population from Pulau Pinang were characterized by a unique head size in relation to others indicating that the physical habitat heterogeneity is probably the trait deformity agents. This suggests that environment conditions such as food source, changes in water temperature, contaminated organic compound and other possible environment factors will influence the development of fish trait which need not correlated with underlying genetic differences. Besides, our molecular data on DNA barcoding approach did not revealed possible cryptic species of *C. striata* among all studied sites in Malaysia although this species shows high population divergence in South-East Asia. This is supported by a monophyletic divergence pattern of *C. striata* reflected in the phylogenetic tree inferred using both NJ and Bayesian inference method with high statistical support. Nevertheless, another supportive results of low intra-species genetic distance (0.60%) compared to its con-generic distance (1.70%) and this

intra-species variation (0.60%) was less than the threshold value (1.60%) under Hebert's *et al.*, 2004a 10X benchmark, strongly suggest that there is no possible existence of hidden cryptic species at the present study. The molecular mtDNA data in this research study were able to illustrate the historical dispersal pattern of *C. striata*. Although physical barrier separation tends to cause population divergence, for example, unique haplotype was found at populations in the central division of Peninsular Malaysia which are separated by physical barrier; however, the historical coalescence of this candidate species was mainly explained by geomorphologic structure and ancient natural climatic changes that affect the area of expose land mass. The unexpected overlapping haplotype distribution and non-significant population difference between the mainland and the island of Borneo are most probably explained by the incident of historical connection during Pleistocene Epoch. The understood population divergence pattern revealed in this study provides valuable information for subsequent spatial conservation planning based on different aspects and also informative to ecologist to further study the water quality on this candidate's morphology as *C. striata* tends to be a potential bio-marker in respond to the changes in dissolved chemical compounds.

ABSTRAK

Murrel snakehead, *Channa striata*, atau dikenali sebagai “Haruan” di Malaysia terkenal dengan tujuan farmaseutikal dan dengan itu banyak dikaji bagi sebatian biokimia yang terlibat dalam perubatan manakala pengetahuan mengenai struktur populasi jarang dikaji. Di sini, dengan menggunakan dua pendekatan, iaitu molekul dan morfologi, kajian kami adalah mencorakkan ciri-ciri populasi *C. striata* yang mewakili populasi liar yang telah dikutip secara meluas di seluruh Malaysia bertujuan untuk mencadangkan pemuliharaan ruang pada spesies ini serta menyiasat kemungkinan spesies samar *C. striata* yang tersembunyi di Malaysia. Bukti data kami menunjukkan kewujudan variasi di *C. striata* dengan corak kecapahan molekul tidak kongruen dengan corak kecapahan morfologi. Dalam kajian ini, sebagai contoh, populasi dari Negeri Sembilan dan Johor adalah serupa dari segi genetik, namun, kedua-dua populasi ini bersama-sama dengan populasi Pulau Pinang telah dicirikan oleh saiz kepala yang unik berbanding dengan populasi lain menunjukkan bahawa kepelbagaian habitat fizikal merupakan kemungkinan ejen yang mengakibatkan deformasi trait. Ini menunjukkan bahawa keadaan persekitaran seperti sumber makanan, perubahan dalam suhu air, kompaun organik yang tercemar dan faktor persekitaran lain akan mempengaruhi perkembangan sifat ikan yang tidak perlu selari dengan asas perbezaan genetik. Selain itu, data molekul kami atas pendekatan DNA barkod tidak mendedahkan kemungkinan spesies samar *C. striata* di kalangan populasi yang dikaji di Malaysia walaupun spesies ini menunjukkan kecapahan populasi yang tinggi di Asia Tenggara. Kenyataan ini disokong oleh corak kecapahan monophyletic *C. striata* yang dicerminkan oleh pokok filogenetik disimpulkan dengan kedua-dua NJ dan kaedah Bayesian dengan sokongan statistik yang tinggi, jarak genetik antara-spesies (0.60%) yang

rendah berbanding dengan jaraknya secara con-generic (1,70%) dan variasi antara-spesies ini (0.60%) adalah kurang daripada nilai ambang (1.60%) di bawah Hebert's *et al*, 2004a 10X penanda aras. Molekul data mtDNA menunjukkan corak penyebaran sejarah *C. striata* di mana corak migrasi ini telah diramalkan berdasarkan kemungkinan sisi penyebaran *C. striata*. Walaupun pemisahan halangan fizikal cenderung untuk menyebabkan kecapahan populasi, sebagai contoh, haplotype yang unik telah ditemui di antara populasi yang dipisahkan oleh halangan fizikal iaitu populasi di bahagian tengah Semenanjung Malaysia, namun, tautan sejarah spesies ini terutamanya akan dijelaskan oleh struktur geomorfologik dan purba semula jadi perubahan iklim yang memberi kesan kepada kawasan tanah yang didedah. Pertindihan haplotype yang sama dan perbezaan populasi yang tidak signifikansi antara Semenanjung Malaysia dan Pulau Borneo secara tidak dijangka kemungkinan besar adalah bukti sambungan bersejarah semasa Epoch Pleistosen. Corak kecapahan penduduk yang difahamkan dalam kajian ini memberikan maklumat yang berharga untuk perancangan pemuliharaan ruang yang berikutnya berdasarkan aspek-aspek yang berbeza dan juga akan memberi maklumat kepada ahli ekologi untuk terus mengkaji kualiti air pada morfologi *C. striata* disebabkan spesies ini cenderung untuk menjadi potensi bio-penanda dalam bertindak balas terhadap perubahan bahan kimia di dalam air yang tercemar.

ACKNOWLEDGEMENT

First and foremost, a deepest thanks to my supervisor, Dr. Subha Bhassu for her guidance and contribution. Her valuable guidance was lead to the succession of my research study. An especially thanks to University of Malaya for research funding to continuity supporting this research project financially (Grant No: PV071/2011A).

I would like to acknowledge Department of Fisheries Malaysia, Freshwater Fisheries Research Division, Jelevu for their provision of samples. I also thank to Department of Fisheries Malaysia of Perak state for their permission to access the Temenggor Lake for sampling. A deeply thanks to all the staff of Fisheries and my lab member, Kaviarasu Munian during the field trip and their physical support are deeply appreciated.

A special thanks to all the Molecular Markers Laboratory members and post-doc for their encouragement and support. Especially thanks to the post-doc, Eleanor Adamson for her guidance and contribution during the discussion of statistical analyses.

Nevertheless, a deepest thanks to my dearest mother, Lam Yet Hong, and my sister, Song Si Min and also all the family members for their care and concern during this study. Their support and understanding had encouraged me during the completion of the thesis. I hereby dedicated this thesis to them and to those who assisted me during my study.

TABLE OF CONTENTS

	PAGE
ABSTRACT	iv
ABSTRAK	vi
ACKNOWLEDGEMENTS	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF SYMBOLS & ABBREVIATIONS	xv
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	7
2.1: VARIATION IN NATURAL POPULATION	7
2.1.1: INTRODUCTORY VARIATION	7
2.1.2: SOURCE OF VARIATION	8
2.1.3: ASSESSMENT OF VARIATION	10
2.1.4: PHYLOGEOGRAPHY	13
2.2: FRESHWATER FISH	15
2.2.1: STATISTICS ON FRESHWATER FISH	15
2.2.2: TAXONOMIC CLASSIFICATION OF FISH	16
2.2.3: <i>Channa striata</i>	18
2.3: MORPHOLOGICAL APPROACH	23
2.3.1: TRADITIONAL TAXONOMIC CLASSIFICATION OF SPECIES	23
2.3.2: MULTIVARIATE ANALYSIS	25
2.4: GENETIC APPROACH	26
2.4.1: DNA BARCODING	26
2.4.2: DNA BARCODING AND DISCOVERY OF CRYPTIC SPECIES	29
2.4.3: DNA BARCODING AND CONSERVATION BIOLOGY	31
2.5: MITOCHONDRIAL GENOME	33
2.5.1: MITOCHONDRIA	33
2.5.2: DISCOVERY OF NUMT IN MITOCHONDRIA	34
2.5.3: MITOCHONDRIA AS MOLECULAR SPECIES TAQ	36

2.6: CYTOCHROME C OXIDASE SUBUNIT I (COI)	39
2.6.1: INTRODUCTION OF COI	39
2.6.2: COI AS DNA BARCODE	40
CHAPTER 3: METHODOLOGY	42
3.1: SAMPLING LOCATION	42
3.2: GENOMIC ANALYSIS	47
3.2.1: DNA EXTRACTION	47
3.2.2: POLYMERASE CHAIN REACTION (PCR)	50
3.2.3: GEL ELECTROPHORESIS	51
3.2.4: GEL PURIFICATION	52
3.2.5: DNA SEQUENCING	54
3.2.6: MOLECULAR DATA ANALYSIS	55
3.3: MORPHOMETRIC ANALYSIS	57
CHAPTER 4: RESULT	60
4.1: GENETIC AMPLIFICATION	60
4.2: DNA BARCODING	61
4.3: MOLECULAR DATA (COI) OF <i>C.striata</i>	67
4.4: MORPHOMETRIC DATA OF <i>C.striata</i>	71
CHAPTER 5: DISCUSSION	80
5.1: DISCOVERY OF CRYPTIC SPECIES	80
5.2: MOLECULAR POPULATION DIVERGENCE (COI)	83
5.2.1: GENETICS DISPERSAL OF <i>C. striata</i>	83
5.2.2: PHYLOGEOGRAPHIC STRUCTURE	88
5.3: PREVIOUS MICROSATELLITE STUDY	91
5.3.1: HIDDEN BIODIVERSITY OF <i>C. striata</i>	91
5.3.2: RECENT DISPERSAL PATTERN OF <i>C. striata</i>	92
5.4: MORPHOLOGICAL POPULATION DIVERGENCE	93
5.4.1: ECOLOGICAL HETEROGENEITY	93
5.4.2: FEEDING BEHAVIOR AND PREDATION	95
5.5: CONSERVATION ASSESSMENT	97

CHAPTER 6: CONCLUSION	99
REFERENCES	101
APPENDIX	xvii
APPENDIX A	xvii
APPENDIX B	xxii
APPENDIX C	xxiii
APPENDIX D	xxiv
APPENDIX E	xxv

LIST OF TABLES

TABLE		PAGE
3.1	List of 27 selected freshwater fishes species used in DNA barcoding study. Numbers and letters under location refer to Figure 3.1 .	45
3.2	List of <i>Channa striata</i> specimens used in both molecular and morphometric analysis.	46
4.1	List of 46 unique haplotypes generated using software DnaSP v5 (Librado & Rozas, 2009).	63
4.2	Comparison of divergence distance (K2P percent) calculated using K2P model.	65
4.3	Pairwise population distance represented by Fst value calculated using mtDNA COI sequence with bold and “*” indicate significant Fst ($p < 0.05$).	70
4.4	Range of morphometric measurement of all 38 characters of <i>Channa striata</i> measured to the nearest centimeter with value in parentheses indicate mean of each measurement.	73
4.5	List of standardized coefficients of canonical variables based on all <i>Channa striata</i> populations surveyed in Peninsular Malaysia. Measurements in “bold” and “underline” indicate greatest contribution of characters to each root. The variables are corresponding to Truss Network Measurement illustrated in Figure 3.2 .	76
4.6	List of standardized coefficients of canonical variables based on all <i>Channa striata</i> populations surveyed across entire Malaysia. Measurements in “bold” and “underline” indicate greatest contribution of characters to each root. The variables are corresponding to Truss Network Measurement illustrated in Figure 3.2 .	79
Appendix A	List of specimen details of 27 freshwater fish species used in this study.	xvii
Appendix B	A list of optimum PCR mixture used in this study.	xxii
Appendix C	A list of optimum thermal cycling profile used in this study.	xxiii
Appendix D	List of components of 10X TBE buffer	xxiv

LIST OF FIGURES

FIGURE		PAGE
2.1	Map of Pleistocene sea level for Southeast Asia and Austral-Asia with depth contour of 120m below present level. Figure cited in Voris (2000).	14
2.2	Taxonomic classification of fishes from the common domain of Eukaryota up to different classes whereby diversification of fish species started from each class.	17
2.3	Taxonomic classification of <i>C. striata</i> .	19
2.4	Image of <i>Channa striata</i> . Image obtained from Department of Fisheries Malaysia, Freshwater Fisheries Research Division, FRI Glami Lemi, Jelebu, 71650 Titi, Negeri Sembilan, Malaysia; www.pppat.gov.my; ffrc@pppat.gov.my.	20
2.5	Distribution map of <i>Channa striata</i> in Asia with darken regions indicate its natural geographical ranges.	20
2.6	Statistic for the global capture and global aquaculture of <i>Channa striata</i> from the year of 1950 to 2007 (FAO 2010).	22
2.7	Diagrammatic representation of the “conserved” vertebrate mitochondrial genome (Pereira, 2000).	33
2.8	Diagrammatic representation of selected 5’ end of mitochondrial cytochrome <i>c</i> oxidase subunit 1 (COI) gene as the barcode region (Waugh, 2007).	38
3.1	Sampling locations of 220 specimens of <i>Channa striata</i> and additional 83 freshwater fishes constituting selected 26 freshwater fish species used in this study. Exact latitude and longitude for each sampling sites are being recorded. Number 1-8 and alphabet A-D correspond to locations listed in Table 3.1 .	44
3.2	(a) & (b): Schematic diagram of 38 morphometric measurement of <i>Channa striata</i> (Total length (tl), standard length (sl), caudal fin length (cfl), eyes diameter (ed), snout width (SnW), snout length (SnL), measurement of each body part (a1-a12, b1-b4, c1-c3, d1-d2, e1-e2, f1-f2, g1-g4, h1-h3)) based on the Truss Network Measurement (Bhassu & Rashid, 2009).	59

4.1	Phylogenetic tree of mtDNA COI barcodes of 27 fish species with number of individuals of each haplotype is listed in brackets after the species name. Posterior probability and bootstrap among different haplotypes of same species are shown by * and ` respectively.	66
4.2	(a): Haplotype distribution of seven haplotypes generated from eight populations of <i>Channa striata</i> across Malaysia. (b): Haplotype network that connecting all generated haplotypes; the size of each haplotype is proportional to the number of each individual of corresponding haplotype; each line represents one mutational step.	69
4.3	Multivariate discriminant analysis represented by root 1 and root 2 axis based on all <i>Channa striata</i> populations surveyed in Peninsular Malaysia.	75
4.4	Multivariate discriminant analysis represented by root 1 and root 2 axis based on all <i>Channa striata</i> populations surveyed across entire Malaysia.	78
5.1	Distribution of haplotypes over entire Malaysia with predicted possible migration routes of <i>C. striata</i> . Details on the stages of migration routes can be referred to section 5.2.1 .	87
Appendix E	Banding pattern of COI amplicon with optimized annealing temperature in the range of 45 – 55° and optimum amount of 25mM MgCl ₂ of 2µl per reaction volume.	xxv

LIST OF SYMBOLS & ABBREVIATIONS

OD _{260/280nm}	Absorbency at wavelength 260nm/280nm
BOLD	Barcode of Life Data Systems
CBOL	Consortium for the Barcode of Life
bp	Base pair
°C	Degree Celsius
DNase	Deoxynuclease
dNTP	2'-deoxynucleoside 5'-triphosphate
DNA	Deoxyribonucleic acid
dH ₂ O	Desterilized water
EDTA	Ethylediamine tetraacetic acid
FAO	Food and Agriculture Organization
g	Gram
indel	Insertion and deletion
MgCl ₂	Magnesium chloride
Mg ²⁺	Magnesium ions
µg	Microgram
µl	Microlitre
µM	Micromolar
ml	Mililitre
mM	Milimolar
min	Minutes
M	Molar

ng	Nanogram
nm	Nanometer
et al.	Others
%	Percentage
PCR	Polymerase chain reaction
RNase	Ribonuclease
RNA	Ribonucleic acid
rpm	Rotation per minute
sec	Seconds
UV	Ultraviolet
V	Volume
v/ml	Volume per milliliter
w/v	Weight per volume