GENETIC VARIATION AND DIFFERENTIATION BETWEEN CULTURED AND WILD POPULATIONS OF *CHANNA STRIATA* (HARUAN) FROM MICROSATELLITE MARKERS

JOTHI ANNAVADIVOO A/P BALASINGAM

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

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2011

ABSTRACT

The snakehead fish (*Channa striata*) is a freshwater fish species indigenous to Malaysia. *Channa striata* (*C. striata*), lo cally known as *Haruan* is valued as a natural remedy in traditional medicine as well as a reputable source of protein. The rapid progress of scientific validation research on the therapeutic properties of the species has spurred the equally rapid expansion of commercialised products that capitalises on these unique yet beneficial qualities of *C. striata*. Dependence on the wild in order to meet the increasing demand on the species is no longer a feasible option whereby, deteriorating natural abundance due to anthropogenic effects and over-exploitation has left an opening for the potential growth of the aquaculture industry of *C. striata* in Malaysia.

The present study employed seven polymorphic microsatellite loci to investigate levels of genetic variation and differentiation of *C. striata* at selected cultured and wild populations in peninsular Malaysia. The study included cultured populations from three separate commercial farms (Kajang, Malacca and Rawang) as well as wild populations from three different states; Johore, Kedah and Pahang located at southern-, northern- and eastern-region of the peninsular, respectively. The results showed that cultured populations (mean number of alleles per locus, A = 7.71–9.29; allelic richness, A_R = 6.752–8.108; mean effective number of alleles per locus, n_e = 3.31–4.38; observed heterozygosity, H_o = 0.24–0.98; expected heterozygosity, H_e = 0.33–0.88) had significantly higher genetic variation relative to the wild (A = 3.86–4.86; A_R = 3.843–4.838; n_e = 2.01–2.20; H_o = 0.07–0.90; H_e = 0.07–0.69) populations. However, AMOVA analysis revealed that the greater percentage of variation (79.2%) in the total genetic diversity of the surveyed populations is primarily due to differences at the

individual level and neither between, nor within, pooled cultured and pooled wild group.

Departure from Hardy Weinberg Equilibrium (HWE) was observed in all cultured populations and the wild population of Kedah. Isolated cases of inbreeding and recent population bottleneck were detected among the cultured and wild populations. The results also displayed low to extensive genetic differentiation among the cultured and wild populations as revealed by pair-wise measures; F_{ST}, R_{ST} and D_{est}. AMOVA analysis showed slight divergence (F_{CT} = 0.083) between pooled cultured and pooled wild groups. A positive correlation was inferred between population genetic delineation and historical patterns of gene flow of *C. striata* in the wild populations.

Genetic variation along with Bayesian clustering (STRUCTURE) analyses indicate assayed cultured populations are highly admixed. It is implied that the cultured populations were founded from sufficient number of broodstocks. Based on the investigation, the domestication process is in an early stage. Further studies are needed for comprehensive determinations of genetic varieties of current broodstocks and successive offsprings of the cultured populations with increased number of *C. striata* sample collections.

ABSTRAK

Ikan snakehead (*Channa striata*) ialah sejenis spesies ikan air tawar asli di Malaysia. *Channa striata* (*C. striata*) atau nama tempatannya, ikan Haruan, adalah penting sebagai remedi semulajadi dalam bidang perubatan tradisional serta merupakan sumber protein yang dipercayai dan bereputasi. Di samping itu, perkembangan pesat dalam pembangunan produk-produk komersial *C. striata* telah berlangsung sejajar dengan kemajuan dalam penyelidikan pengesahan saintifik berkenaan dengan ciri-ciri terapeutik *C. striata*. Walaubagaimanapun, pergantungan kepada spesies liar *C. striata* untuk memenuhi permintaan yang semakin meningkat tidak lagi merupakan pilihan yang rasional dimana, kemerosotan habitat semulajadi yang disebabkan oleh kesan–kesan antropogenik serta pengeksploitasian alam tanpa batasan telah membuka peluang kepada pertumbuhan industri akuakultur *C. striata* di Malaysia.

Penyelidikan semasa ini melibatkan tujuh lokus mikrosatelit polimorfik bertujuan untuk menyiasat tahap kepelbagaian dan pembezaan genetik *C. striata* dalam populasi kultur dan populasi liar terpilih di Semenanjung Malaysia. Kajian ini terdiri daripada populasi-populasi kultur dari tiga ladang ternakan ikan komersial (Kajang, Melaka dan Rawang) bersertakan populasi-populasi liar dari tiga buah negeri berlainan, iaitu; Johor, Kedah dan Pahang yang masing-masing terletak di bahagian selatan, utara dan timur Semenanjung Malaysia.

Hasil kajian menunjukkan bahawa populasi kultur (purata nombor alel per lokus, A = 7.71–9.29; kekayaan alel, A_R = 6.752–8.108; purata efektif nombor alel per locus, n_e = 3.31–4.38; heterozigositi cerapan, H_o = 0.24–0.98; heterozigositi jangkaan, H_e = 0.33–0.88) mempunyai signifikan variasi genetik yang lebih tinggi berbanding dengan

populasi liar (A = 3.86–4.86; A_R = 3.843–4.838; n_e = 2.01–2.20; H_o = 0.07–0.90; H_e = 0.07–0.69). Namun demikian, analisis AMOVA mendedahkan bahawa sebahagian besar (79.2%) daripada jumlah variasi yang ditemui dalam populasi-populasi yang dikaji adalah disebabkan oleh perbezaan pada peringkat individu dan bukan di antara, atau dalam kumpulan populasi kultur dan populasi liar *C. striata*.

Ketidakpatuhan dari Keseimbangan Hardy Weinberg (HWE) telah diperhatikan dalam populasi liar Kedah dan semua populasi kultur dalam kajian ini. Tambahan pula, kejadian pembiakbakaan dalam (inbreeding) dan populasi "bottleneck" telah dikesan di kalangan populasi kultur dan liar. Index-index anggaran pembezaan genetik berpasangan; F_{ST}, R_{ST} and D_{est} mempamerkan nilai-nilai yang rendah hingga tinggi di antara populasi kultur dan liar. Disamping itu, analisis AMOVA menunjukkan perbezaaan kecil (F_{CT} = 0.083) di antara populasi kumpulan kultur and populasi kumpulan liar. Korelasi positif dapat disimpulkan di antara pembezaan genetik populasi dengan corak sejarah aliran gen *C. striata* dalam populasi liar.

Analisa kluster Bayesian (STRUCTURE) bersama dengan kepelbagaian genetik menunjukkan bahawa populasi kultur yang dikaji adalah amat bercampur-aduk (admixed). Ini bermaksud bahawa populasi kultur *C. striata* dalam siasatan ini diasaskan oleh bilangan stok induk (broodstock) yang memadai. Hasil kajian ini juga mengimplikasikan bahawa proses pendomestikan ikan *C. striata* masih berada di peringkat awal. Penyelidikan lanjut perlu dijalankan untuk mendapatkan pemahaman yang komprehensif mengenai kepelbagaian genetik stok induk populasi kultur *C. striata* sekarang dan generasi progeni yang berturutan dengan menambahkan bilangan sampel yang sedia ada.

ACKNOWLEDGEMENTS

It would not have been possible to write this dissertation without the help and support of the kind people around me. Above all, I would like to thank my parents, for their unconditional love, patience, support and unwavering faith in me. I dedicate this thesis to my father, Mr Balasingam, who taught me the importance of education, of having strong values, perseverance, endurance and resilience in overcoming life's many obstacles. My siblings, Lavania, Vicki and Priya albeit in their own ways, has each inspired me to rise up to the challenge.

This dissertation would not have been possible without the crucial contribution, encouragement and motivation from my principal supervisor, Dr Subha Bhassu. She is a dedicated and passionate supervisor and is invaluable on both an academic and a personal level. I also wish to express my warm and sincere thanks to Assoc Prof Dr Zul for his valuable advices and offers of help.

I owe my most sincere gratitude to Dr Eleanor "Ellie" Adamson, for her extensive discussions around my data analysis and interpretations; her interesting explorations and insights that have been really helpful in this study. My senior Yasmeen Salah, whom I am forever grateful to, for enlightening me about the intricate workings of a researcher. A truly beautiful soul with a compassionate nature. My "partner in crime", Shairah "Shira" Abd. Razak, I truly treasure her support and the times that we worked together, digging our way out together towards the LIGHT. We've made it !! Appreciation also goes out to Dr Teh Ser Huy, for her helpful technical expertise and guidance on the uber expensive capillary sequencer.

Last but by no means least, I would like to thank my lab-mates and makan-kakis; Li Min, Tian Tian, Shaun, Iszam, Peter, Saras, Neena, Dr Mahfuj for their guidance and that life in the lab is never dull without you guys. I am also grateful to my junior, Arena without whose assistance this study would not have been successful.

TABLE OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATION

µg	microgram
mg	milligram
µL	microlitre
mL	millilitre
ρmole	picamole
™	trademark
°C	degree Celsius
10×	ten times
1×	one time
A	mean number of alleles per locus
AMOVA	Analysis of Molecular Variance
A_R	allelic richness
bp(s)	basepair(s)
c	centi
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
dNTP	deoxyribonucleotide triphosphate
dTTP	deoxythymidine triphosphate
ddH_2O	double distilled water
DNA	deoxyribonucleic acid
FDR	False Discovery Rate
F_{IS}	Inbreeding Coefficient
F_{ST}	Fixation Index
g(s)	gram(s)
h	hour
H_e	Expected heterozygosity
H_O	Observed heterozygosity
$\bar{H}$	Mean heterozygosity
HWE	Hardy-Weinberg Equilibrium
IAM	Infinite Allele Mutation model
K	Kilo
LD	linkage disequilibrium
m(s)	metre(s)
M	molar
mA	milliampere
$MgCl_2$	magnesium chloride
n_a	observed number of alleles
n_e	effective number of alleles
N_e	effective population size
NJ	Neighbor-Joining
PCR	Polymerase Chain Reaction

PIC	Polymorphism Information Content
T_a	annealing temperature
rpm	rotations per minute
s	second
SMM	Stepwise Mutation model
SD	standard deviation
SSR	simple sequence repeat
UPGMA	Unweighted Pair-Group Method of Arithmetic Averages
V	volt