THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Doctor of Philosophy (Genetics)

DEGREE

Genetics	Genetics
FIELD	**DEPARTMENT**

TITLE: Genetic Variability and Differentiation of Whisker Sheatfish *Micronema bleekeri* (Günther, 1864) Populations Based on Microsatellite Marker

NAME: Mrs. Piyamas Phongkaew

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR

(Assistant Professor Vipa Hongtrakul, Ph.D.)

THESIS CO-ADVISOR

(Associate Professor Nitsri Sangduen, Ph.D.)

THESIS CO-ADVISOR

(Associate Professor Amara Thongpan, Ph.D.)

THESIS CO-ADVISOR

(Miss Uraiwan Arunyawat, Dr.rer.nat.)

DEPARTMENT HEAD

(Associate Professor Somsak Apisitwanich, Dr.Agr.Sci.)

APPROVED BY THE GRADUATE SCHOOL ON _____

DEAN

(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

GENETIC VARIABILITY AND DIFFERENTIATION OF WHISKER SHEATFISH *Micronema bleekeri* (Günther, 1864) POPULATIONS BASED ON MICROSATELLITE MARKER

PIYAMAS PHONGKAEW

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy (Genetics)
Graduate School, Kasetsart University
2012

Piyamas Phongkaew 2012: Genetic Variability and Differentiation of Whisker Sheatfish *Micronema bleekeri* (Günther, 1864) Populations Based on Microsatellite Marker. Doctor of Philosophy (Genetics), Major Field: Genetics, Department of Genetics. Thesis Advisor: Assistant Professor Vipa Hongtrakul, Ph.D. 114 pages.

Microsatellite markers were developed for whisker sheatfish using the enrichment strategy. Forty-two primer pairs were designed at the flanking regions of microsatellite sequences. Thirteen microsatellite markers were polymorphic and eleven markers were selected to investigate genetic variation of seven whisker sheatfish populations, namely the Nong-Han Lake (NH), Nam Kum basin (KB), the Mekong river at the lower area of Nakhon Phanom province (MK), the Mekong river at the upper area of Nakhon Phanom province (MG), Songkhram river basin at Srisongkharm district (SK), Songkhram river basin at Seka district (SE) and Lam Nam Oon (NA). Five whisker sheatfish populations (NH, KB, MG, SE, NA) were conformed to Hardy-Wienberg and linkage equilibrium, whereas the remaining two populations (MK, SK) exhibited Wahlund effect. The average number of alleles per locus for each population ranged from 2.36 (NH) to 5.54 (MK). Mean allelic richness varied from 3.33 (NH) to 5.27 (MK), while NA exhibited two private alleles at MB401 and MB618 loci. The observed heterozygosity ranged from 0.223 (KB) to 0.415 (MK), whereas the expected heterozygosity varied from 0.237 (KB) to 0.657 (MK). F_{IS} had positive values in NH, KB, MG and SE populations, indicating inbreeding effect. Five whisker sheatfish populations exhibited a great population differentiation with the mean F_{ST} of 0.263. The decline in recent migration rate of whisker sheatfish and small effective population size were observed. Geographical barrier and over-exploitation were concerned for whisker sheatfish conservation and management in order to prevent the genetic erosion and to sustain this valuable species.

_____ _____ __ / __ / __

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Assistant Professor Dr. Vipa Hongtrakul, and co-advisors, Associate Professor Dr. Nitsri Sangduen, Associate Professor Dr. Amara Thongpan and Dr. Uraiwan Arunyawat for their guidance and helpful assistance on my research work.

I gratefully thank Associate Professor Dr. Lerdluck Ngernsiri, Assistant Professor Dr. Chuta Boonphakdee, Professor Pradit Pongthongkum, Associate Professor Dr. Surin Peyachockanakul, Dr. Akarapong Swatdipong and all the lecturers at the Department of Genetics for their comment and suggestion on my thesis. I would like to thank Professor Dr. Uthairat Na-Nakorn, Department of Aquaculture, Faculty of Fisheries, for her technical advice.

I also would like to thank Mr. Pramook Rukaewma and the officers of Sakon Nakhon Inland Fisheries Research and Development Center, Department of Fisheries, for their help in collecting specimens. I wish to thank my colleagues at the Faculty of Natural Resources and Agro-Industry, and the Department of Genetics for their friendliness and helpfulness.

This work was granted by the Office of the Higher Education Commission (CHE490140), and Kasetsart University Research and Development Institute (2007-2008).

Most of all, I would like to gratefully thank and I am deeply indebted to my father, my mother and my sister for their love, kindness and encouragement during this long ordeal. Finally, I deeply appreciate Major Weerachai Phongkaew, my husband, who devoted time and support and also my daugther who always gives me heartfelt love during my graduate study.

Piyamas Phongkaew

April 2012

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

AFLP	=	Amplified Fragment Length Polymorphism
T_a	=	annealing temperature
bp	=	base pairs
$^{\circ}C$	=	degree Celcius
dNTP	=	deoxynucleotide triphosphate
DNA	=	deoxyribonucleic acid
ddH_2O	=	double-distilled water
N_e	=	effective population size
E. coli	=	*Escherichia coli*
EtOH	=	ethanol
H_e	=	expected heterozygosity
EST	=	Expressed Sequence Tag
F	=	Fixation index
ISSR	=	Inter Simple Sequence Repeat
L	=	Liter
T_m	=	melting temperature
µg	=	microgram
µl	=	microliter
mg	=	milligram
ml	=	milliliter
mtDNA	=	mitochondrial DNA
H_o	=	observed heterozygosity
PCR	=	Polymerase Chain Reaction
RAPD	=	Random Amplified Polymorphic DNA
RFLP	=	Restriction Fragment Length Polymorphism
STR	=	Short Tandem Repeat
SSR	=	Simple Sequence Repeat
SNP	=	Single Nucleotide Polymorphism
temp.	=	temperature

GENETIC VARIABILITY AND DIFFERENTIATION OF WHISKER SHEATFISH *Micronema bleekeri* (Günther, 1864) POPULATIONS BASED ON MICROSATELLITE MARKER

INTRODUCTION

Whisker sheatfish (*Micronema bleekeri* Günther, 1864) is an economic freshwater fish species. Its meat is appreciated by local people and the species is thus become economically important (Na-Mahasarakarm, 2007). This species distributes itself in the wide areas of Southeast Asia (Rainboth, 1996; Coates *et al*., 2003). In Thailand, they are found throughout the Chao Phraya, Pasak, Maeklong, Tapi, Southeast river systems, Peninsular Thailand and the Mekong River (Vidthayanon *et al*., 1997). Generally, whisker sheatfish are almost totally harvested from the nature although attempts have been made to artificially culture them for the higher demand of market but the production is not very successful. The decline in whisker sheatfish population is due to habitat degradation, over exploitation and water pollution (Fishery Information Technology Center, 2006). Moreover, the water level of freshwater resources is considerably fluctuated. At the peak of the summer, some marginal areas become shallow resulting in the emerging geographical barriers preventing the fish to freely contact. The isolated fish, then subjected to restricted gene flow leading to high level of inbreeding, loss of genetic diversity, and reduction of its ability to adapt to environmental changes (Wu, 2005). Species which restricted gene flow as a consequence of adaptive divergence often exhibit population structuring (Campos *et al*., 2006). Identification of population structuring is intended to help maintain genetic variability in declining populations. Genetic variation is an important factor in the process of evolution in natural populations. In Thailand, the genetic diversity and the population structure have been investigated in other commercial freshwater fishes such as Nile tilapia (Rutten *et al*., 2004), walking catfish (Na-Nakorn *et al*., 2004), giant catfish (Na-Nakorn *et al*., 2006) but not whisker sheatfish. Therefore, genetic structure of whisker sheatfish in Thailand is not characterized and remains unclear. Knowledge of genetic variation and population

structure of this valuable species is essential for protecting genetic resource by the reduction of genetic erosion and for fishery management.

The Northeastern Thailand provides different areas for collecting whisker sheatfish samples such as the Mekong river- the international river (Coates *et al*., 2003), Songkhram river basin- the Mekong River tributaries and wetland model of the Mekong River Commission (MRC) (Coates *et al*., 2003), the Nong Han Lake- large wetland, and small river channel isolated from the main river by drainage structure (Na-Mahasarakarm, 2007). These sites contain a great diversity of whisker sheatfish for observation and research work. However, natural populations usually have a range of geographic distribution and are exposed to different environments at different locations. Under differential selection, individuals tend to adapt to their local environmental conditions resulting in a pattern of local adaptation (Lenormand, 2002) and adaptive divergence (Hendry, 2001).

There are many types of genetic marker used to detect the genetic variation of the organisms (Frankham *et al*., 2009). Microsatellite is a well-known DNA marker. It consists of iterated simple sequence, usually occurring as tandem repeat of 1 to 6 bases generating repeat unit from a single base pair to thousands of base pairs (Ellegren, 2004; Chistiakhov *et al*., 2006; Quan *et al*., 2006). This marker is a superior tool to use in several fields of genetic linkage mapping (Liao *et al*., 2007), conservation genetics (Su *et al*., 2007), genome analysis (Zheng *et al*., 2007), and aquaculture (van Herwerden *et al*., 2006) as well as population genetics (Cui *et al*., 2005; Quan *et al*., 2006). Microsatellite marker has advantages over other types such as AFLPs or RAPDs. This marker is locus-specific, co-dominant inheritant, evolved neutrally, highly abundant, highly polymorphic (Ellegren, 2004), while only small amounts of DNA are needed for PCR amplification and the results are highly reproducible (Goldstein and Pollock, 1997). It has been successfully used to estimate genetic variation of wild and hatchery stocks in many fishes (Hogan and May, 2002; Na-Nakorn *et al*., 2006; So *et al*., 2006). However, there is no report on the use of this marker on whisker sheatfish. Therefore, the microsatellite marker was chosen to study genetic variation of whisker sheatfish populations for this project.

OBJECTIVES

1. To develop microsatellite markers from *Micronema bleekeri* (Günther, 1864) using the enrichment strategy.

2. To evaluate the genetic diversity of *Micronema bleekeri* (Günther, 1864) populations collected from seven areas in Northeastern, Thailand, and to identify population genetic differences using the developed microsatellite markers.

LITERATURE REVIEW

1. Taxonomy of whisker sheatfish

Whisker sheatfish is classified in (Rainboth, 1996);

Kingdom: Animalia

Phylum: Chordata

Class: Actinopterygii (ray-finned fishes)

Order: Siluriformes (catfish)

Family: Siluridae (Sheatfishes)

Genus: *Micronema*

Species: *Micronema bleekeri*

Phalacronotus bleekeri (synonym)

2. Biology of whisker sheatfish

Genus *Micronema* consists of three species, *i.e.*, *Micronema apogon*, *Micronema micronema* and *Micronema bleekeri*.

Micronema bleekeri, whisker sheatfish, is identified by its having soft rays with the length of 77-85 cm. The dorsal fin is absent, while the pectoral fins is much shorter than head. They show a denticulated spine. The pelvic fins is small but the anal fin is long (Taki, 1974). The mandibular barbels are shorter than eye width. The vomerine teeth are smoothly curved band and the maxillary barbel is not extended beyond jaw (Rainboth, 1996).

Whisker sheatfish inhabit in rivers, streams and lakes as well as in impoundments, throughout the Chao Phraya river, Pasak river, Tapi river and the Mekong river and its tributaries (Rainboth, 1996). It is a migratory species (Hill and Hill, 1994). Long-distance migrations within main river and tributaries are normally referred to as 'longitudinal' whilst those from the main river and tributaries into flood plain areas during the flood season and back again during the dry season are referred

to as 'lateral' (Bao *et al*., 2001). Whisker sheatfish undertakes lateral migration (Sokheng *et al*., 1999). Migrations are triggered by the first rainfall at the end of the dry season, as well as water level changes (Eric, 2006). They return to the river from the floodplain and tributaries on, or immediately before, the full moon (Rainboth, 1996). They feed on small fishes, shrimps and aquatic insect larvae (Ukkatawewat, 1984). However, migratory pattern of whisker sheatfish is cryptic between tributaries and swamp especially in disconnected flow that disrupted by barrier. Habitat factors can strongly influence the extent and direction of movement whereas physical distance may be the most fundamental in its effects on movement.

3. Genetic variation

The importance of genetic variation to population adaptability in changing environments or under stressful conditions has long been recognized (Frankel and Soule, 1981; Allendorf *et al*., 1987). The loss of genetic variation, due to prolonged selection, loss of alleles or heterozygosity due to (random) inbreeding or isolation may result in a decrease of the potential adaptability of a population (Ferguson *et al*., 1995).

The general goals of population genetic studies are to characterize the extent of genetic variation within species and account for this variation (Weir, 1996). The frequency of genes and the forces that affect their frequencies, such as migration, mutation, selection and random genetic drift (Gall, 1987) can determine the amount of genetic variation within and between populations.

To describe the genetic constitution of a group of individuals it is necessary to specify their genotypes and enumerate each genotype. One useful measure of genetic diversity is population heterozygosity (*H*), defined as the mean percentage of loci heterozygous per individual (or equivalently, the mean percentage of individuals heterozygous per locus). Heterozygosities may also be estimated from observed frequencies of alleles (rather than genotypes), assuming the population is in Hardy-Weinberg equilibrium (Frankham *et al*., 2002).

Moreover, allelic richness is to allow the genetic diversity of samples to be compared regardless of their respective sample sizes. Observed allelic richness needs correction for the different sample size of populations using the rarefaction and/or extrapolation techniques (Leberg, 2002).

If analyses are to be made using allele frequencies, rather than genotypic frequencies, it is necessary to ensure that the populations are in Hardy-Weinberg equilibrium. This law states that in a large random mating population with no selection, mutation or migration, the allele frequencies and the genotype frequencies are constant from generation to generation and that, furthermore, there is a simple relationship between the gene frequencies and the genotype frequencies (Wright, 1969). A population with these genotype frequencies is said to be in Hardy-Weinberg equilibrium (HWE) at the locus under investigation (Guo and Thompson, 1992).

The linkage (or gametic) disequilibrium, LD, measures the lack of fit of observed two-locus gametic frequencies to those anticipated based on the product of the single locus allelic frequencies. There are two major phenomena responsible for linkage disequilibrium or non-random association of alleles between two loci on a chromosome. They are epistatic natural selection and random genetic drift. Furthermore, if the population at issue went through a bottleneck which reduced the effective population size (N_e) to a small number of breeding adults, it might be expected to see a significant linkage disequilibrium value for several generations (May and Kruger, 1990).

The detection of HWE and LD is the essential steps in the study of population genetics. Populations, which deviate from Hardy-Wienberg and linkage equilibrium, are due to many reasons. Avise *et al*. (2004) suggested that homozytgote exceed or heterozygote deficit might be due to several factors. First, the locus is under selection. Second, “null alleles” may be present which are leading to a false observation of excess homozygotes. Third, inbreeding may be common in the population. Fourth, the presence of population substructure may lead to Wahlund effect (Wahlund, 1928). From these reasons, Castric *et al*. (2002) concluded that heterozygote deficiency that

resulted from Wahlund effect has been recently documented in lake populations of the brook charr whose movement in riverine habitats was restricted. In addition, So *et al*. (2006) reported that the three samples collected at the spawning sites showed heterozygote deficiency, implying that they could be composed of a mixture of sutchi catfish from differentiated gene pools called Wahlund effect. Furthermore, null alleles affected the population genetic analyses of whisker sheatfish. In addition, Chapuis and Estoup (2007) reported that the presence of null alleles led to overestimation of both F_{ST} and genetic distance. In the migration model, larger bias in F_{ST} was observed for high null allele frequencies and low levels of gene flow. Also null alleles are commonly encountered in species especially in population for which N_e is not necessarily large (Dakin and Avise, 2004) as whisker sheatfish populations.

Wright's (1951) *F*-statistics have proved to be useful for illuminating the pattern and extent of genetic variation residing within and among natural populations of animal and plant species. In addition, genetic differences between subpopulations will evolve in the course of time if there is little or no gene flow between them (Chakraborty and Leimar, 1987). That means restriction on gene flow may lead to genetic subdivision. Moreover, Wade (1988) revealed that the effect of random genetic drift could be an important reason for the significant differentiation. In particular, marine species show lower levels of genetic population differentiation than fresh water or anadromous species, probably because there are potentially fewer barriers to migration and gene flow (Carvalho and Hauser, 1995). For a total population that is subdivided into many subpopulations, Wright (1951) defined three *F*-statistics (correlation between uniting gametes), to relate the deviation from Hardy-Weinberg in the total population (F_{IT}), to the genetic divergence among subdivisions (F_{ST}) and to averaged deviation from Hardy-Weinberg within subdivisions (F_{IS}) (Yang, 1998). F_{IT} values are seldom used since any type of departure from a single panmictic population will lead to a significant F_{IT} value. In addition, F_{IS} values help detect departures from Hardy-Weinberg by measuring the amount of heterozygote deficiency or excess observed in the sample. Also it can be defined as the probability that two genes at a given locus, one taken at random from each of two randomly selected individuals from the population, are identical by descent (Frankham *et al*.,

2009). In addition, F_{ST} values help understand the degree of population differentiation within species.

Gene flow is defined as the movement of genes within and between populations and thus it includes all movement of gametes and individuals that are effective in changing the spatial distributions of genes (Slatkin, 1985). Basically, gene flow is interpreted as a migration rate (m) illustrating the allele frequencies in a population of each generation which are of migrant origin. It is notoriously difficult to monitor gene flow directly so it has to be inferred indirectly from the spatial distributions of genetic markers by statistical approaches (Avise, 1994). In order to measure the rate of migration (m), knowledge of the effective population size (N_e) is required; however, the absolute number of migrants into a population (N_em) is related to the level of genetic differentiation between the source and the native populations.

In terms of effective population size (N_e), Frankham *et al*. (2002) suggested that N_e is the effective number of individuals contributing their genes or alleles to the next generation. The effective population size is one of the most important parameters in evolutionary and conservation biology. Moreover, not only N_e determines the degree of genetic drift and the effectiveness of natural selection, but it also affects population viability in small populations. When the relative importance of genetic drift (chance) is higher, deleterious alleles can become more frequent and 'fixed' in a population due to chance. Any allele, deleterious, beneficial or neutral is more likely to be lost from a small population (gene pool) than a large one. In case of a wild population, population bottleneck may occur due to population isolation or dramatic reduction of effective population size. Even with random mating, when the effective population size is small, a decrease of heterozygosity can occur (Crow and Kimura 1970). Population bottlenecks typically affect most loci across the genome, and therefore, can be verified by investigating different loci (Piry *et al*., 1999).

The genetic study of natural populations is dependent on the availability of polymorphic neutral markers. Although electrophoresis of proteins has been widely used for the direct study of genetic variation in fish populations, DNA markers are

becoming more popular to obtain information on gene flow, allele frequencies and other parameters that are crucial in population biology (Neigel, 1997).

4. Polymerase Chain Reaction

Microsatellite alleles are amplified using the polymerase chain reaction (PCR). The purpose of PCR is to facilitate analysis of a specific DNA sequence by increasing the number of copies of that sequence and also minimizing the influence of non-target DNA (Mullis and Faloona, 1987). Often, only a small amount of DNA is available from a sample and by itself is insufficient for a given analytical procedure. For example, starting with 10^{-6} µg of template DNA, 0.5-1.0 µg of target sequences up to 2 kb in length can be obtained after 30-35 PCR cycles (Sambrook *et al*., 1989).

A single PCR cycle consists of three steps: denaturation, annealing, and extension (Palumbi, 1994). During the first step, the template DNA is heat denatured, at approximately 94°C, to form two single DNA strands. Two oligonucleotide primers that have been designed to be complimentary to known sequences that flank the target sequence, are annealed to the single strands of DNA in the second lower temperature step. During the third step, a heat stable DNA polymerase then binds to the template/oligonucleotide combination and creates a DNA strand complimentary to the template (Saiki *et al*., 1988). When complete, a single PCR cycle effectively doubles the number of copies of the sequence of interest.

5. DNA marker and their application in freshwater fish

Genetic diversity is the variety of alleles and genotypes present in the group under study (population, species or group of species; Frankham *et al*., 2002). It is usually required for populations to evolve and adapt to environmental change. It has been measured for different traits, *e.g*., continuously varying (quantitative) characters, deleterious alleles, proteins, nuclear DNA loci, mitochondrial DNA (mtDNA), and whole chromosomes. Numerous methods are available for measuring genetic diversity

at the genetic level. DNA marker is the great potential to determine genetic diversity for species.

There are several types of DNA marker. Normally, DNA markers are classified into two categories (Liu and Cordes, 2004): type I are markers associated with genes of known function (RFLP, EST, mtDNA), while type II markers are associated with anonymous genomic segments (RAPD, AFLP, microsatellite, SNP) (O'Brien, 1991). Type I markers have been used in comparative genomics, genome evolution, candidate gene identification, and enhanced communication among laboratories (Liu and Cordes, 2004). Type II marker have found widespread use in population genetic studies in terms of characterizations of genetic diversity and divergence within and among populations (Brown and Epifanio, 2003).

RFLP is a co-dominant marker based on digestion of DNA with restriction enzymes resulting in fragments whose number and size can vary among individuals, populations, and species. Changes in the DNA sequence due to indels, base substitutions, or rearrangements involving the restriction sites can result in the gain, loss, or relocation of a restriction site (Liu and Cordes, 2004). The major disadvantage of RFLP is the relatively low level of polymorphism. In addition, sequence information or probes are required, making it difficult and time-consuming to develop markers in species lacking known molecular information (Liu and Cordes, 2004). Quiniou *et al*. (2005) studied the role of MHC molecules in spontaneous allogeneic cytotoxic responses of channel catfish using RFLP. This study indicated that channel catfish MHC class IIA and class IIB genes are linked. Hashimoto *et al*. (2010) identified hybrids between Neotropical fish *Leporinus macrocephalus* and *Leporinus elongatus* by PCR-RFLP. They found that the nuclear α-tropomyosin gene allowed the clear distinction between the parental and hybrid individual fish.

ESTs provide a rapid method of gene discovery that has been widely applied in humans and other species. ESTs are partial sequences of randomly selected complementary DNA (cDNA) clones. EST has applications in discovery of new genes, mapping genomes and identifying coding regions in genomic sequences

(Savan and Sakai, 2002). One advantage of EST analysis is its ability to measure the relative abundance of expressed transcripts (Chu *et al*., 2006), which allows the evaluation of the relative importance of expressed genes. Chu *et al*. (2006) studied profile analysis of expressed sequence tags derived from the ovary of tilapia whereas Wang *et al*. (2007) developed novel EST-SSR markers in common carp by data mining from public EST sequences.

mtDNA can be considered as a haplotype and inherited mainly maternally. It evolves faster than nuclear genes but not the hypervariable sequences (Penzes *et al*., 2002). It has been used extensively for species identification, and recently has been validated as a method of species identification for the major vertebrate phyla. It needs low amount of DNA, but middle-laborious use, very high reproducibility and reliability (Penzes *et al*., 2002). mtDNA markers are also effective in assessing genetic, demographic and phylogeographic patterns in intraspecific populations (Eguia *et al*., 2004). Variation at mtDNA may be analysed mainly with two different approaches: RFLP analysis of whole purified mtDNA and RFLP analysis or DNA sequencing of small segments of the mtDNA molecule obtained by means of PCR amplification (Okumus and Ciftci, 2003). Eguia *et al*. (2004) analyzed genetic diversity in farmed Asian Nile and red hybrid tilapia stocks by mtDNA-RFLP. It produced 14 restriction morphs which corresponded to eight distinct haplotypes. Jondeung *et al*. (2007) developed the complete mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*) and studied the phylogenetic relationships among Siluriformes. They found the complete nucleotide sequence (16,533 bp) of the mitochondrial genome of this species.

RAPD markers are inherited as Mendelian markers in a dominant fashion. This marker uses the single short primers and relatively low annealing temperatures (often 36-40 $^{\circ}$C) for PCR to randomly amplify anonymous segments of nuclear DNA with an identical pair of primers 8-10 bp in length. Genetic variation and divergence within and between the taxa of interest are assessed by the presence or absence of each product. RAPD polymorphisms can occur due to base substitutions at the primer binding sites or to indels in the regions between the sites. The potential power is

relatively high for detection of polymorphism; typically, 5-20 bands can be produced using a given primer pair. RAPD has advantages that primers are commercially available and do not require prior knowledge of the target DNA sequence or gene organization (Dinesh *et al*., 1995). However, RAPD markers are subjected to poor reproducibility due to the low annealing temperature used in the PCR amplification (Liu and Cordes, 2004). RAPD makers have been used for species identification in fishes (Partis and Wells, 1996). Liu *et al*. (1999) studied gene mapping and analyzed genetic variation of catfish by RAPD. It generated 462 polymorphic bands ranging from 200 to 1,500 base pairs. Barman *et al*. (2003) observed genetic variation between four species of Indian major carps using RAPD. It was found that 45% of the scorable RAPD bands were specific to each species.

AFLP is a multi-locus fingerprinting technique. It is inherited as dominant markers. Polymorphism of this marker includes indels between restriction sites and base substitutions at restriction sites. It also includes base substitutions at PCR primer binding sites (Liu and Cordes, 2004). The unique feature of the technique is the addition of adaptors of known sequence to DNA fragments generated by digestion of whole genomic DNA. The power of AFLP analysis is tremendously high in revealing genomic polymorphisms and different enzymes can be used to scan the genomes. The major strengths of the AFLP method include large (over 100) numbers of revealed polymorphisms, high reproducibility due to high PCR annealing temperatures, and relative economy on a per marker basis (Yue *et al*., 2004). AFLP does not require any previously known genetic information (Mickett *et al*., 2003). Its major weakness includes the need for special equipment such as automated gene sequencers for electrophoretic analysis of fluorescent labels. Traditional electrophoretic methods can also be employed but they require the use of radioactive labels or special staining techniques such as silver staining (Liu and Cordes, 2004). Agresti *et al*. (2000) produced linkage maps from several three-way and four-way crosses of Nile tilapia using AFLP map. Mickett *et al*. (2003) studied genetic analysis of diversity among and within domestic populations of channel catfish using AFLP markers. They found 454 polymorphic bands from 16 populations. Poompuang and Na-Nakorn (2004) constructed the first linkage map for walking catfish, using AFLP. They found the

AFLP map consists of 134 loci placed into 31 linkage groups, with 12 loci remaining unlinked. Simmons *et al*. (2006) studied genetic diversity between domestic and wild channel catfish populations. They found 396 polymorphic bands across the 269 fish.

SNPs describe polymorphisms caused by point mutations that give rise to different alleles containing alternative bases at a given nucleotide position within a locus (Liu and Cordes, 2004). SNPs occur more frequently in the genome than microsatellites, but mutation rates in SNPs are significantly lower than in microsatellites (Nielsen, 2000). This stability of SNPs is an advantage in evolutionary studies, population biology and pedigree studies. However, SNPs can also be detected in the coding sequences, thus, SNPs may affect protein function and expression levels directly (Rengmark *et al*., 2006). SNPs study of freshwater fish still was limited informations, whereas, provides more information for marine fish. Rengmark *et al*. (2006) studied genetic variability in wild and farmed Atlantic salmon strains estimated by SNP. They found the average observed heterozygosity value of 0.41 and observed heterozygosity across population ranging from 0.05 to 0.77. Hubert *et al*. (2009) developed single nucleotide polymorphism markers for Atlantic cod using expressed sequences. They used thirty-three SNPs to test for variability in fish from two Atlantic Canadian populations. They found that significantly 30 SNPs tested were polymorphic in wild Atlantic cod.

6. Microsatellite marker and their application in freshwater fish

Microsatellite or simple sequence repeats (SSRs) represent a unique type of tandemly repeated genomic sequences, which are abundantly distributed across genomes in many living organism (Okumus and Ciftci, 2003). Microsatellites are stretches of DNA consisting of tandemly repeated short units of 1-6 base pairs (bp) in length (Liu and Cordes, 2004). It is inherited in a Mendelian fashion as co-dominant markers (Penzes *et al*., 2002) in diploid sexual organisms. One allele is inherited from the mother and another allele is inherited from the father that can cause information loss in the case of dominant markers, because the genotypization is not so exact (Penzes *et al*., 2002). Di-nucleotide repeats dominate, followed by mono- and tetra-

nucleotide repeats, and tri-nucleotide repeats are least dominant (Ellegren, 2004). Rat and human revealed $(CA)_n$, $(AAAT)_n$, and $(AG)_n$ repeat motifs. Plants exhibited common motif, $(AT)_n$ (Ellegren, 2004), whereas, mono- and tetra-nucleotide repeats are common in bird (Yang *et al*., 2008). In addition, microsatellite DNA has been widely distributed as often as every 10 kb in fish genome (Wright, 1993). Generally, di-nucleotide repeats, especially $(CA)_n$ and $(TG)_n$ are widely dispersed in fish genome (Cui *et al*., 2005). Microsatellite repeat was divided into three types. Perfect repeats consist of a specific unit (*e.g*. CA) repeated in sequence, compound repeats are composed of a series of one unit (*e.g*. CA) followed by a series of another unit (*e.g*. GA) as CACAGAGAGA, and interrupted repeats are sequences of a specific unit that are separated by other nucleotides as CACACAGTAATCACACACA. The key feature of SSRs as molecular markers is their hypermutability and, hence, their hypervariability in species and populations. The microsatellite mutation rate is estimated at 10^{-2}-10^{-6} per locus per generation (Ellegren, 2000). Two models have been suggested to explain microsatellite generation and evolution (Ellegren *et al*. 1995). The first model derives from the stepwise mutation model (SMM), which involves with polymerase slippage during DNA replication, resulting in insertions or deletions of repeat units relative to the template strand (O' Connell and Wright, 1997; Altukhov and Salmenkova, 2002; Ellegren, 2004; Lucentini *et al*., 2006) and generate the differences in the number of repeat units (Liu and Cordes, 2004). The second model is the infinite-allele model (IAM) which derives from replication slippage and can also occur during PCR amplification of microsatellite sequences *in vitro*. Therefore, microsatellite polymorphism is based on size differences due to varying numbers of repeat units contained by alleles at a given locus (Penzes *et al*., 2002).

Microsatellites can be simply and rapidly detected by the polymerase chain reaction (PCR) using two unique oligonucleotide primers that flank the microsatellite and hence define the microsatellite locus (Chistiakov *et al*., 2006). Because of their multiallelic nature, co-dominant inheritance, small length, extensive genome coverage and relative abundance, microsatellites have been successfully applied in a wide variety of research fields and practical disciplines. Microsatellite loci are often highly polymorphic due to variation in the number of repeat units which can easily be

determined by estimating the size of the entire microsatellite and can result in a large number of alleles at each microsatellite locus in a population. The high polymorphism of microsatellite markers enables higher statistical power and increased discrimination among genotypes (Halkett *et al*., 2005) making them very useful for applications such as parent-offspring identification in mixed populations, while others have lower numbers of alleles and may be more suited for population genetics and phylogeny (O'Connell and Wright, 1997). Moreover, primers developed for one species may be used to analyze the genetic diversity in closely related species (Okumus and Ciftci, 2003; Gu *et al*., 2006). However, although the number of completed genomes is increasing with the availability of new sequencing technologies, only 114 eukaryote genomes have been sequenced until today (Techen *et al*., 2010). As for the whisker sheatfish there are none and only rare sequence information of its related species in public databases to design as primer for this species.

Although strength of microsatellite is their abundance, genomic distribution, small locus size and high polymorphism, the major drawback of microsatellites is that they need to be isolated for the first time from most species examined (Zane *et al*., 2002). In addition, this marker shows the species-specific character (Okumus and Ciftci, 2003). The microsatellite marker for closely related species may be unapplicable because the conservation of flanking region might be insufficient across the taxa (Gu *et al*., 2006). This is due to the fact that microsatellites are usually found in non-coding regions where the nucleotide substitution rate is higher than in coding regions. So, the strategy of designing the universal primers matching conserved sequences is more problematic for microsatellites (Zane *et al*., 2002). Therefore, high polymorphism observed in a species does not guarantee that similar polymorphism will be found in related species especially when increasing the evolutionary distance (Morin *et al*., 1998). In addition, microsatellite is probably rarely useful for higher-level systematics. Across highly divergent taxa two problems arise such as the microsatellite primer sites may not be conserved, homoplasy becomes much more likely, so, can no longer safely assume that two alleles identical in state are identical by descent (Vali *et al*., 2008). Moreover, null alleles have long been known in DNA electrophoresis with reduced or absent expression of a DNA product (Utter *et al*.,

1987; Murphy *et al*., 1996) and more recently have been observed for microsatellite loci (Callen *et al*., 1993). A null allele can be defined as any allele at a microsatellite locus that is only weakly amplified or not visible after amplification and separation (O'Connell and Wright, 1997) and is recognized, together with population subdivision, as a major factor in depression of observed heterozygosity, compared with that expected on the basis of Hardy-Weinberg equilibrium. Therefore, when some microsatellite loci tend to produce null alleles, statistical and data interpretation should be undertaken with caution (Gu *et al*., 2006). Each microsatellite locus has to be identified and its flanking regions need to be sequenced for the design of PCR primers (Liu and Cordes, 2004). Zane *et al*. (2002) discussed the other problem of microsatellite that the isolation of microsatellite can be quite involving in terms of effort and time because it traditionally consists of screening genomic libraries with appropriate microsatellite probes. The number of positive clones containing microsatellites that can be obtained by traditional method usually ranges from 12% to less than 0.04%. However, the statistical power depends not only on the number of scored loci but also on other factors such as the degree of polymorphism of each locus and the sample size, and so the use of a limited number of loci might fail to provide sufficient information (Gu *et al*., 2006). Traditional strategies are less useful when dealing with taxa with a very low frequency of microsatellites, when a large number of microsatellites are required as in case studies on the genetic distances between populations. However, microsatellite marker can provide strong foundations for the identification of species or population and the assessment of fish genetic diversity, and also are currently the tool of choice for looking at population genetic structure, phylogeny and parentage analysis of freshwater fish (Gu *et al*., 2006).

Since the enrichment strategy has been developed, it raises the efficiency of microsatellite and reduces the cost and time-consuming for microsatellite development (Penzes *et al*., 2002; Zane *et al*., 2002; Gu *et al*., 2006). There are many reports on enrichment strategy used for microsatellite development. Hogan and May (2002) discovered novel microsatellite primer for the migratory Asian catfish family Pangasiidae. They found 27 microsatellite markers from 130 positive clones which showed polymorphic markers at least one species of five catfish. Sukmanomon and

Poompuang (2003) developed microsatellite primers for walking catfish (*Clarias macrocephalus*), they found 41 sequences contained microsatellite from 2,841 recombinant clones. Sukumasavin *et al*. (2004) identified and characterized microsatellite markers from endangered cyprinid, the seven-line barb. They discovered 83 sequences contained microsatellites from more than 5,000 screened colonies whereas only 20 sets of PCR primers were designed and tested. Moeser and Bermingham (2005) isolated microsatellite for Neotropical freshwater catfish. They characterized CA repeat unit and found that 23 clones of 150 recombinant clones contained arrays of at least 10 uninterrupted repeats. Na-Nakorn *et al*. (2006) developed microsatellite loci for the endangered Mekong giant catfish. They received 10 primer pairs from 82 sequenced inserts which was able to amplify DNA of this species. In addition, Ngamsiri *et al*. (2006) characterized microsatellite markers from the critically endangered species, *Pangasianodon gigas*, using $(GT)_{15}$ probe. They found that 39 clones containing microsatellite could be obtained from more than 3,000 of screened colonies and also successfully amplified four related species. Moreover, Yang *et al*. (2008) developed microsatellite markers from mud carp using $(CA)_{15}$ probe. They received 56 sequences containing microsatellites from 252 positive colonies whereas only 12 primer sets showed polymorphic loci.

Microsatellite is a famous DNA marker which is a useful tool at the level of populations and individuals to evaluate the genetic diversity in freshwater fish (Gu *et al*., 2006). There are many reports using microsatellite markers to evaluate genetic variation in freshwater fish. Hogan and May (2002) observed genetic variation using cross-species amplification of 27 microsatellite loci for 5 species of Pangasidiiae and found 15 polymorphic loci in at least three species. Bhassu *et al*. (2004) observed genetic structure of *Oreochromis* spp. populations in Malaysia. They found mean heterozygosity ranging from 0.628-0.704 and genetic distance value showing a clear separation between two populations. Rutten *et al*. (2004) evaluated 4 strains of Nile tilapia, and found mean number alleles per locus ranging from 5.0-7.5 with expected heterozygosity ranging from 0.624-0.711. Sukumasavin *et al*. (2004) evaluated two closely related species of endangered cyprinid, and found that 6 out of 20 primer sets provided number of alleles per locus ranging from 7-16 and expected heterozygosity

ranging from 0.47-0.91. Moeser and Bermingham (2005) characterized 23 individuals of Neotropical freshwater catfish and found number of alleles per locus varied from 7-23 whereas the observed heterozygosity ranging from 0.522 to 0.909. Since 2006, microsatellite markers have been developed from the Mekong giant catfish and succeeded in amplifying closely related species of *Pangasius*. Na-Nakorn *et al*. (2006) and Ngamsiri *et al*. (2006) found 2-6 and 2-4 alleles per locus, with observed heterozygosity ranging from 0.05-0.95 and 0.13-0.68, respectively. In addition, Ohashi *et al*. (2006) evaluated genetic variability in founder stock and F_1, and found an average 2.8 and 1.6 alleles per locus with average observed heterozygosity of 0.41 and 0.32, respectively. So *et al*. (2006) studied genetic diversity in 3 populations of the migratory sutchi catfish *Pangasianodon hypophthalmus* in the Mekong river. They found that the mean allelic richness, mean observed heterozygosity, mean expected heterozygosity and mean inbreeding coefficient to be 8.5, 0.773, 0.771 and 0.028, respectively. Ngamsiri *et al*. (2007) inspected genetic diversity of wild Mekong giant catfish *Pangasianodon gigas* collected from Thailand and Cambodia. They found an average 2.8 and 3 alleles per locus with average observed heterozygosity of 0.418 and 0.433, respectively. Sriphairoj *et al*. (2007) observed genetic aspect in broodstock management of the critically endangered Mekong giant catfish in Thailand using 7 microsatellite loci. They found that the genetic variation within stocks was relatively low as number of alleles per locus, effective number of alleles per locus, observed heterozygosity and expected heterozygosity of 2.29±0.76-4.00±1.83, 1.84±0.51-3.04±1.04, 0.58±0.34-0.80±0.12, and 0.43±0.21-0.66±0.11, respectively. In addition, Thai *et al*. (2007) inspected genetic diversity and population structure of common carp (*Cyprinus carpio* L.) in Vietnam using 4 microsatellite loci. The mean number of alleles per locus and mean observed heterozygosity were 4.25-11 and 0.40-0.83, respectively. Yang *et al*. (2008) evaluated genetic diversity of cultivated and wild populations of the mud carp using 12 microsatellite loci. They found the mean number of alleles per locus at 9.08, whereas the mean observed heterozygosity of wild and cultivated populations was at 0.6361 and 0.6417, respectively. In addition, the genetic distance between two populations was 0.1546 with G_{ST} at 0.0473. Ha *et al*. (2009) revealed the genetic differentiation between hatchery and contemporary wild populations of striped catfish, *Pangasianodon hypopthalmus* in Veitnam using five

microsatellite loci. The mean number of alleles per locus, allelic richness, mean effective number of alleles per locus, observed heterozygosity, expected heterozygosity of wild population were 4.80-6.20, 4.54-5.06, 2.86-3.20, 0.62-0.65, 0.62-0.64, respectively, and of hatchery populations were 4.60-5.20, 4.10-4.83, 2.80-3.11, 0.61-0.66, 0.61-0.64, respectively. Na-Nakorn and Moeikum (2009) reported the genetic diversity of striped catfish in Thailand for broodstock management using 5 microsatellite loci. They found allelic richness of the hatchery populations with the non-original introduction and without introduction of 6.53-8.06 and 3.18-6.06, respectively. Heterozygosity was high in the majority of populations (observed heterozygosity=0.633-0.763; expected heterozygosity=0.593-0.834). Also the wild population was not genetically differentiated from the hatchery populations.

7. Description of sample sites

Freshwater fish species show a greater average degree of genetic structuring among locations than those inhabiting estuarine or marine environments. This is thought to be at least partly because it is difficult for obligated freshwater species to move between river systems, often resulting from the isolation of fish populations among drainages. Therefore, freshwater ecosystems will be the most endangered ecosystems in the world and the decline in freshwater biodiversity is far greater than in most affected terrestrial ecosystems (Sala *et al*., 2006).

Whisker sheatfish were collected from seven areas of northeastern Thailand. The Nong-Han Lake is the largest inland water resource of northeastern, Thailand. Limnologically, the significant outflow of the lake is to the Nam Kam basin, which further drains southeast to the world's great Mekong river. The water level of the Nong-Han Lake is controlled by the Nam Kam Gate comprising a spillway and two sluice gates that have been operated for over 25 years (Nielsen and Strom, 1984). The Nam Kam basin is located in northeastern Thailand, discharging into the Mekong River. At present, a hydropower project has been constructed in the basin for irrigation (Nielsen and Strom, 1984). The Mekong river is the largest river in South-East Asia and the twelfth largest in the world. From its source in the Tibetan

Himalayas it flows through six countries China, Myanmar, Thailand, the Lao PDR, Cambodia and Vietnam. There are many tributaries which connect to the Mekong river as well as Songkhram river basin. An estimated 1,700 species of fish are believed to inhabit the waters of the Mekong river. The Mekong river supports one of the largest inland fisheries in the world (Bao *et al*., 2001). Songkhram river basin covers parts of Nakhon Phanom, Sakon Nakhon, Udon Thani and Nong Khai provinces. The Songkhram river basin is an open riverine system intimately linked to the Mekong river mainstream in terms of ecology and hydrology (Poulsen *et al*., 2004). This basin shows a clear geographic barrier. When the Mekong river level is high in some years, there will be a reverse flow of the Mekong river into the Songkhram river basin for many kilometers. The Songkhram river basin is noted for its abundant and biodiverse capture fishery and associated living aquatic resources (Poulsen *et al*., 2004). The Lam Nam Oon is the second largest sub-basin of the Songkhram river basin, covering an area of 3,566 km^2. It features a broad floodplain with tracts of remnant seasonally flooded forest, but much forest has been cleared for agriculture in the last two decades. The river hydrology has been radically altered by an upstream large storage reservoir and irrigation scheme at Lam Nam Oon, built in the late 1960s. The village is situated on elevated land about 1 km south of the river and land use is a mixture of paddy fields, upland cash crops and new conversion to Sakon Nakhon and Nakhon Phanom provinces.

Whisker sheatfish populations in the Mekong river and Songkhram river basin are not obstructed by drainage structure, whereas populations in the Nong-Han lake, Nam Kum basin and Lam Nam Oon are separated by drainage construction leading to small and isolated populations of whisker sheatfish. Each location represents different environments for adaptivity of whisker sheatfish. Therefore, study on genetic variation of whisker sheatfish collected from these different locations will be useful to understand the effect of environmental changes from both natural and man-made on genetic structure of fish in order to prevent genetic erosion and lead to a suitable management in the future.

MATERIALS AND METHODS

1. Sample collection and DNA extraction

Whisker sheatfish samples were collected from seven different geographic regions, namely the Nong-Han lake (NH), Nam Kum basin (KB), the Mekong river at lower Nakhon Phanom province (MK), the Mekong river at upper Nakhon Phanom province (MG), Songkhram river basin at Srisongkharm district (SK), Songkhram river basin at Seka district (SE) and Lam Nam Oon (NA). The locations and the abbreviations for the populations are listed in Appendix Table A1. Sample identification was based on character in the fishery species identification of Rainboth (1996) such as maxillary barbels, anal-fin rays, vomerine teeth. Thirty adult whisker sheatfish per population were collected. The samples were collected during flood until dry season. Tissue samples (about 10 mg) of each individual were obtained from the distal portion of the caudal fin and were stored in absolute ethanol during transportation. DNA was extracted from the caudal fin using Aqua Pure Genomic DNA Isolation kit (Bio-Rad, USA). The detail of the method is shown in Appendix B. DNA concentration was estimated by 1.0% agarose gel electrophoresis, compared with Gene Ruler^{TM} DNA ladder mix (Fermentas, USA) and spectrophotometer. The concentration of genomic DNA was adjusted to 100 ng/µl and 50 ng/µl prior use on the microsatellite enrichment strategy and genetic differentiation analysis, respectively.

2. Microsatellite development and genotyping

2.1 Microsatellite development

A single whisker sheatfish collected from the Mekong river was genomic DNA extracted to develop an enrichment of a microsatellite library. High quality genomic DNA of 100 ng/µl was fragmented using restriction enzyme, *Mse*I (Fermentas, USA). The fragmented DNAs were ligated to specific adapters (5′-GATGAGTCCTGAGT AAC-3′/ 5′-TACTCAGGACTCATC A-3′), then subjected to

PCR amplication using primers complementary to the adapter sequences. The PCR reaction was performed on a XP cycler (BIOER, China) in 30 µl of a mixture containing 50 ng DNA, 20 mM Tris-HCl pH 8.4, 50 mM KCl, 20 pmol of each primer, 1.5 mM $MgCl_2$, 200 µM of each dNTP and 1U of *Taq* DNA polymerase (Invitrogen, Norway). The amplification cycle consisted of 3 min denaturation at 94°C, followed by 30 cycles of 94°C denaturation for 30 s, annealing at 45°C for 1 min and 72°C extension for 1 min. Cycling concluded with a 5 min extension at 72°C in a thermal cycler. The amplified products were electrophoresed in 2% agarose gel, followed by ethidium bromide staining and visualized under UV light. The products were then size-selected to preferentially obtain small fragments of about 300-1,000 bp and purified using Extract II kit (Nucleospin, Germany).

To identify and enrich PCR fragments containing microsatellite DNA, the PCR products were denatured by heating at 95°C for 10 min and hybridized with four biotinylated microsatellite probes: $(CA)_{15}$, $(GA)_{15}$, $(ACC)_{10}$, $(CCT)_{10}$ at 65°C overnight. After hybridization, genomic DNA fragments complementary to the microsatellite motifs were recovered by streptavidin-magnetic bead capture technique (Dynabead® Myone™, Invitrogen, Norway) and several washes to remove non-specific binding. The DNAs were eluted and recovered by PCR amplification using adapter specific primers. The PCR products were purified using Extract II kit (Nucleospin, Germany). The enriched DNAs were ligated into pGEM®-T easy Vector (Promega, USA) and transformed into competent *E. coli* JM109 cells (Promega, USA). Recombinant transformants were selected by growing them on Luria Bertani (LB) agar plates containing 100 mg/ml ampicillin and 50 mg/ml X-gal and left at 37°C overnight. The white clones were selected and tested for DNA insertion by PCR using M13 universal primers. The reaction was carried out by 3 min denaturation at 94°C, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 50°C for 30 s, 72°C extension for 1 min, and final extension at 72°C for 5 min. Analysis of the amplified products was achieved by electrophoresis on a 2% agarose gel and only positive clones were taken for further analysis. To confirm the presence of microsatellite sequences, the positive clones were subjected to Southern hybridization using repeat-containing probes and North2South® Chemiluminescent Hybridization

and Detection kit (PIERCE, USA). The detail of the method is shown in Appendix B. The recombinant clones were directly sequenced. CAP3 program (Huang and Madan, 1999) was used to assemble sequences into contigs to facilitate the identification of duplicate sequences and to remove vector sequence. Flanking regions were recovered to design primers for the amplification of each microsatellite sequence using PRIMER 3 (Rozen and Skaletsky, 2000).

2.2 Genotyping using microsatellite markers

Genomic DNAs of all whisker sheatfish samples were used for PCR amplification of 13 microsatellite loci: MB71, MB79, MB81, MB320, MB354, MB381, MB401, MB456, MB515, MB613, MB614, MB618, MB645. PCR reactions were carried out on a XP cycler (BIOER, China) in 12.5 µl volumes containing 50 ng of genomic DNA, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 2.0 mM dNTP, 2 pmol of each primer and 1U *Taq* DNA polymerase (Invitrogen, Norway). Details for thermal profiles for all loci consisted of an initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C denaturation for 45 s, annealing at 53°C-61°C depending on the locus for 30-45 s and extension at 72°C for 30-45 s. Cycling concluded with a 3 min extension at 72°C to complete the extension. Loci were examined for successful amplification by running the PCR products (12.5 µl) on 6% (w/v) denaturing polyacrylamide gel (Bio-Rad, USA) and visualized using silver-staining (0.2% silver nitrate) procedure described by Karnsomdee and Meckvichai (2002). Sizes of alleles were determined based on PhiX 174 *Hin*fI marker (Promega, USA) and 100 bp ladder marker (Fermentas, USA).

3. Statistical analysis of microsatellite data

Prior to the analysis of data for population substructure, it is essential to test the genetic variation found at microsatellites to ensure that the basic assumptions upon which the subsequent theory based are not violated. Three main assumptions need to be tested. First, the selective neutrality of each locus should be analyzed. Second, the presence of "null alleles" (alleles which are not detected via PCR

analysis) should be identified. Finally, before the data from various loci are combined, the independent assortment of the loci must be tested (Murray, 1996).

3.1 Microsatellite diversity, Hardy-Wienberg and linkage disequilibrium

The DNA bands were scored by size estimation comparing with 50 bp standard markers (Fermentas, USA). Evidence for genotyping error in the scoring of the microsatellite alleles for all loci and populations was assessed using MICRO-CHECKER 2.2.3 (van Oosterhout *et al*., 2004). The program helps detect large allele drop out, the presence of null alleles and errors in allele scoring due to stutter. Deviations from Hardy-Weinberg and linkage equilibrium were tested using Genepop 3.4 (Raymond and Rousset, 2003). Sequential Bonferroni correction (Holm, 1979) was employed to account for multiple testing. When the null hypothesis was rejected, the F_{IS} statistic of Wright (1951) estimated following Weir and Cockerham (1984), was used as an indicator for heterozygote excess or deficit. Linkage disequilibrium (LD) between pairs of loci within locality was tested using GENEPOP 3.4 (Raymond and Rousset, 2003). Genotype data from 11 microsatellite loci of whisker sheatfish were used to calculate the number of alleles per locus (A), allelic frequencies, effective number of alleles (A_e), observed heterozygosity (H_o) and expected heterozygosity (H_e) using POPGEN 32 software package (Yeh and Yang, 1999). Allelic richness as a standardized measure of the number of alleles per locus independent of the sample size was calculated using FSTAT 2.9.3.2 (Goudet, 2001).

3.2 Genetic differentiation between populations and population structure

To evaluate the extent of differences within and among populations, the fixation index, F, was calculated as an estimator of inbreeding (F_{IS}). The significance of the inbreeding value was determined within populations and over loci using FSTAT version 2.9.3.2 (Goudet, 2001). Inter-population genetic differentiation was calculated using the multilocus F_{ST} estimator of Weir and Cockerham (1984) with 1,000 permutation using FSTAT 2.9.3.2 (Goudet, 2001). Balloux and Moulin (2001) interprets the F_{ST} values. Low level of genetic differentiation is 0-0.05. Moderate

level of genetic differentiation is 0.05-0.15. Great level of genetic differentiation is 0.15-0.25. Very great level of genetic differentiation is above 0.25.

COLONY version 2.0 can be used among others in estimating full- and half-sib relationships, assigning parentage, inferring mating systems (polygamous / monogamous) and reproductive skew in both diploid and haplo-diploid species (Wang, 2004). To define individual for population, BAPS version 3.1 (Corander and Marttinen, 2006) was performed.

We also inferred the possibility of hidden population structure using a Bayesian model based clustering approach. The genetic stratification could be done with STRUCTURE 2.2 (Pritchard *et al*., 2000) (50,000 burn in period, 10,000 Markov chains Monte Carlo replications for each of K = 1 to 8, where K is the number of genetically distinct clusters and 3 replicates run per K in order to examine the consistency of result) using the admixture model. STRUCTURE employs a Bayesian model-based clustering method using multilocus genotype data to infer population structure and assign individuals to a genetic cluster (Pritchard *et al*., 2000). For the number of clusters best represented by the data, only individuals with probabilities greater than 0.9 for a specific cluster were retained in that population for subsequent analyses. The criterion implemented in STRUCTURE to determine K is the likelihood of the data for a given K, $L(K)$. This clustering method estimates the most probable number of discrete sub-populations with no *a priori* assumptions of population structure.

3.3 Estimation of effective population size and migration rate

The experiment estimated both the long term and contemporary effective population size (N_e) for the same seven groupings of the data as for the bottleneck tests. The 'recent' effective population size, N_e, was estimated from a single temporal sample using an approximate Bayesian computation framework (Tallmon *et al*., 2004) implemented in the program ONeSAMP 1.1 (Koyuk *et al*., 2008; Tallmon *et al*., 2008). This program is expected to provide more accurate and precise N_e estimates

compared to alternative single-sample methods relying on one genetic parameter (Tallmon *et al*., 2008). The lower and upper bounds of N_e were set for 2 and 1,000, respectively, according to a preliminary survey using the linkage disequilibrium method (Hill, 1981) implemented in NeEstimator 1.3 (Peel *et al*., 2004).

The MIGRATE-n 3.2.6 (Beerli, 2008) was used based on the maximum likelihood (ML)-based coalescent Markov Chain Monte Carlo (MCMC) approach. This program was used to infer the 'historical' effective population size, θ ($\theta = 4N_e\mu$; N_e is the effective population size; μ is the mutation rate per site) and the historical migration rate, M ($M = m/\mu$; m is the immigration rate per generation) among whisker sheatfish populations. For each locus in the data set, the ML was run for ten short and one long chains with 50,000 and 500,000 recorded genealogies, respectively, after discarding the first 10,000 genealogies (burn-in) for each chain. To predict the mutation rate of the whisker sheatfish that is unknown, but required for the calculate m (from $M\mu$ in MIGRATE-n 3.2.6), MSVAR 1.3 (Beaumont, 2004) was used. The MSVAR simulations were run for 2×10^8 iterations and the first 10% of the output was discarded as the burn-in period. Hence, the remained output was information used to estimate the mutation rate of the whisker sheatfish.

To estimate the recent migration and its direction, a Bayesian method based on multilocus genotypes implemented in BAYESASS was used (Wilson and Rannala, 2003). It is assumed that the loci in the source population were in linkage equilibrium. The method is based on MCMC to estimate the posterior probabilities of the migration matrix among neighboring populations.

3.4 Population bottleneck

The reflecting demographic changes or substantial reduction in population size caused bottleneck effects on the populations of the species. To test this hypothesis, we used the two programs for detecting a genetic signature of a recent reduction in the effective population size (N_e) in the populations. Population bottlenecks may also initiate gaps in the size distribution of microsatellite alleles. The

gaps in distributions can be quantified as the *M* ratio, the mean ratio of the number of alleles to the allele size range across all loci. The *M* statistic values were calculated according to Garza and Williamson (2001) to detect recent reductions in effective population size. $M = k/r$ simply measures the ratio of the number of alleles k to the range in allele size r across microsatellite loci. Specifically, all monomorphic loci within a particular population were excluded for over-estimation. In addition, BOTTLENECK 1.2.02 (Piry *et al*., 1999) using the Wilcoxon sign-rank test (Luikart *et al*., 1998a) was used to assess significant deviations from the null hypothesis of a stationary population, and a mode-shift indicator test (Luikart *et al*., 1998b) was performed. This program is very sensitive as it exploits the tendency of bottlenecked populations to loose allelic diversity before heterozygosity. Thus, recently bottlenecked populations will temporarily display a heterozygosity excess compared to that of a population that is presumed to be at mutation-drift equilibrium. Following the recommendation of Piry *et al*. (1999), we calculated H_{eq} under the two-phase mutation model with 95% one-step mutations and 5% multistep mutations, which is believed to be a better fit for microsatellite loci than a pure stepwise mutation model. The program performs optimally with a minimum of 12 loci and 30 individuals (Piry *et al*., 1999). We acknowledge that there may be a reduction in statistical power to detect recent bottlenecks given that we have used only eleven loci.

RESULTS AND DISCUSSIONS

This experiment consisted of 3 parts which are genomic DNA extraction, development of microsatellite markers for whisker sheatfish, and genetic diversity analyses of whisker sheatfish populations using the developed microsatellite markers.

1. Genomic DNA extraction

Genomic DNAs extracted from whisker sheatfish samples displayed a single band of high molecular weight. Smear bands of degraded DNA were not observed. Thus, genomic DNAs extracted from the whisker sheatfish using Aqua Pure Genomic DNA isolation kit (Bio-Rad, USA) were in high quality, providing intense bands without the contamination of protein and RNA (Figure 1).

Figure 1 Genomic DNAs of whisker sheatfish (lane M is100 bp DNA ladder marker; lanes 1- 21 are genomic DNA of whisker sheatfish).

A_{260}/A_{280} ratio of genomic DNA extracted for microsatellite development was 1.8 whereas those of genomic DNAs for population genetic analysis were in the range of 1.7-1.9. Generally, a pure sample of genomic DNA has the A_{260}/A_{280} ratio at 1.8, the ratio that is relatively free from protein contamination (Sambrook *et al*., 1989). Our results of genomic DNA extracted from whisker sheatfish populations maintained good quality and had the amount enough to be used for genetic study based on the microsatellite markers.

2. Development of microsatellite markers for whisker sheatfish

Genomic DNA of a whisker sheatfish was digested with *Mse*I. The smear DNA bands were detected in agarose gel electrophoresis (Figure 2). After ligation the DNA fragments to adapters and PCR amplification, the products were size-selected to obtain a small fragments of 300 to 1,000 bp. Most of the DNA fragments were approximately 500 bp in size (Figure 3). The DNA fragments were then purified, denatured and hybridized with microsatellite probes, $(CA)_{15}$, $(GA)_{15}$, $(ACC)_{10}$, $(CCT)_{10}$. The DNA fragments carrying microsatellite sequences were recovered, re-amplified and purified before ligation to vector according to the enrichment strategy. A total of 610 positive clones (white colonies) were obtained and were tested for inserted DNA fragments by colony PCR (Figure 4).

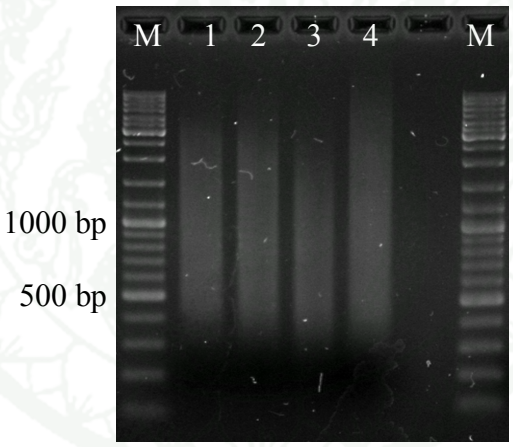


Figure 2 Genomic DNAs of whisker sheatfish after digestion by *Mse*I restriction endonuclease (lane M is 100 bp DNA ladder marker; lanes 1-4 are *Mse* I digested DNAs).

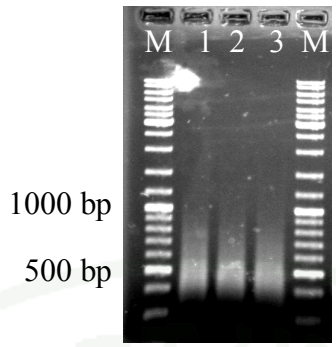


Figure 3 The size-selected DNA fragments for microsatellite development (lane M is 100 bp DNA ladder marker; lanes 1-3 are *Mse* I digested DNA fragments).

In total, five-hundred and twenty-five positive clones were selected. From those selected, eighty-five positive clones giving ambiguous bands or 300 bp PCR product in size were discarded. Normally, negative clones (blue colonies) which did not carry the inserted DNA produced 300 bp PCR products in size. Some clones did carry the linearized vector as an inserted DNA fragment causing false positive with no LacZα produced and no blue colonies formed. This could be proved by *Eco*RI cutting. In our study, the digested fragments ranging from 300-700 bp were observed together with plasmid fragment (Figure 5). It implied that the white colonies carried target inserted DNA fragments, not a linearized vector.

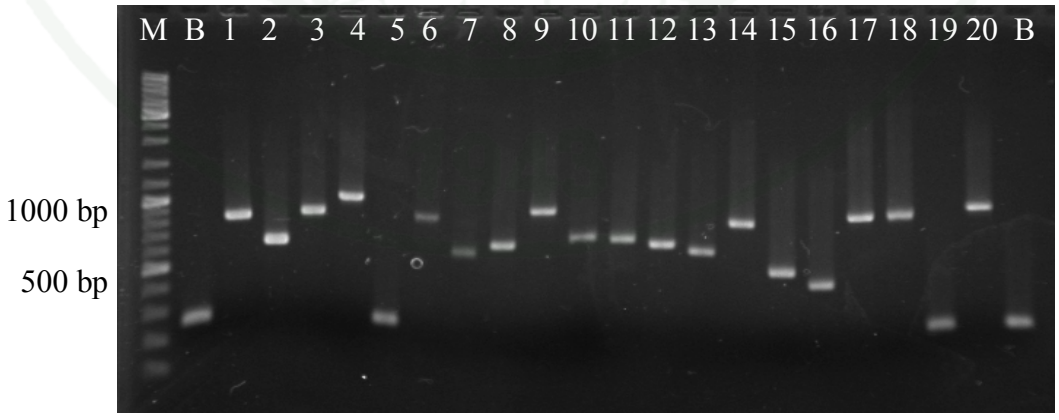


Figure 4 PCR patterns of colonies containing putative microsatellite motifs (lane M is 100 bp DNA ladder marker; lane B is 300 bp in size of blue colony; lanes 1-20 are white colonies carrying various sizes of inserted PCR fragments).

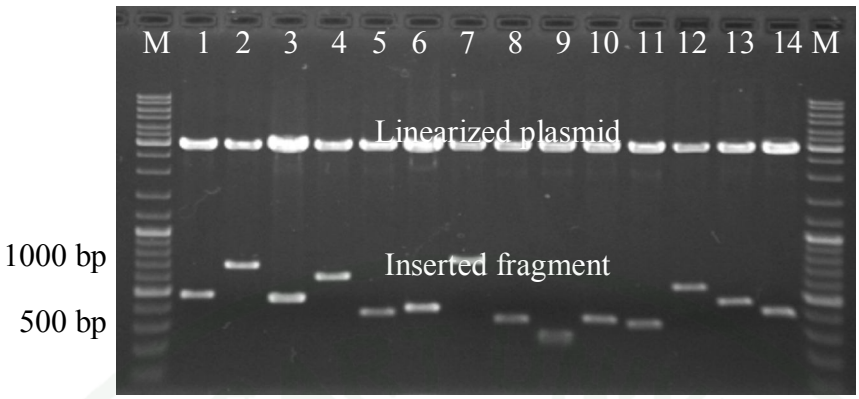


Figure 5 DNA fragments of recombinant plasmids after cutting with *Eco*RI. [lane M is 100 bp DNA ladder marker; lanes 1-14 are recombinant plasmids containing linearized plasmid fragments (upper lines) and *Eco*RI cut fragments (lower lines)].

Two-hundred and fifty clones out of 525 clones (47.61%) were confirmed to contain microsatellite sequences using dot blot hybridization (Table 1, Figure 6). Only 103 clones (19.61%) exhibited strong signal and were sent for sequencing. In total, 99 sequences (18.85%) carried microsatellite DNAs containing 36 di-nucleotide repeats, 30 tri-nucleotide repeats and 18 multiple-motif repeats (Table 2). Fifteen sequences (2.85%) were found more than one time. Identical microsatellite sequences in high copy number were produced in our manipulation of microsatellite enrichment. Zane *et al*. (2002) suggested that sequences from positive clones were redundant and were presented 5-20 times in the microsatellite libraries. The existence of identical microsatellite sequences increased the cost of microsatellite development. Grouping microsatellite sequences by cutting with a set of restriction enzymes should be performed before sequencing.

Two hundred and seventy-five clones (52.38%) were discarded during microsatellite development for whisker sheatfish. Similar results were reported from other experiments. Moeser and Bermingham (2005) concluded that more than 50% of recombinant clones were eliminated during development in Neotropical freahwater catfish, whereas Ngamsiri *et al*. (2006) reported that 57 of 96 positive clones

(59.37%) were discarded in the critically endangered giant catfish microsatellite development.

Table 1 Number of clones, sequences and primers during microsatellite development process.

Microsatellite development process	Number from microsatelite enrichment library (percentages)
1. Number of positive recombinant clones from colony PCR	525
2. Number of clones confirmed by dot bot hybridization	250 (47.61%)
3. Number of clones sent for sequencing	103 (19.61%)
4. Number of sequences containing microsatellite sequences	99 (18.85%)
5. Number of sequences used for designing primers	42 (8.00%)
6. Number of primers producing PCR products	17 (3.23%)
7. Number of polymorphic markers	13 (2.47%)

The number of restriction enzyme and microsatellite probe used are the most critical steps in obtaining microsatellite sequences (Hamilton *et al*., 1999). Only one restriction endonuclease (*Mse*I) and four biotinylated probes: $(CA)_{15}$, $(GA)_{15}$, $(ACC)_{10}$, $(CCT)_{10}$ were used in our study, and 99 (39.60%) out of 250 positive recombinant clones carrying microsatellite DNAs were obtained. Na-Nakorn *et al*. (2006) used *Mse*I and six biotinylated probes: $(AG)_{10}$,$(TG)_{10}$, $(TAC)_{10}$, $(CAA)_{10}$, $(CAG)_{10}$, $(GAT)_{10}$ to develop microsatellite markers for the Mekong giant catfish (*Pangasianodon gigas*) and obtained 53 (64.63%) out of 82 inserted sequences containing microsatellite repeat. Ngamsiri *et al*. (2006) used *Rsa*I, *Hinc*II and *Hae*III with two $(AC)_{15}$ and $(GT)_{15}$ biotinylated probes to develop microsatellite markers for the Mekong giant catfish and found 39 out of 96 positives clones (40.62%) containing microsatellites.

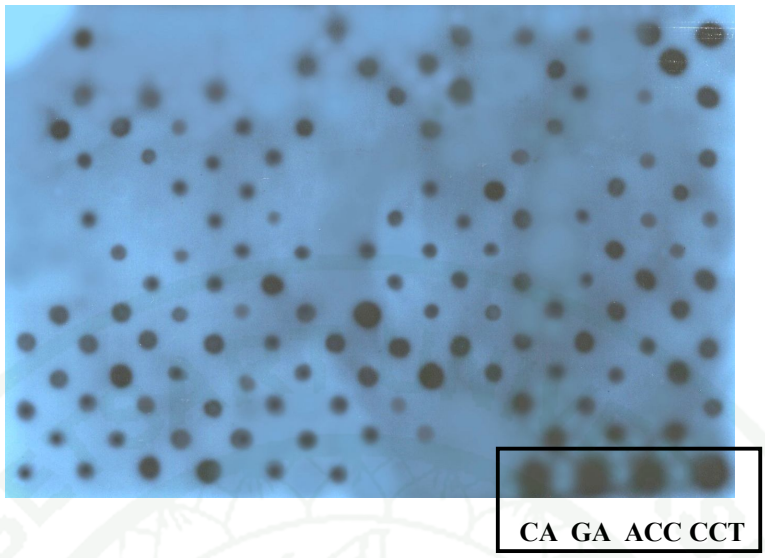


Figure 6 Confirmation of colonies containing microsatellite repeat sequences by dot blot hybridization [positive signal of four oligo-biotinylated probes (CA, GA, ACC, CCT) are shown in the box below the figure].

Microsatellite sequences obtained from our study could be divided into three groups; perfect repeat sequences, $(CA)_n$, $(GT)_n$, $(CCT)_n$, compound repeat sequence (*i.e.,* $(CT)_n(GA)_n$) and interrupted repeat sequences $(A)_nN(CT)_n$, $(GA)_nN(GA)_n$, $(GA)_nN(GGA)_n$, $(GGA)_nN\ (GGA)_n$ where N is the bases interrupting motif sequences. Di-nucleotide repeat was commonly found in whisker sheatfish (Table 2), corresponding with the results shown by Zane *et al*. (2002). They revealed that di-nucleotide was expected more frequency than tri- or tetra-nucleotide repeats in all analysed taxa. Moreover, Chistiakov *et al*. (2006) suggested that the CA/GT repetitive unit is common in fish as also reported by Sukumasavin *et al*. (2004) and Ngamsiri *et al*. (2006) in the Mekong giant catfish, and also by Moeser and Bermingham (2005) for the Neotropical freshwater catfish. Yang *et al*. (2008) developed microsatellite markers for mud carp (*Cirrhina molitorella*). They concluded that 95.56% of the sequences contained CA/GT microsatellites and 4.44% were found to contain CT/GA repeats. In contrast to whisker sheatfish, CA/GT repeat unit was not usually observed but GA/CT was typical microsatellite repeat motif found. Zane *et al*. (2002) revealed that the repeat preference was very different from species to species, for examples, Sanetra and Meyer (2005) found CT repeat unit for the Holarctic freshwater fish and

Chen *et al*. (2010) found GA microsatellite repeat unit most common in the Chinese sucker. Our result could be indicated that GA/CT microsatellite repeat unit was the repeat preference and it was abundance in genome of whisker sheatfish.

From 99 sequences containing microsatellites, forty-two sequences were successfully used to design specific primers flanking the region of microsatellite sequences. These 42 sequences contained microsatellite sequences at the center of the sequences, providing adequate flanking regions for primer development. Seventeen primer pairs were able to amplify the specific DNA fragments. Four markers were monomorphic, whereas 13 markers were polymorphic. Sequence information and characteristics of microsatellite markers are shown in Table 3 and Appendix D.

The number of microsatellite markers that were able to amplify the specific DNA fragments comparing with the total number of sequences containing microsatellites (17 out of 99 sequences; 17.17%) of the whisker sheatfish was slightly lower than other reports on freshwater fish. Moeser and Bermingham (2005) obtained 17 out of 74 sequences (22.97%) for developing microsatellite markers of the Neotropical freshwater catfish *Pimelodella chagresi*. Sukumasavin *et al*. (2004) could obtain 20 out of 83 sequences (24.09%) for developing microsatellite markers of the seven-line barb (*Probarbus jullieni* Sauvage, 1880). In addition, Na-Nakorn *et al*. (2006) reported 15 out of 53 sequences (28.30%) containing microsatellites and could be used to evaluate genetic variation for the endangered Mekong giant catfish, *Pangasianodon gigas*.

Table 2 Groups of microsatellite repetitive unit of whisker sheatfish.

Repetitive unit	Number of sequences
1. Perfect repeat sequence	
Di-nucleotide repeat	
$(CA)_n$ or $(AC)_n$	2
$(GA)_n$ or $(AG)_n$	5
$(GT)_n$ or $(TG)_n$	5
$(CT)_n$ or $(TC)_n$	2
Tri-nucleotide repeat	
$(AGG)_n$ or $(GGA)_n$ or $(GAG)_n$	17
$(CCT)_n$ or $(TCC)_n$ or $(CTC)_n$	2
2. Compound repeat sequence	
Di-nucleotide repeat	
$(CT)_n(GA)_n$	1
3. Interrupted repeat sequence	
Di-nucleotide repeat	
$(CA)_nN(CA)_n$	2
$(CT)_nN(CT)_n$	1
$(TG)_nN(TG)_n$	4
$(GA)_nN(GA)_n$	7
$(TC)_nN(TG)_n$	1
$(GA)_nN(TG)_n$	6
Tri-nucleotide repeat	
$(GGA)_nN(GGA)_n$	11
Multiple motif repeat	
$(A)_nN(CT)_n$	1
$(A)_nN(GT)_n$	1
$(T)_nN(GA)_n$	1
$(CA)_nN(GGA)_n$	2
$(GA)_nN(GGA)_n$	11
$(GT)_nN(GGA)_n$	1
$(CT)_nN(CCT)_n$	1
Total	84

Our results showed that 13 microsatellite markers out of 42 sequences (30.95%) provided polymorphism in whisker sheatfish. The results were not different from those of Ngamsiri *et al*. (2006). They reported 11 polymorphic markers out of 39 sequences (28.21%) in *Pangasianodon gigas*. All microsatellite markers obtained from our study possessed more than 12 repeats of the motif. Microsatellite markers containing motif more than 12 repeats were reported to be able to produce high polymorphic information content (Weber and May, 1990) and produce polymorphism in both interspecies and intraspecies (Smulders *et al*., 1997). Therefore, the method of enrichment established in this project has proved to be efficient and satisfactory for developing microsatellite markers in whisker sheatfish.

Table 3 Characteristics of 13 microsatellite loci isolated from whisker sheatfish.

Locus name	Primer sequence (5′→3′)	Repeat motif	Annealing temp. (°C)	Product size expected	GenBank accession no.
MB71	F: CCTGGACAAACGCTCCAAAC	$(TG)_7A(TG)_6$	57	199	GQ128432
	R: GCTCCTGCTTCACGAACAC				
MB79	F: GTTATCTGTCTTGTGTAGACG	$(GA)_{20}N$	57	297	GQ128432
	R: AACACTCCTGTCTGTCCAGCC	$(GA)_{15}$			
MB81	F: CAGCAGCAAGAAGCAGACG	$(GT)_{15}$	57	161	GQ128433
	R: CAAGGTCAGTAGGGAAGTGTG				
MB320	F: AGAAGCAGCTCGGGTCTGC	$(GAG)_{21}$	60	216	GQ128436
	R: TCCATGGCCTGGTTGAACTC				
MB354	F: ACAACACAATGAGCGATCGAC	$(A)_{16}N$	59	176	GQ128437
	R: ACGGCCTCGGAAATCACTG	$(TC)_{19}$			
MB381	F: CCTGGACAAACGCTCCAAAC	$(CT)_{19}$	57	134	GQ128432
	R: GCTCCTGCTTCACGAACAC				
MB401	F: TAGAGATCAAGAGACCTTTAG	$(AG)_{13}$	61	119	GQ128439
	R: CACTCTTTCATCTGTGCGG				

Table 3 (continued)

Locus name	Primer sequence (5′→3′)	Repeat motif	Annealing temp. (°C)	Product size expected	GenBank accession no.
MB456	F: GCTGATCTAAAGTTCATTAGC R: ACACCATACTGCAGGAGG	$(GA)_{20}$ $N(T)_{11}$	56	166	GQ128440
MB515	F: CCTGGACAAACGCTCCAAAC R: GCTCCTGCTTCACGAACAC	$(AG)_{19}N$ $(TG)_{13}$	57	196	GQ128432
MB613	F: CCTGGACAAACGCTCCAAAC R: GCTCCTGCTTCACGAACAC	$(GT)_{12}(GA)_8$	57	162	GQ128432
MB614	F: CCTGGACAAACGCTCCAAAC R: GCTCCTGCTTCACGAACAC	$(CA)_7N$ $(GGA)_5$	57	249	GQ128432
MB618	F: CCTGGACAAACGCTCCAAAC R: GCTCCTGCTTCACGAACAC	$(GT)_{11}$	57	133	GQ128432
MB645	F: GAAGGATGTAACTATGGAAGC R: ATACGGAGTGATGTGTTAGAG	$(GA)_{20}$	57	121	HQ660229

F is forward primer; R is reverse primer; N is the bases interrupting motif sequences.

3. Genetic diversity analyses of whisker sheatfish populations

The analysis consisted of six parts; microsatellite scoring error analysis and populations defined, Hardy-Wienberg and linkage equilibrium analyses, genetic diversity within populations, genetic diversity among populations and population structure, migration rate and migration pattern and finally effective population size and population bottleneck.

3.1 Microsatellite scoring error analysis and population defined

Error in scoring the microsatellite bands can occur due to the presence of bands stuttering, large allele drop out and null alleles. From our study, null alleles were found in all populations studied. MB71 and MB381 presented strong evidence of null alleles for all populations, whereas the remaining loci exhibited null allele in some populations (Table 4).

The presence of null alleles in our study was not different from other reported in fish. Chistiakov *et al*. (2006) indicated that microsatellite null alleles created the heterozygote deficits, and have been found in a wide range of taxa and very common in fish. Ngamsiri *et al*. (2006) concluded that the heterozygote deficits could be due to null alleles in the Mekong giant catfish. So *et al*. (2006) suggested that the occurrence of null alleles has been regularly recorded as a major explanation for the observed deficit of heterozygotes in the migratory sutchi catfish. Ha *et al*. (2009) reported the main reason for departure of HWE was null alleles in striped catfish (*Pangasainodon hypophthalmus*).

Dakin and Avise (2004) suggested that one potential cause of microsatellite null alleles was poor primer annealing due to primer/template mismatch. Mismatch annealed region was caused by extremely high levels of DNA sequence variation or nucleotide sequence divergence (*i.e.* involving point mutations or indels) in one or both priming sites. Furthermore, Paetkau and Strobeck (1995) reported that mutations at the priming regions would cause some individuals to have

only one allele amplified, or failure to amplify both alleles, leading to false scoring as homozygotes of heterozygotes.

Table 4 The presence of null allele at each locus in each population studied.

Loci/Populations	NH	KB	MK	MG	SK	SE	NA
MB71	✓	✓	✓	✓	✓	✓	✓
MB79	✓		✓		✓		
MB81			✓	✓	✓	✓	
MB320		✓	✓				
MB354					✓	✓	
MB381	✓	✓	✓	✓	✓	✓	✓
MB401			✓				
MB456		✓			✓		
MB515			✓	✓		✓	
MB613			✓		✓		
MB614			✓				
MB618			✓		✓	✓	
MB645				✓	✓	✓	

To overcome null allele problem, new primers were designed and PCR condition was adjusted. Fortunately, the problem of null alleles could be eliminated from almost all microsatellite loci, except for MB71 and MB381 loci. Lehmann *et al*. (1996) suggested that the redesigning primers to bind to different regions of the flanking sequence, or adjusting PCR conditions could often ameliorate null alleles. Nonetheless, null alleles still remained significant in some cases and as some null genotypes were still observed even redesigning primers or adjusting PCR condition had been performed. Callen *et al*. (1993) reported that types of microsatellite repeat unit might have an effect on generation of null alleles, for example AC repeat unit usually promoted null alleles in human. Lehmann *et al*. (1996) revealed that PCR failure was due to an A to G change at the priming sites. In addition, van Treuren (1998) identified two clones that differed only in a G to T transversion in the

microsatellite flanking regions leading to null alleles in the oystercatcher (*Haematopus ostralegus*). Moreover, Ha *et al*. (2009) concluded that the adjust of null alleles could not conform all populations to HWE. Therefore, MB71 and MB381 loci were excluded from population genetic analyses because the problem of null alleles.

3.2 Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD)

Deviations from Hardy-Weinberg equilibrium (HWE) after Sequential Bonferroni correction (P<0.0045) were detected across loci and populations, except population of NA. In addition, pair-wise linkage disequilibrium (LD) was found in 4 populations (P<0.0045). Fifteen of 55 LD were observed in MK population (P<0.0045). One of 55 LD was observed in MG population. Six of 55 LD were observed in SK population and one of 55 LD was observed in SE population but none was found in NH, KB and NA populations.

When the null hypothesis of HWE was not applicable, like in our result, the F_{IS} statistic of Wright (1951) was estimated following Weir and Cockerham (1984) and used as an indicator of heterozygote deficit. In this study, all populations except NA showed positive value F_{IS}, indicating homozygote exceed. Surprisingly, MK and SK populations showed positive value of F_{IS} for all loci. Angel *et al*. (2006) suggested that not only inbreeding but also Wahlund effect might cause homozygote exceed, resulting in deviation from HWE. The result of COLONY did not show strong evidence for full- and half-sib relationships for all whisker sheatfish populations. Forty-five families were found in 7 populations. Ten families carried many members obtained from one population. Therefore, these families were re-sampling by random to discard the sampling bias. In addition, if deviation from HWE was caused from Wahlund effect, the individual should be defined for obscured genotype. In this study, the result of Bayesian Analysis of Population Structure (BAPS) showed that some individuals gave ambiguous genotype. Once individuals have been identified, it is imperative that the test of these should be repeat in the laboratory to ensure that they are valid prior to further analyses. After individual defined, some genotype data was eliminated because they were ambiguous source of

whisker sheatfish. Finally, 28 genotypes data of NH population, 26 genotypes data of KB population, 21 genotypes data of MG population and 24 genotypes data of SE population were used for further analyses. After adjusting genotypic data, five populations (NH, KB, MG, SE, NA) were found to conform to Hardy-Weinberg and linkage equilibrium. However, both MK and SK populations still deviated from HWE and showed some LD.

From the result obtained, Wahlund effect was the major cause of deviation from HWE for MK and SK populations. Chakraborty *et al*. (1992) revealed that the Wahlund effect refers to the reduction of heterozygosity in a population caused by population substructure. In addition, the occurrence of LD confirmed population substructure as mentioned by Na-Nakorn and Moiekum (2009). They found the presence of non-random association of loci of striped catfish might be attributed to mixing of more than one genetically distinct population in sample. For our study, when gene flow was restricted, whisker sheatfish populations might show different level of allele frequency. Therefore, one collected site such as main river and large tributaries might consist of several sources of whisker sheatfish populations, leading to obtaining the different allelic frequencies. Generally, if two or more sub-populations have different observed allelic frequencies, then the overall heterozygosity was reduced and generated the Wahlund effect or population substructure. Finally, MK and SK populations were excluded prior to population genetic analyses. However, some genetic diversity index, *i.e*. heterozygosity, number of allele per locus, allelic richness could be reported to evaluate the genetic diversity at those collecting sites.

3.3 Genetic diversity of whisker sheatfish populations

In this study, genetic variation in 189 individuals of 7 populations of whisker sheatfish were investigated using 11 microsatellite loci.

3.3.1 Allele size range

Allele size range of across 11 microsatellite loci were 101-288 bp. Allele size range for each locus are shown in Appendix C. The allele size range of our experiment was not different from the expected size range (Table 3) except at MB320 locus. This locus contained tri-nucleotide repeat and generated larger size range than the expected size. Schlotterer and Harr (2001) suggested that tri-nucleotide repeat mutations encompassed changes by increasing several repeat units and gains was significantly more likely than losses. Thus, the increase in number of repeat units was not surprising for MB320 locus.

3.3.2 Allelic diversity

The numbers of alleles for each population were 26, 29, 61, 36, 43, 35 and 33 for NH, KB, MK, MG, SK, SE and NA populations, respectively (Table 4). Seventy-one alleles were obtained across loci and populations. Mean number of allele per locus (*A*) were 2.36, 2.63, 5.54, 3.27, 3.90, 3.18 and 3.00 for NH, KB, MK, MG, SK, SE and NA populations, respectively. Number of alleles across populations (N) was found ranging from 3 (MB320) - 11 (MB613). In addition, the average of allelic diversity across loci and populations was 6.45. Allele distribution patterns are shown in Appendix E. The average allelic diversity of whisker sheatfish was lower than other freshwater (9.1±6.1), anadromous (10.8±7.2) and marine fish (19.9±6.2) (DeWoody and Avise, 2000). However, allelic diversity of whisker sheatfish was still more than the critically endangered giant catfish (3.2) (Ngarmsiri *et al*., 2007). Moreover, all loci except MB320 contained at least four alleles. Barker (1994) suggested that, for the microsatellite selection standard, any locus that had at least four alleles would be excellent for genetic evaluation. Therefore, our microsatellite markers were useful to be used in population genetic study of whisker sheathfish.

Considering all populations, NH population displayed lower genetic diversity than other populations. This location was isolated from other sites by drainage construction. Frankham *et al*. (2002) suggested that random genetic drift

which predominated in small and isolated population influenced the loss of allelic diversity. Therefore, NH population was proned to the reduction in allelic diversity by random genetic drift. Good management of whisker sheathfish was needed for protecting genetic resource at this location.

The dominant alleles were found in many loci. MB320 (285 bp), MB354 (164 bp), MB515 (194 bp), MB613 (152 bp), MB614 (249 bp) and MB645 (111 bp) exhibited the dominant allele at the frequency of 0.8 (Figures 7, 8, 9, 10, 11 and 12). In terms of effective number of allele (A_e), it ranged from 1.40 (KB) to 3.20 (MK) with a mean of 1.86. This result was lower than in the striped catfish (5.0 and 7.4) (Na-Nakorn and Moeikum, 2009) and the critically endangered species, the Mekong giant catfish (1.9) (Ohashi *et al*., 2007). Frankham *et al*. (2009) revealed that when all alleles are equally frequent, the effective number of allele should be very close to the actual value. It means that all of the alleles were equally important for those populations and provided long-term genetic diversity for populations.

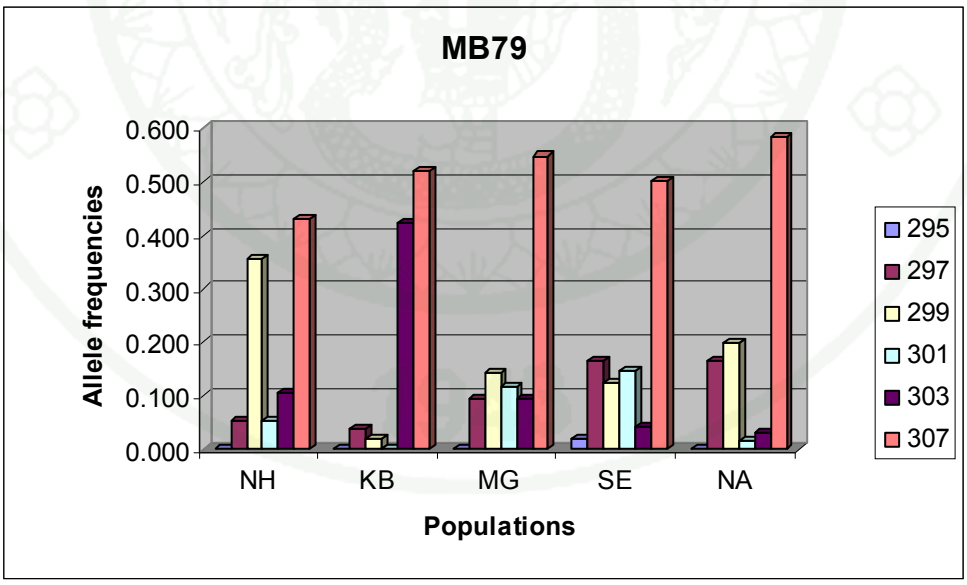


Figure 7 Allele frequencies of MB79 locus for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

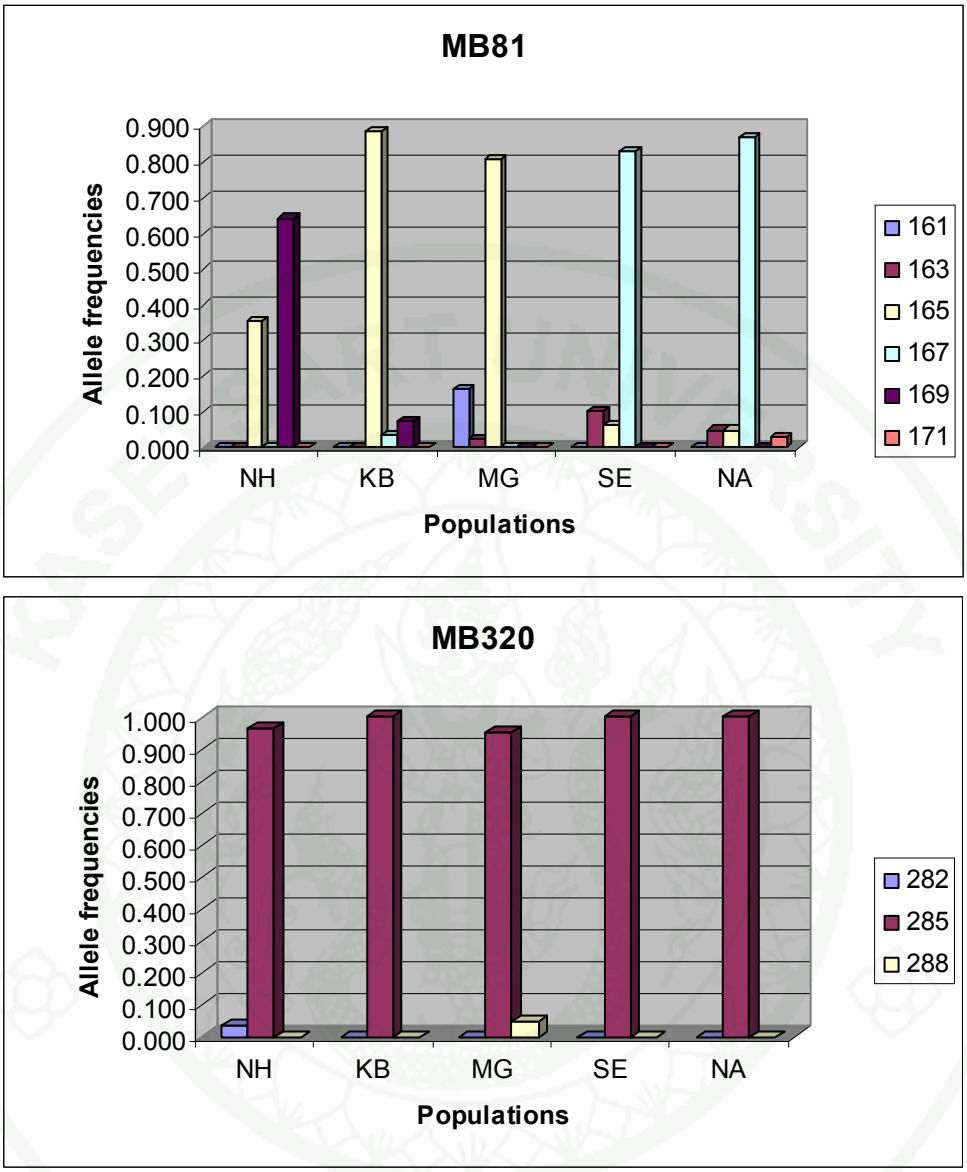


Figure 8 Allele frequencies of MB81 and MB320 loci for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

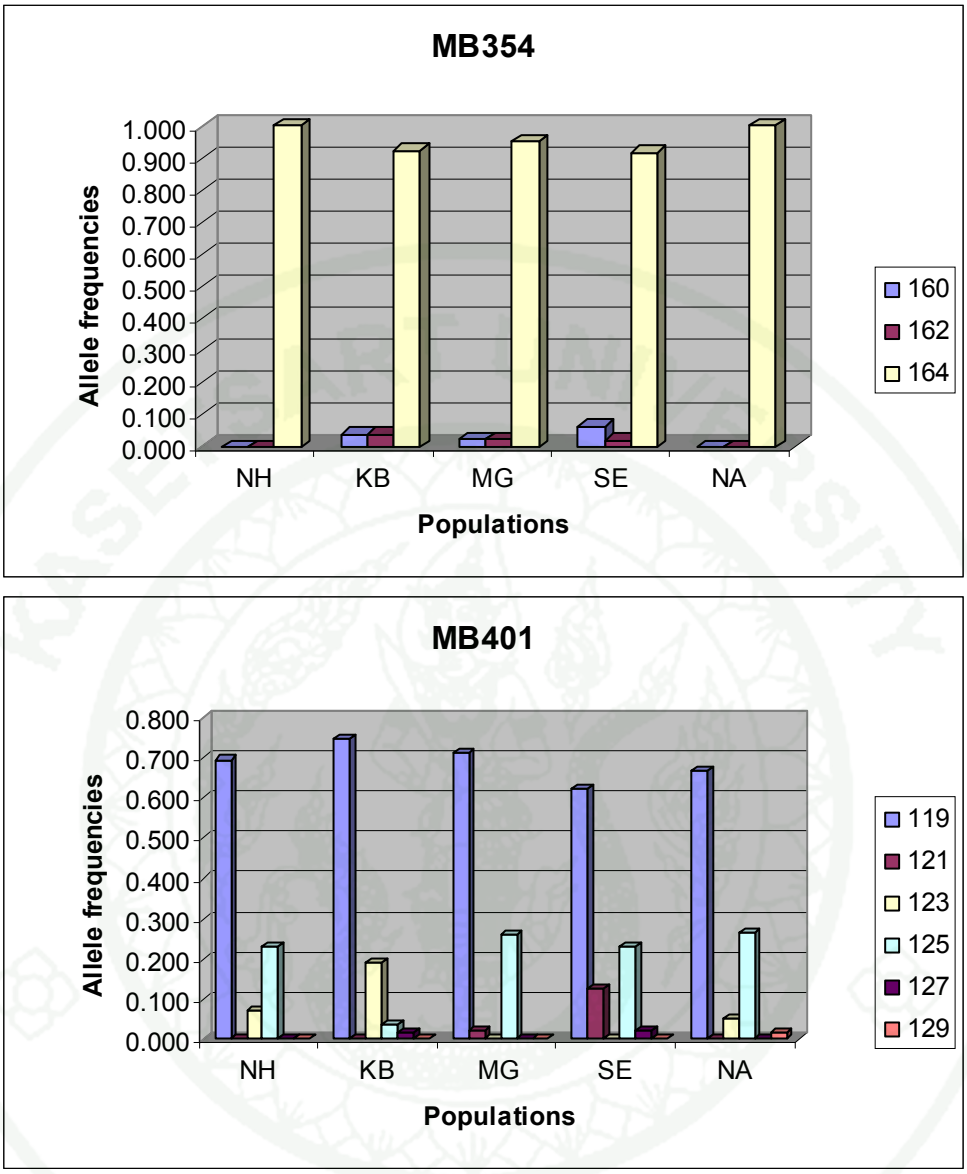


Figure 9 Allele frequencies of MB354 and MB401 loci for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

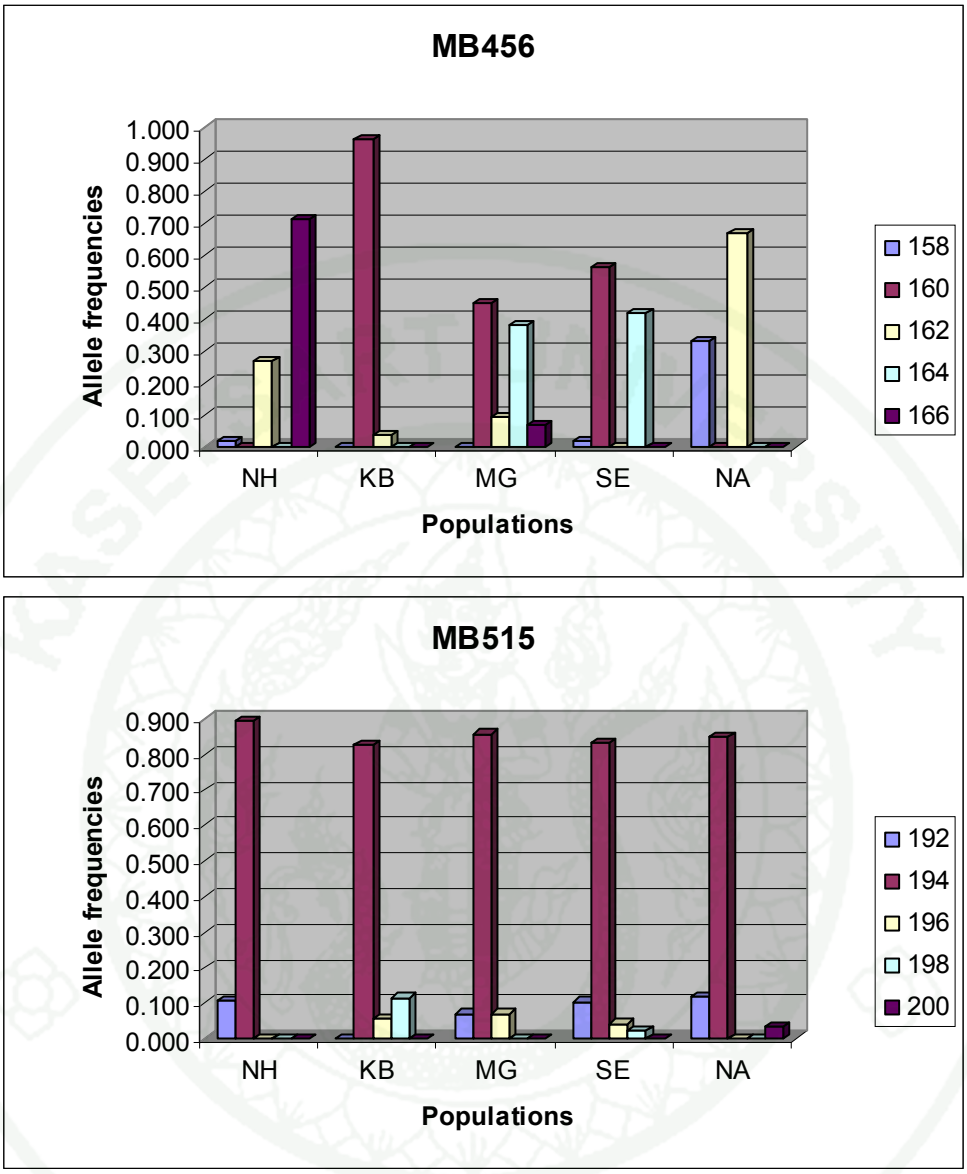


Figure 10 Allele frequencies of MB456 and MB515 loci for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

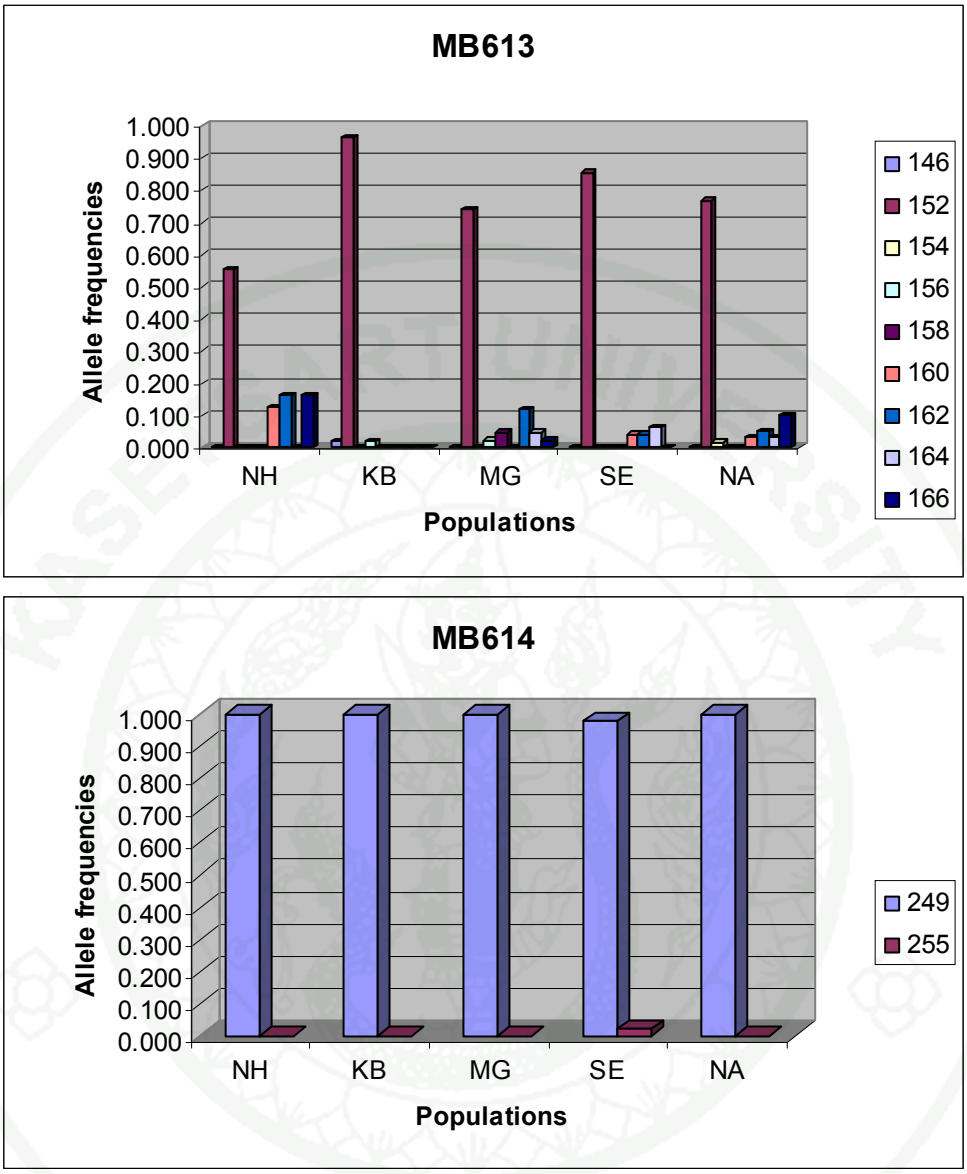


Figure 11 Allele frequencies of MB613 and MB614 loci for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

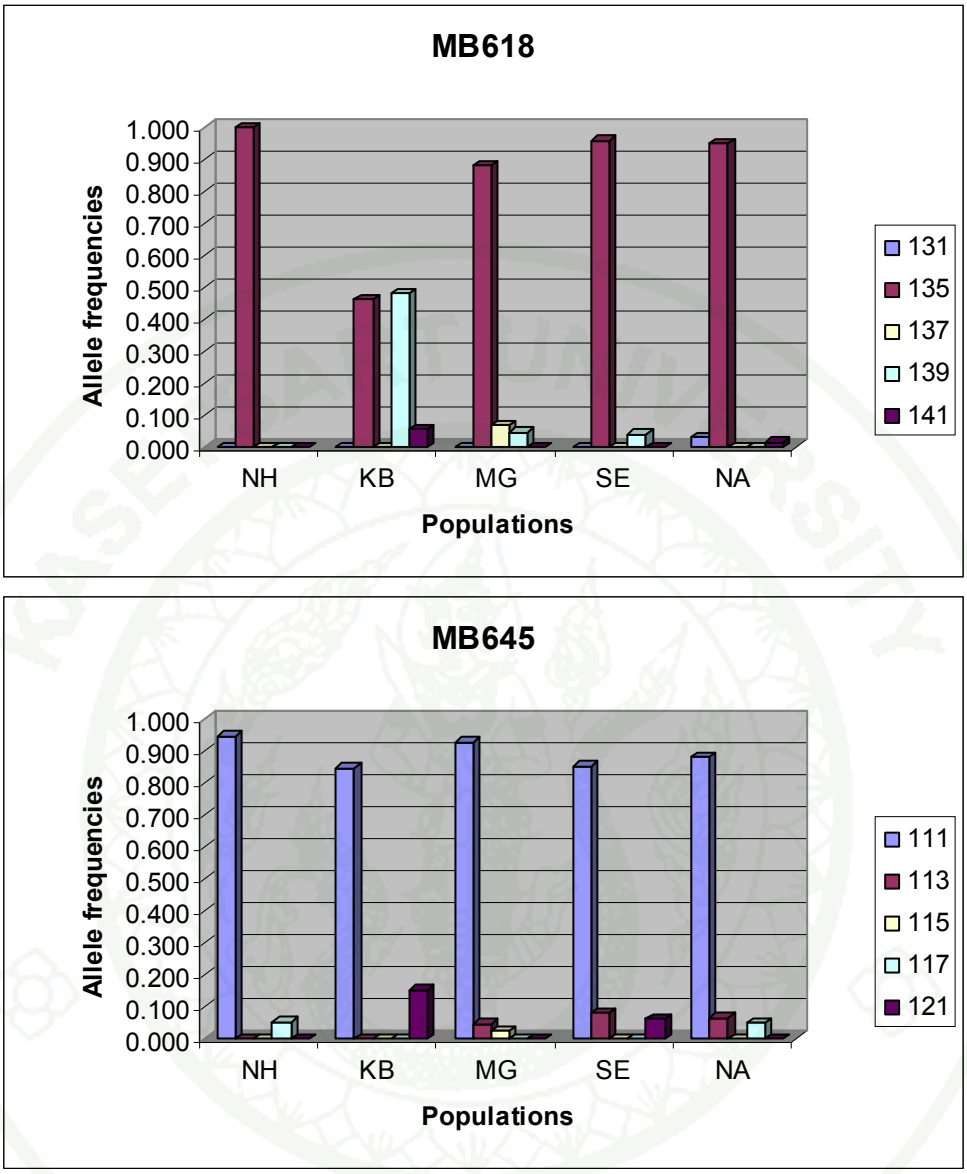


Figure 12 Allele frequencies of MB618 and 645 loci for 5 whisker sheatfish populations (the number in the box is allele size in the unit of bp).

Table 5 The number of allele per locus (A), allelic richness (A_R), effective number of allele (A_e) and number of allele per locus across populations (N) of all loci and populations.

Locus	A/ A_R/ A_e							N
Name	NH	KB	MK	MG	SK	SE	NA	
MB79	5/ 4.97/ 3.04	4/ 3.77/ 2.22	7/ 6.39/ 3.58	5/ 5.00/ 2.83	6/ 5.67/ 3.47	6/ 5.86/ 3.15	5/ 4.61/ 2.44	8
MB81	2/ 2.00/ 1.84	3/ 2.96/ 1.26	6/ 5.95/ 3.94	3/ 3.00/ 1.46	4/ 3.99/ 2.97	3/ 2.99/ 1.41	4/ 3.86/ 1.32	6
MB320	2/ 1.94/ 1.07	1/ 1.00/ 1.00	3/ 2.99/1.78	2/ 2.00/ 1.09	2/ 1.97/ 1.10	1/ 1.00/ 1.00	1/ 1.00/ 1.00	3
MB354	1/ 1.00/ 1.00	3/ 2.93/ 1.16	4/ 3.99/ 2.62	3/ 3.00/ 1.10	5/ 4.97/ 3.16	3/ 2.87/ 1.18	1/ 1.00/ 1.00	6
MB401	3/ 2.99/ 1.83	4/ 3.77/ 1.66	6/ 5.56/ 1.92	3/ 3.00/ 1.72	2/ 2.00/ 1.92	4/ 3.87/ 2.17	4/ 3.67/ 1.92	8
MB456	3/ 2.75/ 1.71	2/ 1.96/ 1.07	5/ 4.90/ 3.22	4/ 4.00/ 2.74	4/ 3.91/ 2.62	3/ 2.87/ 2.03	2/ 2.00/ 1.80	5
MB515	2/ 2.00/ 1.23	3/ 2.99/ 1.42	5/ 4.67/ 3.24	3/ 3.00/ 1.34	5/ 4.52/ 1.79	4/ 3.86/ 1.41	3/ 2.91/ 1.35	5
MB613	4/ 4.00/ 2.67	3/ 2.61/ 1.08	8/ 7.36/ 4.15	6/ 6.00/ 1.77	5/ 4.67/ 2.86	4/ 3.97/ 1.35	6/ 5.50/ 1.65	11
MB614	1/ 1.00/ 1.00	1/ 1.00/ 1.00	4/ 3.70/ 1.88	1/ 1.00/ 1.00	1/ 1.00/ 1.00	2/ 1.87/ 1.04	1/ 1.00/ 1.00	3
MB618	1/ 1.00/ 1.00	3/ 2.99/ 2.23	5/ 4.91/ 2.96	3/ 3.00/ 1.27	3/ 2.67/ 1.43	2/ 1.98/ 1.08	3/ 2.61/ 1.10	6
MB645	2/ 1.98/ 1.10	2/ 2.00/ 1.35	8/ 7.61/ 5.92	3/ 3.00/ 1.15	6/ 5.58/ 2.82	3/ 2.99/ 1.35	3/ 2.97/ 1.27	10
Mean	2.36/ 3.33/ 1.58	2.63/ 2.54/ 1.40	5.54/ 5.27/ 3.20	3.27/ 3.27/ 1.59	3.90/3.72/ 2.29	3.18/ 3.10/ 1.56	3.00/ 2.83/ 1.44	71

The lower effective number of allele than the number of alleles in whisker sheatfish indicated that some alleles were dominant and could lead to fix in population, whereas the allele which had low allele frequency might be risk to be lost from population leading to reduction in allelic diversity especially in small or isolated whisker sheatfish populations. Moreover, Frankham *et al*. (2002) suggested that low effective number of allele reflected the high number of low frequency alleles. Under a severe random genetic drift, low frequency adaptive alleles might also be lost from the population, which might reduce the ability of populations to respond to the environmental changes. Whisker sheatfish population exhibited low allelic diversity and dominant alleles. Thus, these populations might lose their genetic diversity in the short-term.

In this experiment, long distance (Table 9) between collecting sites of whisker sheatfish and geographical barrier resulted in low rate of gene flow between populations. Bartfai *et al*. (2003) suggested that the presence of private alleles in the populations could be caused by restricted gene flow among populations, or low levels of genetic exchange between populations. The number of private alleles over seven populations of whisker sheatfish was relatively low. Two private alleles were detected at MB401 (129 bp) and MB618 (131 bp) for only NA population. Slatkin (1985) defined the private alleles to be those which are present in one population and absent from all others tested. Therefore, these private alleles might be used for NA specific marker. In this study, private alleles were detected at locus MB401 and MB618 with the frequencies of 0.0167 and 0.0333, respectively. The number of private alleles and allele frequency was not different from the results reported by Ha *et al*. (2009). They found three low frequencies of private alleles in wild populations and only one private allele in hatchery population of striped catfish. However, Allendorf and Luikart (2007) suggested that rare alleles (low allele frequency) were those to be lost first under high selection intensity, random genetic drift and population bottleneck events. In addition, Garza and Williamson (2001) concluded that low frequency of private alleles indicating of a reduction in genetic diversity.

Allelic richness was calculated based on minimum sample size of 21 diploid individuals (MG population). Mean of allelic richness for each population was 2.33, 2.54, 5.27, 3.27, 3.72, 3.10 and 2.83 for NH, KB, MK, MG, SK, SE and NA populations, respectively. Allelic richness across loci and populations was quite low (3.29) comparing with previously published surveys of freshwater fish. So *et al*. (2006) revealed that the average allelic richness of migratory sutchi catfish was 9.9. Also Na-Nakorn and Moeikum (2009) reported the allelic richness across loci and population of striped catfish in Vietnam was 5.6. In terms of allelic richness, Foulley and Ollivier (2006) suggested that allelic richness was particularly important for a long-term perspective, because it reflected better past fluctuations in population size.

3.3.3 Heterozygosity

The values of observed heterozygosity (H_o) across loci of whisker sheatfish populations was found ranging from 0.223 (KB) to 0.415 (MK) with the average H_o of 0.279, whereas the expected heterozygosity (H_e) across loci was in the range of 0.237 (KB) to 0.657 (MK) with the average H_e of 0.357 (Table 6). The mean heterozygosity was 0.318 that was lower than other freshwater (0.54±0.25), anadromous (0.68±0.12) and marine fish (0.77±0.19) (DeWoody and Avise, 2000). Moeser and Bermingham (2005) found that observed and expected heterozygosity of eight microsatellite loci in Neotropical freshwater catfish was 0.73 and 0.85, respectively. Ha *et al*. (2009) reported that observed and expected heterozygosity of five microsatellite loci in striped catfish was 0.64 and 0.63, respectively. However, the heterozygosity of whisker sheatfish was less than the critically endangered giant catfish. Na-Nakorn *et al*. (2006) found that observed and expected heterozygosity using ten microsatellite loci was 0.55 and 0.49, respectively. Ohashi *et al*. (2006) concluded that expected heterozygosity was 0.41 in founder. Ngarmsiri *et al*. (2007) found that observed and expected heterozygosity using ten microsatellite loci was 0.425 and 0.472, respectively. In terms of heterozygosity, whisker sheatfish populations showed low level. It implied that these populations had little genetic variability. Thus, they were vulnerable to be lost from those collecting site.

Table 6 Observed and expected heterozygosity and *P*-value after Sequential Bonferroni correction.

Populations	H_o	H_e	Bonferroni corrected *P*-value
NH	0.266	0.277	0.022
KB	0.223	0.237	0.015
MK	0.415	0.657	0.000*
MG	0.264	0.302	0.008
SK	0.260	0.489	0.000*
SE	0.272	0.289	0.046
NA	0.254	0.252	0.339
Mean	0.279	0.357	na

* Deviation from HWE after Sequential Bonferroni correction (P<0.0045).
H_o is observed heterozygosity. H_e is expected heterozygosity.
na is not applicable.

In summary, the mean value of A, A_R, A_e, H_o and H_e were 3.41, 3.43, 1.86, 0.279 and 0.357, respectively. It implied that genetic diversity at microsatellite loci of whisker sheatfish was characterized by low, comparing with previously published surveys of freshwater fish population variability such as those of DeWoody and Avise (2000); Quan *et al*. (2006); So *et al*. (2006); Li *et al*. (2007); Ngamsiri *et al*. (2007) and Ha *et al*. (2009). Therefore, whisker sheatfish population from this study faced the decline genetic diversity.

3.4 Genetic diversity among populations and population structure

3.4.1 Inbreeding coefficient (F_{IS})

In this experiment, whisker sheatfish populations displayed the positive value of inbreeding coefficient. NH population showed positive F_{IS} value of

3 loci. KB population exhibited positive F_{IS} value of 4 loci. MG population revealed positive F_{IS} value of 6 loci. SE population showed positive F_{IS} value of 5 loci. NA population exhibited positive F_{IS} value of 3 loci (Table 7). In addition, a broad variation was found on the F_{IS} value among populations, ranging from NA (-0.009) to MG (0.129). Generally, positive values of F_{IS} indicate a deficit of heterozygotes and negative values indicate an excess of heterozygotes (Frankham *et al*., 2002). In addition, exceed of homozygote could be explained by non-random sampling (sample bias), intra-population structure (Wahlund effect) (Castric *et al*., 2002), founder event (inbreeding), and existence of null alleles (Angel *et al*., 2006). Inbreeding effect was found in many freshwater fish. Parra-Bracamonte *et al*. (2011) found inbreeding evidence in traditional channel catfish. Na-Nakorn and Moeikum (2009) found in stock of striped catfish. In our experiment, although positive F_{IS} value was not found in every locus (Table 7) but inbreeding coefficient showed positive value for all whisker sheatfish populations except only in NA population.

3.4.2 Population differentiation (F_{ST})

Pairwise matrix of F_{ST} values was calculated based on Weir and Cockerham (1984). The level of F_{ST} was ranged from 0.15 (KB vs. MG; SE vs. NA) to 0.40 (KB vs. NA) (P<0.01) with mean F_{ST} value of 0.26 (Table 8). This value indicated very high genetic differentiation among populations. Generally, when F_{ST} = 0, it means all sub-populations are having the same gene frequencies with no variance among sub-populations and no Wahlund effect (Hoarau *et al*., 2002).

Table 7 Inbreeding coefficient (F_{IS}) for all loci and populations of whisker sheatfish after populations defined.

	Inbreeding coefficient (F_{IS})				
Populations	NH	KB	MG	SE	NA
MB79	0.378	0.110	-0.153	-0.077	-0.169
MB81	-0.385	-0.079	0.565	-0.125	0.056
MB320	-0.019	na	1.000	na	na
MB354	na	-0.042	-0.013	0.216	na
MB401	0.156	-0.236	0.118	-0.057	-0.090
MB456	-0.093	1.000	0.051	0.203	-0.033
MB515	-0.102	0.248	0.459	0.166	0.129
MB613	0.049	-0.010	-0.070	-0.088	-0.074
MB614	na	na	na	0.000	na
MB618	na	0.184	-0.075	-0.022	-0.024
MB645	-0.038	-0.163	0.661	0.534	0.542
All	0.039	0.060	0.129	0.058	-0.009

na is not applicable.

The mean F_{ST} value obtained in whisker sheatfish (0.26) was higher than the average degree of genetic differentiation estimated for 49 species of freshwater fish (G_{ST} =0.222) (Ward *et al*., 1994). So *et al*. (2006) reported significant F_{ST} in the Mekong river sutchi catfish ranging from 0.0009 to 0.0128. Also Pereira *et al*. (2009) found significant F_{ST} in the migratory catfish *Pseudoplatystoma corruscans* ranging from 0.022 to 0.151. In general, the level of genetic differentiation of freshwater fish was higher than in marine and anadromus fish. Freshwater fishes in river systems often exhibit population structure that follows the isolation by distance model (Makinen *et al*., 2006). The degree of population differentiation depended on population size, the rate of gene flow and migration. In this experiment, high degree

of F_{ST} value between whisker sheatfish populations was observed due to long distance and geographical barrier between populations and also small effective population size of all whisker sheatfish populations.

Table 8 Pairwise matrix of F_{ST} values (below diagonal) and correction for multiple comparisons ($P < 0.01$) and distance in Km. (above diagonal) among 5 whisker sheatfish populations.

Populations	NH	KB	MG	SE	NA
NH	-	78.81	180.65	396.97	468.44
KB	0.35**	-	101.84	318.16	389.63
MG	0.19**	0.15**	-	216.32	287.79
SE	0.28**	0.29**	0.16**	-	293.96
NA	0.26**	0.40**	0.26**	0.15**	-

** Highly significant F_{ST} value (P<0.01 from FSTAT)

Surprisingly, the F_{ST} between NH and MG (F_{ST} = 0.19) was lower than F_{ST} between NH and KB (F_{ST} = 0.35) comparing with the distance between sites. This indicated that gene flow between NH and KB was restricted. Generally, the Nong-Han Lake (upstream) is connected to Nam Kum basin and water flows to the Mekong river (downstream). Most whisker sheatfish populations in the Nong-Han Lake and Nam Kam basin move to the Mekong river for feeding ground. On the other hand, whisker sheatfish from the Mekong river moves back to Nam Kum basin and the Nong-Han Lake for spawning ground. In this case, Nam Kam basin was only fish passageway. Therefore, low level of gene flow was observed between NH and KB, whereas high level of gene flow was observed between NH and MG.

Additionally, the F_{ST} between SE and MG (0.16) populations was lower than F_{ST} between NA and MG (0.26) populations. This could be described that the whisker sheatfish populations could move through Songkhram river basin to the Mekong river. The Songkhram river basin is their spawning ground, whereas the

Mekong river is their feeding ground. In addition, Lam Nam Oon is a small tributary of Songkhram river basin which is located far away from the Mekong river and in the dry season this area becomes shallow and acts as geographical barriers among whisker sheatfish populations. Therefore, MG and SE populations had a chance to exchange gene between them whereas NA population was interrupted by the natural barrier and long distance.

In addition, the value of F_{ST} observed between MG and KB was 0.15, and between MG and SE to be 0.16 which were lower than comparing with other sites. Naturally, the Mekong river is connected to both Nam Kam basin and Songkhram river basin. Whisker sheatfish among these locations could migrate freely for feeding and spawning grounds, resulting in high rate of gene flow among populations. Meffe and Vrijenhoek (1988) proposed that the influence of the barriers on gene flow generally results in a genetic signature where populations within the same tributaries or connected areas are more genetically similar than the distant regions.

Futhermore, F_{ST} values between KB and NA (0.4) was higher than that of NH and NA (0.26). In the past, Lam Nam Oon was connected to the Nong-Han lake which acts as a spawning ground, whisker sheatfish populations could move along between sites and passed to Nam Kum basin (see migration rate). Gene flow could occur among three populations. At present, the Nam Oon dam and drainage structure are installed, the dam separate the Nong-Han lake and Lam Nam Oon into two disconnected sites. However, some alleles derived from the same common ancestor are still kept within both NH and NA populations whereas some alleles have already lost from KB population.

From the results obtained, the high F_{ST} values between populations suggested that migration between populations were restricted. Inbreeding should have an effect on these isolated populations. Therefore, fishery and genetic management of whisker sheatfish should be implemented to prevent genetic erosion in the near future.

3.4.3 Population structure

STRUCTURE and BAPS program were performed to define population structure. The result from both programs confirmed population substructure for MK and SK populations. The STRUCTURE exhibited that whisker sheatfish populations collected from the Mekong river (MK) was divided into 5 clusters with log probability of data -636.8 and 8 clusters with log probability of data -473.4 for Songkhram river basin (SK) (Figure 13). In addition, BAPS showed 6 clusters for MK population and 5 clusters for SK population.

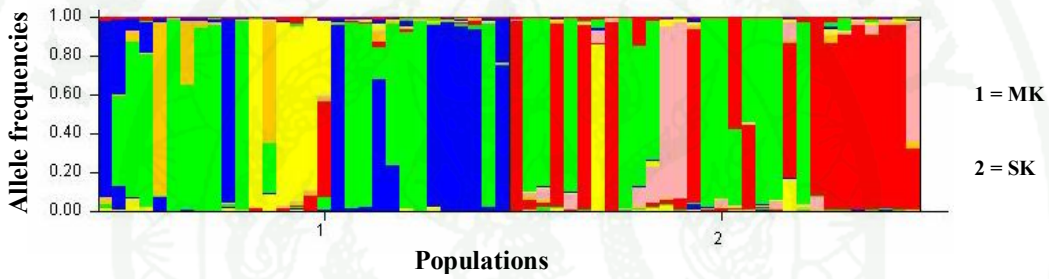


Figure 13 Bayesian analysis of the genetic structure based on eleven microsatellite loci in bar plot. Each individual is represented by a vertical line, which is coloured according to the assigned groups at estimated $K = 6$ (MK); $K = 8$ (SK); (x axis is populations and y axis is allele frequency).

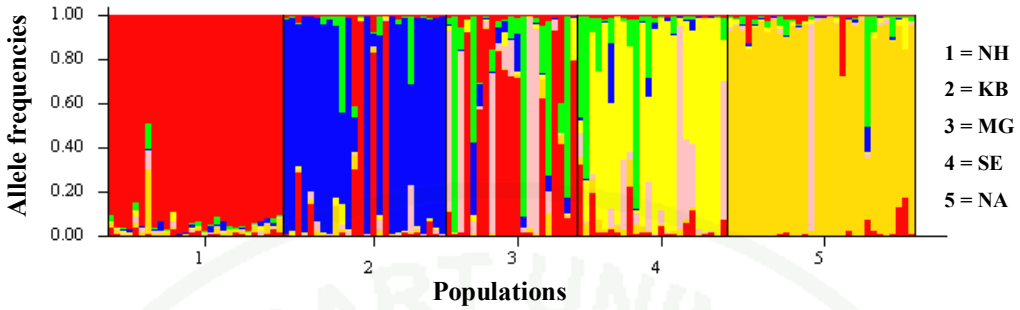
The Mekong and Songkhram basin are large and long river generating varities of fish communities. Whisker sheatfish is the migratory fish which has the potential to migrate over extensive distances. From the results obtained, the collections at one location may be composed of individuals originating from more than one spawning area, generating the differences in allele frequencies. Individuals of MK could be composed of individuals from other populations such as NH, KB and unknown sites whereas SK population consisted of MG, NA and unknown sites. Our results agreed with those reported by Pujolar *et al*. (2011) in marble trout. They found that when other individuals with different allele frequency were introgressed or added into the population, population mixture occurred. Furthermore, Arnold (1999) and Hatanaka *et al*. (2006) reported that most commercial fish were long-distance

migratory species and could exhibit population sub-structuring. In addition, Duncan and Lockwood (2001) suggested that the extent of structuring could be influenced by natural and artificial barriers. Therefore, MK and SK populations exhibiting strong sub-structuring were eliminated from further analysis of population structure.

For five remaining populations, the most likely value for *K* is 5 based on the highest average estimated likelihood value of -1554.2. Further increasing of *K* up to 7 did not give new information. Therefore, our results exhibited the 5 clusters that mostly corresponded to geographical areas of collecting sites. The three clusters, KB, MG and SE, showed many different plotting (Figure 14). The results indicated that gene flow of those three populations might occur. The level of gene flow corresponded with the moderately level of F_{ST} among the three populations which were at 0.15 for MG and KB and at 0.16 for MG and SE.

In the context of gene pool conservation, Waits *et al*. (2001) and Cegelski *et al*. (2003) suggested that different genetic populations must be considered as independent units to be managed if one wants to conserve the genetic diversity of a species. Therefore, all of whisker sheatfish populations, which generated low genetic diversity, and high level of genetic differentiation were concerned for conservation and management strategies.

(a)



(b)

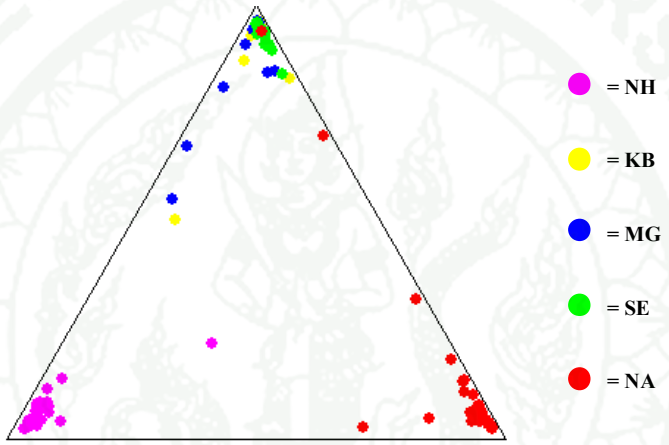


Figure 14 Bayesian analysis of the genetic structure based on eleven microsatellite loci in bar plot. Figure (a): each individual represented by a vertical line, which is coloured according to the assigned groups at estimated $K = 5$ (x axis is populations, y axis is allele frequency). Figure (b): each individual represented by dot, which is coloured to the assigned groups at estimated $K = 5$.

3.5 Migration rate and migration pattern

In this experiment, the historical migration rate of whisker sheatfish was estimated by MIGRATE-n (Beerli, 2008) and was found ranging from 0.530 (SE→NH) to 2.7003 (MG→SE) (Table 9, Figure 15). For migration direction, MG population moved to NH and SE populations with high migration rate (*m*) of 2.5126 and 2.7003, respectively, whereas SE population directly moved to KB population (*m*

= 2.2660) that did not pass into NH. Migration rate of KB→MG was more than (*m* = 1.2616) KB→NH (*m* = 0.6057). Migration rate of NA→NH was more than (*m* = 2.1310) NA→SE (*m* = 0.5933). Migration rate of SE→NA was higher (*m* = 1.3255) than the opposite direction (NA→SE).

Table 9 MCMC estimates historical migration rate of 5 populations of whisker sheatfish using MIGRATE-n program.

	Loc. Ln(L)	Theta (4Nμ)	M (m/mμ) of receiving population				
			NH	KB	MG	SE	NA
NH	-1115.859	0.35399	-------	0.6608	1.1692	0.8369	1.1347
KB	-1115.859	0.39254	0.6057	-------	1.2616	0.9998	0.9648
MG	-1115.859	0.45185	2.5126	1.5841	-------	2.7003	1.3211
SE	-1115.859	0.66294	0.5304	2.2660	1.4299	-------	1.3255
NA	-1115.859	0.48598	2.1310	0.6945	1.2533	0.5933	-------

Whisker sheatfish is a horizontal migratory fish which is classified as potamodromous; fish migrate within freshwater only (Eric, 2006). In the past, drainage structure was not constructed. Whisker sheatfish could move along between main river and tributaries or even swamp. Bao *et al*. (2001) suggested that fishes migrate for food from place to place and from season to season and fishes must also migrate to lay their eggs. In this study, whisker sheatfish migrated between feeding habitat, the Mekong river and Songkhram river basin, for food on a daily basis, and breeding/spawning habitat such as the Nong-Han Lake and Nam Kum basin. Therefore, high migration rate was found between main river and their tributaries or swamp such as MG→SE, MG→KB, MG→NH and NA→NH.

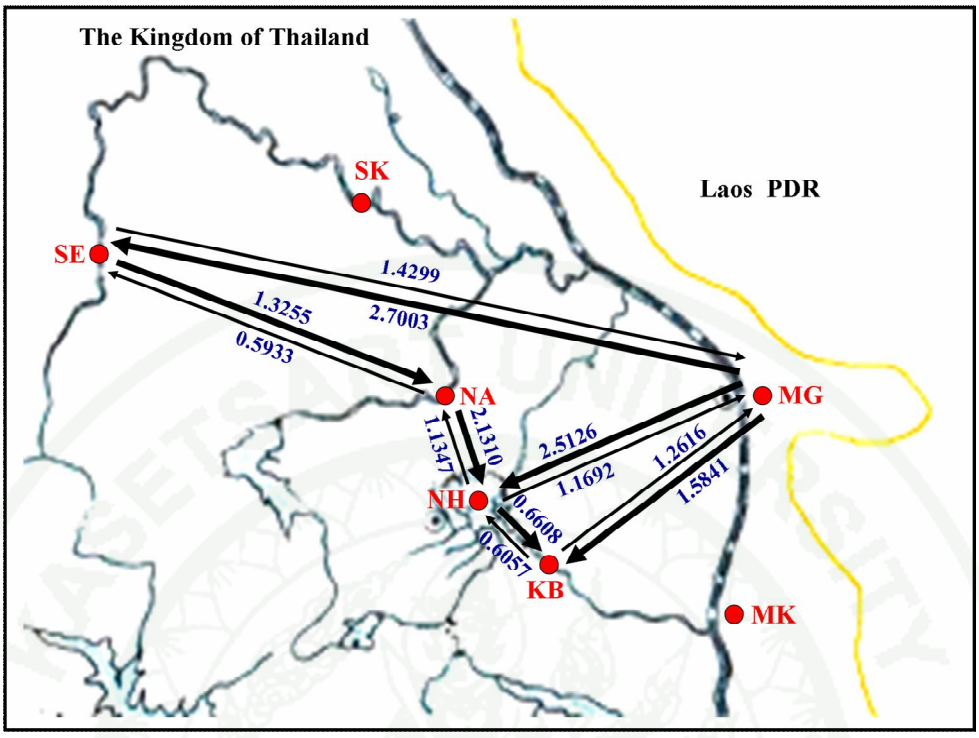


Figure 15 Estimates of the historical migration rate (numbers on arrow) and migration directions on five populations of whisker sheatfish implemented in MIGRATE-n. The bold arrows are the main direction of whisker sheatfish movement.

The recent migration rate was estimated using BAYESASS (Wilson and Rannala, 2003) and was found ranging from 0.0025 (SE→NA) to 0.0129 (MG→SE) (Table 10, Figure 16). The migratory direction of whisker sheatfish changed at three connecting sites (the green arrows in Figure 16). Firstly, the direction of whisker sheatfish movement changed from NA→NH (m = 0.0030) to NA→SE (m = 0.0056). Secondly, the direction changed from MG→KB (m = 0.0041) to KB→MG (m = 0.0108). Finally, it changed from MG→NH (m = 0.0033) to NH→MG (m = 0.0066). Moreover, the recent migration rate of all whisker sheatfish populations was extremely reduced. The rate of NA→NH was reduced 700 times of historical migration rate. In addition, the rate of MG→KB was reduced 386 times of historical migration rate whereas the rate of MG→NH was reduced 761 times of historical migration rate.

Table 10 MCMC estimates the recent migration rate of 5 populations of whisker sheatfish using BAYESASS.

	Recent migration rate of receiving population				
	NH	KB	MG	SE	NA
NH	-------	0.0032	0.0066	0.0043	0.0026
KB	0.0030	-------	0.0108	0.0056	0.0032
MG	0.0033	0.0041	-------	0.0129	0.0026
SE	0.0030	0.0042	0.0058	-------	0.0025
NA	0.0030	0.0036	0.0051	0.0056	-------

Our results corresponded with those reported by Morita and Yamamoto (2002). They found that isolation due to physical barriers is expected to reduce within-population genetic diversity and increase genetic differentiation, and migration rate has been changed in the white-spotted charr *Salvelinus leucomaenis* (Pallas) isolated upstream of dams for 30 years. In addition, Hanfling and Weetman (2006) studied in the sedentary river fish *Cottus gobio* and suggested that the construction blocks the migratory of fish passage from feeding ground to native spawning grounds.

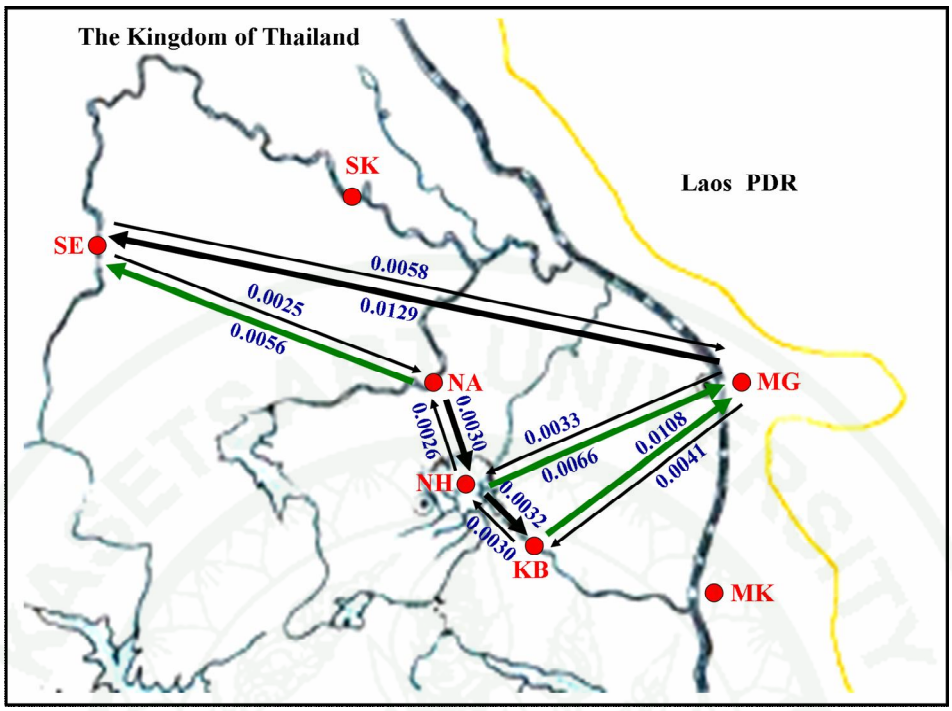


Figure 16 Estimates of the recent migration rate (numbers on arrow) and migration directions on five populations of whisker sheatfish implemented in BAYESASS. The bold arrows are the main direction of whisker sheatfish movement.

From the results obtained, the construction of barrier at the collecting sites degraded the habitat of whisker sheatfish, leading to change the migration pattern and reduce the migration rate between populations. Additionally, the reduction of migration rate provided high degree of population differentiation, affecting inbreeding within population and the reduction in population size.

3.6 Effective population size and population bottleneck

3.6.1 Effective population size

In this study, the maximum likelihood-based coalescent Markov Chain Monte Carlo (MCMC) approach inferred that the mutation scaled population size parameter ($\theta = 4N_e\mu$) were highest in SE population ($\theta = 0.662$) but lowest in NH

population (θ = 0.353) (Table 11). The mutation rate across loci was 2.2-2.7×10^{-4} per generation. These mutation rate values were used to predict the historical migration rate and translate the estimates to the historical effective population size (N_e). The historical N_e of 5 populations were 317, 398, 470, 690 and 407 for NH, KB, MG, SE and NA, respectively.

In this experiment, the θ-estimates from migrate was directly related to effective population size (N_e) whereas the mutation rates for the microsatellites used in this study was estimated from MSVAR 1.3 (Beaumont, 2004) for all studied populations. The θ-estimate of the Songkhram river basin was larger than the estimates of the remaining populations. It implied that population size of SE was large enough to maintain the long-term genetic diversity.

Table 11 Population size parameter (θ), mutation rate (μ), historical effective population size (N_e), recently effective population size (N_e in parenthesis) and 95% confidence interval from ONeSAMP of 5 whisker sheatfish populations.

Populations	θ	μ	N_e	(N_e)	95%confidence interval
NH	0.353	2.7x10^{-4}	317	(36)	24.9-78.4
KB	0.392	2.4x10^{-4}	398	(92)	57.0-246.4
MG	0.451	2.4x10^{-4}	470	(377)	162.4-1335.0
SE	0.662	2.4x10^{-4}	690	(92)	55.8-233.8
NA	0.485	2.2x10^{-4}	407	(39)	28.5-77.5

The estimated recent effective population sizes using the approximate Bayesian computation method were 36, 92, 377, 92 and 39 for NH, KB, MG, SE and NA, respectively (Table 11). The Mekong river carried higher recent effective population size than the remaining populations. However, these values were lower than 500 individuals per population.

The significant reduction in population size of the whisker sheatfish was observed based on the effective population size estimates. Generally, a population of N_e = 500 is considered to be large enough to maintain genetic diversity for key life-history traits, hence genetic diversity would appear depleted only if N_e <500 (Frankham, 1995). In addition, Franklin (1980) revealed that N_e > 50 will prevent inbreeding depression, whereas a detectable value of which N_e < 50 will decrease in viability or reproductive fitness of a population. Furthermore, Horreo *et al*. (2011) suggested that small effective population size has some influence on the reduction of allelic diversity. In this study, low allelic diversity was found for all populations. Therefore, the whisker sheatfish populations were thus currently proned to population extinction in a short period of time, especially in small and isolated populations such as the Nong-Han Lake (NH) and Lam Nam Oon (NA) populations.

3.6.2 Population bottleneck

The observed *M* ratio of all whisker sheatfish populations was higher than the commonly used bottleneck threshold, ranging from 0.705 (NA) to 0.933 (MG) (Table 12). The *M* ratio across loci of whisker sheatfish was 0.782. Values of $M \leq 0.68$ are indicative of a recent reduction in size that suggests bottleneck effect (Garza and Williamson, 2001). However, even population bottleneck signature was not observed in all whisker sheatfish, but the loss of allele size range was found in this study. One allele was not found in MB79 (305 bp), MB618 (133 bp) and MB645 (179 bp). In addition, two alleles were not observed for MB613 (148 and 150 bp) and MB614 (251 and 253 bp). The results corresponded with those reported by Garza and Williamson (2001). They suggested that when a population was reduced in size, genetic drift was enhanced and alleles will eventually be lost.

Under the strict SMM and TPM (95%SMM and 5%IAM), the Wilcoxon's sign-rank test based on detection of heterozygosity excess (Luikart *et al*., 1998a) did not reveal recent bottleneck signal with non-significant value in the whisker sheatfish populations (Table 13). This indicated that the observed gene

diversity in all whisker sheatfish samples was lower than the expected equilibrium gene diversity.

Table 12 Population bottleneck assessed from three methods.

	Populations				
Tests	NH	KB	MG	SE	NA
M-ratio	0.730	0.782	0.933	0.763	0.705
Wilcoxon's					
SMM	0.843^{NS}	0.995^{NS}	0.999^{NS}	0.999^{NS}	0.994^{NS}
TPM	0.679^{NS}	0.995^{NS}	0.998^{NS}	0.999^{NS}	0.994^{NS}
Mode-shift	L-shape	L-shape	L-shape	L-shape	L-shape

^{NS} is non significant ($P>0.05$)

Additionally, the mode-shift test on the allele frequency distributions did not display the evidence of bottleneck (Table 12). The allele frequency distribution pattern as revealed by the mode shift indicator (Luikart *et al*., 1998b) was L-shaped, indicating a larger proportion of low frequency allele classes in whisker sheatfish.

Nonetheless, the test failed to detect any bottleneck in the population dynamics of the whisker sheatfish, despite the fact that a reduction in the allelic diversity was found. However, Hudertmark and van Daele (2010) and Cornuet and Luikart (1996) reported that bottlenecks could be a precursor to reductions in the allelic diversity. In addition, Cornuet and Luikart (1996) suggested that populations and species having undergone severe population reduction or bottlenecks are more likely to become extinct. In this study, the *M* ratio of isolated population, NH (0.730) and NA (0.705), was close to bottleneck threshold, $M \leq 0.68$ (Garza and Williamson, 2001). Many loci showed monomorphism at MB354, MB614 and MB618 for NH

population. In addition, MB320, MB354 and MB614 exhibited monomorphism for NA population. Spear *et al*. (2006) suggested that the *M* ratio may be more sensitive to detect population bottlenecks than other two methods because this parameter detected size-reduced in the population. In this study, loss of allele size ranges were found in MB81, MB401, MB456, MB515, MB613, MB618 and MB645 for NH population whereas NA population exhibited the loss of allele size range at MB81, MB401, MB456, MB515, MB613, MB618 and MB645 loci. Furthermore, the effective population size of both NH and NA populations was lower than 50. These implied that random genetic drift influenced the reduction in population size and caused the loss of allelic diversity in NH and NA populations. In the past, Lam Nam Oon was connected with the Nong-Han Lake, until Nam Oon dam was constructed. The dam blocked and separated these two regions. Whisker sheatfish was lost from Nam Oon in a short time period, and also whisker sheatfish populations were fragmented to small population size and conducted random genetic drift to affect NH and NA populations.

In the wild populations, homozygote excess could be explained by a small population bottleneck (dramatic decrease of natural population). With the evidences of small population size, the bottleneck signatures can be expected (Islam *et al*., 2007). As the whisker sheatfish is considered a rare but exploited species, the reduction in population size is thus highly possible. In general, low genetic diversity such as a small number of alleles, can be caused by population processes, such as historical bottlenecks and low effective population size (Nei *et al*., 1975).

Due to destruction and fragmentation of whisker sheatfish habitats, these fish have been forced into small and isolated populations. They also faced tremendous risk from the effects of environmental variation, demographic stochasticity and reduced genetic diversity. Whisker sheatfish at these areas are facing the reduction in population size, low rate of gene flow among populations, causing by geographical barrier. It is, therefore, a great concern to maintain genetic variation within and between populations to prevent genetic erosion for this valuable species.

CONCLUSION AND RECOMMENDATION

1. Microsatellite sequences of whisker sheathfish were isolated using the enrichment strategy with four microsatellite probes, $(CA)_{15}$, $(GA)_{15}$, $(ACC)_{10}$, $(CCT)_{10}$. Three groups of microsatellite sequences were obtained. These included perfect repeat sequences, $(CA)_n$, $(GT)_n$, $(CCT)_n$, compound repeat sequences, $(CT)_n(GA)_n$) as an example and interrupted repeat sequences $(A)_nN(CT)_n$, $(GA)_nN(GA)_n$, $(GA)_nN(GGA)_n$, $(GGA)_nN$ $(GGA)_n$ where N is the bases interrupting motif sequences. Di-nucleotide repeat, GA/CT, was commonly microsatellite repeat motif found in whisker sheatfish. The efficiency for isolation microsatellites of whisker sheatfish was not different from the previous reports on other freshwater fish species. Forty-two primer pairs were designed and thirteen microsatellite markers were polymorphic. All microsatellite markers possessed more than 12 repeats of the motif. Eleven polymorphic markers were selected and used for the detection of genetic variability in whisker sheatfish populations.

2. Average allelic richness was 3.29 whereas observed heterozygosity was 0.279 and expected heterozygosity was 0.357. Whisker sheatfish populations exhibited low genetic diversity, comparing with other freshwater fish species. Therefore, these populations were vulnerable to genetic erosion and consequently could vanish from these regions.

3. Genetic diversity among populations was very high with F_{ST} value of 0.26 due to the reduction in migration rate and geographical barrier. Conservation management should operate to protect this valuable species.

LITERATURE CITED

Agresti, J.J., S. Seki, A. Cnaani, S. Poompuang, E.M. Hallerman, N. Umiel, G. Hulata, G.A.E. Gall and B. May. 2000. Breeding new strains of tilapia: development of an artificial center of origin and linkage map based on AFLP and microsatellite loci. **Aquaculture** 185: 43-56.

Allendorf, F., N. Ryman and F. Utter. 1987. Genetics and fishery management: past, present and future, pp. 1-20. *In* N. Ryman and F.M. Utter, eds. **Population Genetics and Fishery Management**. University of Washington, Washington.

Allendorf, F.W. and G. Luikart. 2007. **Conservation and the Genetics of Populations**. Blackwell Publishing Ltd., Oxford, UK.

Altukhov, Y.P. and E.A. Salmenkova. 2002. DNA polymorphism in population genetics. **Russ. J. Genet.** 38: 989-1008.

Angel, P., G. Mercedes, L. Philippe, M. Concepcion and A. Jose, 2006. Effects of fishing protection on the genetic structure of fish populations. **Biol.Cons.** 129: 244-255.

Arnold, G. 1999. Unravelling mechanisms of fish migration: a 30-year perspective, pp. 7-21. *In* D.A. Hancock, D.C. Smith and J.D. Koehn, eds. **Fish Movement and Migration.** Bendigo: Fisheries Research and Development Corporation.

Avise, J.C. 1994. **Molecular Markers**. Natural History and Evolution. Chapman and Hall, NY.

Avise, J.C., A.J. Power and D. Walker. 2004. Genetic sex determination, gender identification, and pseudohermaphroditism in the knobbed whelk, Busycon carica (Mollusca; Melongenidae). **Proc. R. Soc. Lond. B.** 271: 641-646.

Balloux, F. and N.L. Moulin. 2001. The estimation of population differentiation with microsatellite markers. **Mol. Ecol.** 11: 155-165.

Bao, T.Q., K. Bouakhamvongsa, S. Chan, K.C. Chhuon, T. Phommavong, A.F. Poulsen, P. Rukawoma, U. Suornratana, D.V. Tien, T.T. Tuan, N.T. Tung, J. Valbo-Jorgensen, S. Viravong and N. Yoorong. 2001. **Local Knowledge in the Study of River Fish Biology: Experiences from the Mekong, Mekong Development Series No. 1**. Mekong River Commission, Phnom Penh, Cambodia.

Barker, J.S.F. 1994. A global protocol for determining genetic distances among domestic livestock breeds, pp. 501-208. *In* **Proceedings of the 5th world congress on genetics applied to livestock production**. Guelph, Ontario, Canada.

Bartfai, R., S. Egedi, G.H. Yue, B. Kovacs, B. Urbanyi, G. Tamas, L. Horvath and L. Orban. 2003. Genetic analysis of two common carp broodstocks by RAPD and microsatellite markers. **Aquaculture** 219: 157-167.

Barman, H.K., A. Barat, B.M. Yadav, S. Banerjee, P.K. Meher, P.V.G.K. Reddy and R.K. Jana. 2003. Genetic variation between four species of Indian major carps as revealed by random amplified polymorphic DNA assay. **Aquaculture** 217: 115-123.

Beaumont, M.A. 2004. **msvar1.3 update**. Available Source: http://www.rubic.rdg. ac.uk/~mab/stuff., August 28, 2010.

Beerli, P. 2008. **MIGRATE-n a maximum likelihood program to estimate gene flow using the coalescent.** Available Source: http://pop gen.scs.fsu.edu/ Migrate-n.html., December 17, 2010.

Bhassu, S., K. Yusoff, J.M. Panandam, W.K. Embong, S. Oyyan and S.G. Tan. 2004. The genetic structure of *Oreochromis* spp. (Tilapia) populations in Malaysia as revealed by microsatellite DNA analysis. **Biochem. Genet.** 42: 217-229.

Brown, B. and J. Epifanio. 2003. Nuclear DNA, p. 458. *In* E.M. Hallermann. **Population Genetics: Principles and Applications for Fisheries Scientists**. American Fisheries Society, Bethesda, MD.

Callen, D.F., A.D. Thompson, Y. Shen, H.A. Phillips, R.I. Richards, J.C. Mulley and G.R. Sutherland. 1993. Incidence and origin of null alleles in the $(AC)_n$ microsatellite markers. **Am. J. Hum. Genet.** 52: 922-927.

Campos, J.L., D. Posada and P. Moran. 2006. Genetic variation at MHC, mitochondrial and microsatellite loci in isolated populations of Brown trout (*Salmo trutta*). **Conserv. Genet.** 7: 515-530.

Carvalho, G.R. and L. Hauser. 1995. Molecular genetics and the stock concept in fisheries, pp. 55-80. *In* G.R. Carvalho and T.J. Pitcher, eds. **Molecular Genetics in Fisheries**. Chapman and Hall, London.

Castric, V., L. Bernatchez, K. Belkhir and F. Bonhomme. 2002. Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus Fontinalis* Mitchill (Pisces, Salmonidae): a test of alternative hypotheses. **Heredity** 89: 27-35.

Cegelski, C.C., L.P. Waits and N.J. Anderson. 2003. Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. **Mol. Ecol.** 12: 2907-2918.

Chakraborty, R. and O. Leimar. 1987. Genetic variation within a subdivided population, pp. 89-120. *In* N. Ryman and F.M. Utter, eds. **Population Genetics and Fishery Management**. University of Washington, Washington.

Chakraborty, R., M. De Andrade, S.P. Daiger and B. Budowle. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. **Ann. Hum. Genet.** 56: 45-57.

Chapuis, M.P. and A. Estoup. 2007. Microsatellite null alleles and estimation of population differentiation. **Mol. Biol. Evol.** 24: 621-631.

Chen, N., H.L. Wang, W.M. Wang and V. Simonsen. 2010. Isolation of polymorphic microsatellite loci from an endangered freshwater species Chinese sucker, *Myxocyprinus asiaticus*. **Conserv. Genet. Resour.** 2: 73-75.

Chistiakov, D.A., B. Hellemans and F.A.M. Volckaert. 2006. Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics**. Aquaculture** 255: 1-29.

Chu, S.L., C.F. Weng, C.D. Hsiao, P.P. Hwang, Y.C. Chen, J.M. Ho and S.J. Lee. 2006. Profile analysis of expressed sequence tags derived from the ovary of tilapia, *Oreochromis mossambicus.* **Aquaculture** 251: 537-548.

Coates, D., O. Poeu, U. Suntornratana, N.T. Tung and S. Viravong. 2003. **Biodiversity and Fisheries in the Mekong River Basin, Mekong Development Series 2.** Mekong River Commission, Phnom Penh, Cambodia.

Corander, J. and P. Marttinen. 2006. Bayesian identification of admixture events using multi-locus molecular markers. **Mol. Ecol.** 15: 2833-2843.

Cornuet, J.M. and G. Luikart. 1996. Description and evaluation of two tests for detecting recent bottlenecks. **Genetics** 144: 2001-2014.

Crow, J.F. and M. Kimura. 1970. **An Introduction to Population Genetics Theory**. Harper and Row Publishers, New York, Evanston and London, UK.

Cui, J.Z., X.Y. Shen, G.P. Yang, Q.L. Gong and Q.Q. Gu. 2005. Characterization of microsatellite DNAs in *Takifugu rubripes* genome and their utilization in the genetic diversity analysis of *T. rubripes* and *T. pseudommus*. **Aquaculture** 250: 129-137.

Dakin, E.E. and J.C. Avise. 2004. Microsatellite null allele in parentage analysis. **Heredity** 93: 504-509.

DeWoody, J.A. and J.C. Avise. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. **J. Fish Biol.** 56: 461-473.

Dinesh, K.R., W.K. Chan, T.M. Lim and V.P.E. Phang. 1995. RAPD markers in fishes: an evaluation of resolution and reproducibility. **Asia-Pac. J. Mol. Biol. Biotechnol**. 3: 112-118.

Duncan, J.R. and J.L. Lockwood. 2001. Extinction in a field of bullets, a search for causes in the decline of the world's freshwater fishes. **Biol. Conserv.** 102**:** 97-105.

Eguia, M.R.R.R., M. Ikeda, Z.U. Basiao and N. Taniguchi. 2004. Genetic diversity in farmed Asian Nile and red hybrid tilapia stocks evaluated from microsatellite and mitochondrial DNA analysis. **Aquaculture** 236: 131-150.

Ellegren, H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. **Trends Genet**. 16: 551-558.

Ellegren, H. 2004. Microsatellites: Simple sequences with complex evolution. **Nat. Rev. Genet.** 5: 435-445.

Ellegren, H., C. R. Primmer, B. Sheldon. 1995. Microsatellite evolution: directionality or bias in locus selection? **Nat. Genet**. 11: 60-62.

Eric, B. 2006. **Fish Migration Triggers in the Lower Mekong Basin and Other Tropical Freshwater Systems, MRC Technical Paper No. 14.** Mekong River Commission, Vientiane.

Ferguson, M. 1995. The role of molecular genetic markers in the management of cultured fish, pp. 81-104. *In* G.R. Carvalho and T.J. Pitcher, eds. **Molecular Genetics in Fisheries**. Chapman and Hall, London.

Fishery Information Technology Center. 2006. **Statistic of the Freshwater Fisheries Product from Natural Reservoir in 2004 No. 12.** Department of Fisheries, Ministry of Agriculture and Cooperative, Bangkok, Thailand. (Thai article)

Foulley, J.L. and L. Ollivier. 2006. Estimating allelic richness and its diversity. **Livest. Sci.** 101: 150-158.

Frankel, O.H. and M.E. Soule. 1981. **Conservation and Evolution**. Cambridge University Press, NY.

Frankham, R. 1995. Conservation genetics. **Ann. Rev. Genet.** 29: 305-327.

Frankham, R., J.D. Ballou and D.A. Briscoe. 2002. **Introduction to Conservation Genetics**. Cambridge University Press, Cambridge, UK.

Frankham, R., J.D. Ballou and D.A. Briscoe. 2009. **Introduction to Conservation Genetics**. Cambridge University Press, NY.

Franklin, I.R. 1980. Evolutionary change in small populations, pp. 135-150. *In* M.E. Soule and B.A. Wilcox, eds. **Conservation Biology: An Evolutionary-Ecological Perspective**. Sinauer, Sunderland, MA.

Gall, G.A.E. 1987. Inbreeding, pp. 47-87. *In* N. Ryman and F. Utter, eds. **Population Genetics and Fishery Management**. University Washington Press, Seattle, WA.

Garza, J.C. and E.G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. **Mol. Ecol.** 10: 305-318.

Goldstein, D.B. and D.D. Pollock. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. **J. Hered.** 88: 335-342.

Goudet, J. 2001. **FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3)**. Available Source: http://www2.unil.ch popgen/softwares/ fstat.htm, January 9, 2012.

Gu, Z.X., T.J. Gou and X.B. Xi. 2006. Application of microsatellite markers in studies of genetics and breeding of fish. **Chinese J. Agr. Biotechnol.** 3: 83-87.

Guo, S. and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. **Biometrics** 48: 361-372.

Ha, H.P.,T.T.T. Nguyen, S. Poompuang and U. Na-Nakorn. 2009. Microsatellites revealed no genetic differentiation between hatchery and contemporary wild populations of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878) in Vietnam. **Aquaculture** 291: 154-160.

Halkett, F., J.C. Simon and F. Balloux. 2005. Tackling the population genetics of clonal and partially clonal organisms. **Trends Ecol. Evol.** 20: 194-201.

Hamilton, M.B., E.L. Pincus, A. Di-Fiore and R.C. Fleischer. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. **Biotechniques** 27: 500-507.

Hanfling, B. and D. Weetman. 2006. Concordant genetic estimators of migration reveal anthropogenically enhanced source-sink population structure in the river sculpin, *Cottus gobio*. **Genetics** 173: 1487-1501.

Hashimoto, D.T., F.F. Mendonca, J.A. Senhorini, J. Bortolozzi, C. de Oliveira, F. Foresti and F. Porto-Foresti. 2010. Identification of hybrids between Neotropical fish *Leporinus macrocephalus* and *Leporinus elongatus* by PCR-RFLP and multiplex-PCR: tools for genetic monitoring in aquaculture. **Aquaculture** 298: 346-349.

Hatanaka, T., F. Henrique-Silva and P.M. Galetti JR. 2006. Population structuring in a migratory freshwater fish *Prochilodus argenteus* (Characiformes, Prochilodontidae) from São Francisco River. **Genetica** 126: 153-159.

Hendry, A.P. 2001. Adaptive divergence and the evolution of reproductive isolation in the wild: an empirical demonstration using introduced sockeye salmon. **Genetica** 112: 515-534.

Hill, M.T. and S.A. Hill. 1994. **Fisheries Ecology and Hydropower in the Lower Mekong River: an Evaluation of Run-of-the-River Projects**. Mekong Secretariat, Bangkok, Thailand.

Hill, W.G. 1981. Estimation of effective population size from data on linkage disequilibrium. **Genet. Res.** 38: 209-216.

Hoarau, A., A.D. Rijnsdorp, H.W. Vander Veer, W.T. Stam and J.L. Olsen. 2002. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: microsatellites revealed large-scale spatial and temporal homogeneity. **Mol. Ecol.** 11: 1165-1176.

Hogan, Z.S. and B.P. May. 2002. Twenty-seven new microsatellites for the migratory Asian catfish family Pangasiidae. **Mol. Ecol. Notes** 2: 38-41.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. **Scand. J. Stat.** 6: 65-70.

Horreo, J.L., J.L. Martinez, F. Ayllon, I.G. Pola, J.A. Monteoliva, M.H. Land and E. Garcia-Vazquez. 2011. Impact of habitat fragmentation on the genetics of populations in dendritic landscapes. **Freshwater Biol.** 56: 2567-2579.

Huang, X. and A. Madan. 1999. CAP3: A DNA sequence assembly program. **Genome Res**. **9**: 868-877.

Hubert, S., J.T. Bussey, B. Higgins, B.A. Curtis and S. Bowman. 2009. Development of single nucleotide polymorphism markers for Atlantic cod (*Gadus morhua*) using expressed sequences. **Aquaculture** 296: 7-14.

Hundertmark, K.J. and L.J. van Daele. 2010. Founder effect and bottleneck signatures in an introduced, insular population of elk. **Conserv. Genet.** 11: 139-147.

Islam, M.N., M.S. Islam and M.S. Alam. 2007. Genetic structure of different populations of walking catfish (*Clarias batrachus* L.) in Bangladesh. **Biochem. Genetics** 45: 647-662.

Jondeung, A., P. Sangthong and R. Zardoya. 2007. The complete mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*), and the phylogenetic relationships among Siluriformes. **Gene** 387: 49-57.

Karnsomdee, P. and W. Meckvichai. 2002. Genetic variation of Thai native fowls *Gallus gallus domesticus* based on microsatellite DNA, pp 126. **Proceeding of the 14th annual meeting of Thai Society for "Biotechnology for Better Living in the New Economy"**. Khonkaen Province, Thailand.

Koyuk, A., A. Bennett and D. Tallmon. 2008. **ONeSAMP 1.1**. Avialable Source: http://genomics.jun.alaska.edu, January 20, 2011.

Leberg, P.L. 2002. Estimating allelic richness: effects of sample size and bottlenecks **Mol. Ecol.** 11: 2445-2449.

Lehmann, T., W.A. Hawley and F.H. Collins. 1996. An evaluation of evolutionary constraints on microsatellite loci using null alleles. **Genetics** 144: 1155-1163.

Lenormand, T. 2002. Gene flow and the limits to natural selection. **Trends Ecol. Evol.** 17: 183-189.

Liao, M., L. Zhang, G. Yang, M. Zhu, D. Wang, Q. Wei, G. Zou and D. Chen. 2007. Development of silver carp (*Hypophthalmichthys molitrix*) and bighead carp (*Aristichthys nobilis*) genetic maps using microsatellite and AFLP markers and a pseudo-testcross strategy. **Anim. Genet.** 38: 364-370.

Liu, Z.J., P. Li, B.J. Argue and R.A. Dunham. 1999. Random amplified polymorphic DNA markers: usefulness for gene mapping and analysis of genetic variation of catfish. **Aquaculture** 174: 59-68.

Liu, Z.J. and J.F. Cordes. 2004. DNA marker technologies and their applications in aquaculture genetics. **Aquaculture** 238: 1-37.

Lucentini, L., A. Palomba, H. Lancioni, L. Gigliarelli, M. Natali and F. Panara. 2006. Microsatellite polymorphism in Italian populations of northern pike (*Esox lucius* L.). **Fisheries Res**. 80: 251-262.

Luikart, G. and J.M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. **Conserv. Biol.** 12: 228-237.

Luikart, G., F.W. Allendorf, J.M. Cornuet and W.B. Sherwin. 1998a. Distortion of allele frequency distributions provides a test for recent population bottlenecks. **J. Hered.** 89: 238-247.

Luikart, G., W.B. Sherwin, B.M. Steele and F.W. Allendorf. 1998b. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. **Mol. Ecol.** 7: 963-974.

Makinen, H.S., J.M. Cano and J. Meril. 2006. Genetic relationships among marine and freshwater populations of the European three-spined stickleback (*Gasterosteus aculeatus*) revealed by microsatellites. **Mol. Ecol.** 15: 1519-1534.

May, B. and C.C. Krueger. 1990. Use of allozyme data for population analysis, pp. 157-171. *In* D.H. Whitmore. **Electrophoretic and Isoelectric Focusing Techniques in Fisheries Management**. CRC Press, Boston.

Meffe, G. K. and R.C. Vrijenhoek. 1988. Conservation genetics in the management of desert fishes. **Conserv. Biol.** 2: 157-169.

Mickett, K., C. Morton, J. Feng, P. Li, M. Simmons, D. Cao, R.A. Dunham and Z. Liu. 2003. Assessing genetic diversity of domestic populations of channel catfish (*Ictalurus punctatus*) in Alabama using AFLP markers. **Aquaculture** 288: 91-105.

Moeser, A.A. and E. Bermingham. 2005. Isolation and characterization of eight microsatellite loci for the Neotropical freshwater catfish *Pimelodella chagresi* (Teleostei: Pimelodidae). **Mol. Ecol. Notes** 5: 363–365.

Morin, P.A., P. Mahboubi, S. Wedel and J. Rogers. 1998. Rapid screening and comparison of human microsatellite markers in baboons: allele size is conserved, but allele number is not. **Genomics** 53**:** 12-20.

Morita, K. and S. Yamamoto. 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. **Conserv. Biol.** 16: 1318-1323.

Mullis, K.B. and F.A. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. **Met. Enzymol.** 155: 335-350.

Murphy, R.W., J.W. Sites, D.G. Buth and C.H. Haufler. 1996. Proteins: Isozyme Electrophoresis, pp. 51-120. *In* D.M. Hillis, C. Moritz and B.K. Mable, eds. **Molecular Systematics**, 2nd edition. Sinauer Associates, Massachusetts.

Murray, B.W. 1996. **The estimation of genetic distance and population substructure from microsatellite allele frequency data**. Available Source: http://helix. mcmaster.ca/brent/brent.html, January 28, 2011.

Na-Mahasarakarm, O.P. 2007. **An Introduction to the Mekong Fisheries of Thailand, Mekong Development Series 5**. Mekong River Commission, Vientiane, Lao PDR.

Na-Nakorn, U., K. Sriphairoj, S. Sukmanomon, S. Poompuang and W. Kamonrat. 2006. Polymorphic microsatellite primers developed from DNA of the endangered Mekong giant catfish, *Pangasianodon gigas* (Chevey) and cross-species amplification in three species of Pangasius. **Mol. Ecol. Notes** 6: 1174-1176.

Na-Nakorn, U. and T. Moeikum. 2009. Genetic diversity of domesticated stocks of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878), in Thailand: relevance to broodstock management regimes. **Aquaculture** 297: 70-77.

Na-Nakorn, U., W. Kamonrat and T. Ngamsiri. 2004. Genetic diversity of walking catfish, *Clarias macrocephalus*, in Thailand and evidence of genetic introgression from introduced farmed *C. gariepinus* genome. **Aquaculture** 240: 145-163.

Nei, M., T. Maruyama and R. Chakraborty. 1975. The bottleneck effect and genetic variability in populations. **Evolution** 29: 1-10.

Neigel, J.E. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. **Annu. Rev. Ecol. Syst.** 28: 105-128.

Ngamsiri, T., M. Nakajima, S. Sukmanomol, N. Sukumasavin, W. Kamonrat, U. Na-Nakorn and N. Taniguchi. 2007. Genetic diversity of wild Mekong giant catfish *Pangasianodon gigas* collected from Thailand and Cambodia. **Fisheries Sci**. 73: 792-799.

Ngamsiri, T., Y.N. Ohashi, S. Sukmanomon, M. Nakajima, U. Na-Nakorn and N. Taniguchi. 2006. Characterization of microsatellite DNA markers in a critically endangered species, Mekong giant catfish, *Pangasianodon gigas*. **Mol. Ecol. Notes 6**: 313-315.

Nielsen, M.D. and B. Storm. 1984. Simulation of water resources development projects in the Nam Kam Basin, Thailand. **Nord. Hydrol.** 15: 297-306.

Nielsen, R. 2000. Estimation of population parameters and recombination rates from single nucleotide polymorphisms. **Genetics** 154: 931-942.

O'Brien, S.J. 1991. Molecular genome mapping: lessons and prospects. **Curr. Opin. Genet. Dev.** 1: 105-111.

O'Connell, M. and J.M. Wright. 1997. Microsatellite DNA in fish. **Rev. Fish Biol. Fisher.** 7: 331-363.

Ohashi, Y., M. Nakajima, N. Sukumasavin, U. Na-Nakorn and N. Taniguchi. 2006. Isolation and characterization of microsatellite DNA markers in endangered M Mekong giant catfish *Pangasianodon gigas.* **Fisheries Sci.** 72: 1066-1071.

Okumus, I. and Y. Ciftci. 2003. Fish population genetics and molecular markers: II-molecular markers and their applications in fisheries and aquaculture. **Turk. J. Fisheries Aqua. Sci.** 3: 51-79.

Paetkau, D. and C. Strobeck. 1995. The molecular-basis and evolutionary history of a microsatellite null allele in bears. **Mol. Ecol.** 4: 519-520.

Palumbi, S.R. 1994. Nucleic Acids II: The polymerase chain reaction, pp. 205-247. *In* D.M. Hillis, C. Moritz and B.K. Mable, eds. **Molecular Systematics**. Sinauer Associates Inc., Sunderland, MA.

Parra-Bracamonte, G.M., A.M. Sifuentes-Rincón, X.F. De la Rosa-Reyna, W. Arellano-Vera and B. Sosa-Reyes. 2011. Inbreeding evidence in a traditional channel catfish (*Ictalurus punctatus*) hatchery in Mexico. **Electro. J. Biotechnol.** 14: DOI http://dx.doi.org/10.2225/vol14-issue6-fulltext-7.

Partis, L. and R.J. Wells. 1996. Identification of fish species using random amplified polymorphic DNA (RAPD). **Mol. Cell. Probes** 10: 435-441.

Peel, D., J.R. Ovenden and S.L. Peel. 2004. NEESTIMATOR: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries. Available Source: http://www2.dpi.qld.gov.au/fishweb/13887.html, January 28, 2011.

Penzes, Z.S., G.Y. Csanadi, G.M. Kovacs and Z.S. Beer. 2002. Molecular marker in ecology. **Tiscia** 33: 9-30.

Pereira, L.H.G., F. Foresti and C. Oliveira. 2009. Genetic structure of the migratory catfish *Pseudoplatystoma corruscans* (Siluriformes: Pimelodidae) suggests homing behaviour. **Ecol. Freshw. Fish** 18: 215-225.

Piry, S., G. Luikart and J. Cornuet. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. **J. Hered.** 90: 502-503.

Poompuang, S. and U. Na-Nakorn. 2004. A preliminary genetic map of walking catfish (*Clarias macrocephalus*). **Aquaculture** 232: 195-203.

Poulsen, A.F., K.G. Hortle, J. Valbo-Jorgensen, S. Chan, C.K. Chhuon, S. Viravong, K. Bouakhamvongsa, U. Suntornratana, N. Yoorong, T.T. Nguyen and B.Q. Tran. 2004. **Distribution and Ecology of Some Important Riverine Fish Species of the Mekong River Basin**. MRC Technical Paper No. 10, Mekong River Commission, Phnom Penh.

Pritchard, J.K., M. Stephen and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. **Genetics** 155: 945-959.

Pujolar, J.M., A.N. Lucarda, M. Simonato and T. Patarnello. 2011. Restricted gene flow at the micro- and macrogeographical scale in marble trout based on mtDNA and microsatellite polymorphism. **Front. Zool.** 8: 7. DOI:10.1186/ 1742-9994-8-7.

Quan, Y.C., X.W. Sun and L.Q. Liang. 2006. Genetic polymorphism of microsatellite DNA in two populations of Northern sheatfish (*Silurus soldatovi*). **Acta Genet. Sin.** 33: 908-916.

Quiniou, S.M.A., M. Wilson, E. Bengten, G.C. Waldbieser, L.W. Clem and N.W. Miller. 2005. MHC RFLP analyses in channel catfish full-sibling families: identification of the role of MHC molecules in spontaneous allogeneic cytotoxic responses. **Dev. Comp. Immunol**. 29: 457-467.

Rainboth, W.J. 1996. **FAO Species Identification Field Guide for Fishery Purposes: Fishes of the Cambodian Mekong**. FAO, Rome.

Raymond, M. and F. Rousset. 2003. Genepop 3.4. an updated version of Genepop V.1.2 (1995): population genetics software for exact tests and ecumenicism. **Heredity** 86: 248-249.

Rengmark, A.H., A. Slettan, O. Skaala, O. Lie and F. Lingaas. 2006. Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. **Aquaculture** 253: 229-237.

Rozen, S. and H.J. Skaletsky. 2000. Primer3 on the WWW for general users and biologist programmers, pp. 365-386. *In* S. Krawetz and S. Misener, eds. Bioinformatics **Methods and Protocols: Methods in Molecular Biology**. Humana Press, Totowa.

Rutten, M.J.M., H. Komen, R.M. Deerenberg, M. Siwek and H. Bovenhuis. 2004. Genetic characterization of four strains of Nile tilapia (*Oreochromis niloticus* L.) using microsatellite markers. **Anim. Genet.** 35: 93-97**.**

Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. **Science** 239: 487-491.

Sala, O.E., F.S. Chapin, J.J. Armesto, R. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L.F. Huenneke, R.B. Jackson, A. Kinzig, R. Leemams, D. Lodge, H.A. Mooney, M. Oesterheld, N.L. Poff, M.T. Sykes, B.H. Walker, M. Walker and D.H. Wall. 2000. Global biodiversity scenarios for the year 2100. **Science** 287: 1770-1774.

Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. **Molecular Cloning: A Laboratory Manual**, 2nd edition. Cold Spring Harbor Laboratory Press, NY.

Sanetra, M. and A. Meyer. 2005. Microsatellites from the burbot (*Lota lota*), a freshwater gadoid fish (Teleostei). **Mol. Ecol. Notes** 5: 390-392.

Savan, R. and M. Sakai. 2002. Analysis of expressed sequence tags (EST) obtained from common carp, *Cyprinus carpio* L., head kidney cells after stimulation by two mitogens, lipopolysaccharide and concanavalin-A. **Comp. Biochem. Phys. B.** 131: 71-82.

Schlotterer, C. and B. Harr. 2001. Microsatellite Instability. **Encyclopedia of Life Sciences**. Nature Publishing Group, pp. 1-4.

Simmons, M., K. Mickett, H. Kucuktas, P. Li, R. Dunham and Z. Liu. 2006. Comparison of domestic and wild channel catfish (*Ictalurus punctatus*) populations provides no evidence for genetic impact. **Aquaculture** 252: 133-146.

Slatkin, M. 1985. Gene flow in natural populations. **Ann. Rev. Ecol. Sys.** 16: 393-430.

Smulders, M.J.M., G. Bredemeijer, W. Rus-kortekaas, P. Arens and B. Vosman. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. **Theor. Appl. Genet.** 97: 264-272.

So, N., G.E. Maes and F.A.M. Volckaert. 2006. High genetic diversity in cryptic populations of the migratory sutchi catfish *Pangasianodon hypophthalmus* in the Mekong River. **Heredity** 96: 166-174.

Sokheng, C., C.K. Chhea, S. Viravong, K. Bouakhamvongsa, U. Suntornratana, N. Yoorong, N.T. Tung, T.Q. Bao, A.F. Poulsen and J.V. Jørgensen. 1999. **Fish Migrations and Spawning Habits in the Mekong Mainstream: a Survey Using Local Knowledge (basin-wide)**. Assessment of Mekong fisheries: Fish Migrations and Spawning and the Impact of Water Management Project (AMFC). AMFP Report 2/99. Vientiane, Lao PDR.

Spear, S.F., C.R. Peterson, M.D. Matocq and A. Storfer. 2006. Molecular evidence for historical and recent population size reductions of tiger salamanders *Ambystoma tigrinumi* in Yellowstone National Park. **Conserv. Genet.** 7: 605-611.

Sriphairoj, K., W. Kamonrat and U. Na-Nakorn. 2007. Genetic aspect in broodstock management of the critically endangered Mekong giant catfish, *Pangasianodon gigas* in Thailand. **Aquaculture** 264: 36-46.

Su, Y., R. Long, G. Chen, X. Wu, K. Xie and J. Wan. 2007. Genetic analysis of six endangered local duck populations in China based on microsatellite markers. **J. Genet. Genomics** 34: 1010-1018.

Sukmanomon, S. and S. Poompuang. 2003. Development of microsatellite primers for walking catfish (*Clarias macrocephalus*), pp. 41-48. *In* **Proceeding of the 41st Kasetsart University Conference (Fisheries).** Kasetsart University, Bangkok.

Sukumasavin, N., M. Nakajima and N. Taniguchi. 2004. Identification and characterization of microsatellite DNA markers developed in an endangered cyprinid of the Mekong River, the seven-line barb (*Probarbus jullieni* Sauvage, 1880). **Mol. Ecol. Notes** 4: 465-467.

Taki, Y. 1974. **Fishes of the Lao Mekong Basin**. United States Agency for International Development Mission to Laos Agriculture Division. 232 p.

Tallmon, D.A., A. Koyuk, G. Luikart and M.A. Beaumont. 2008. ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. **Mol. Ecol. Resour.** 8: 299-301.

Tallmon, D.A., G. Luikart and M.A. Beaumont. 2004. Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. **Genetics** 167: 977-988.

Techen, N., R.S. Arias, N.C. Glynn, Z. Pan, S.A. Khan and B.E. Scheffler. 2010. Optimized construction of microsatellite-enriched libraries. **Mol. Ecol. Resour**. 10: 508-515.

Thai, B. T., C. P. Burridge and C. M. Austin. 2007. Genetic diversity of common carp (*Cyprinus carpio* L.) in Vietnam using four microsatellite loci. **Aquaculture** 269: 174-186.

Ukkatawewat, S. 1984. **The Taxonomic Characters and Biology of Some Important Freshwater Fishes in Thailand.** Manuscript. National Inland Fisheries Institute, Department of Fisheries, Ministry of Agriculture, Bangkok, Thailand.

Utter, F., P. Aebersold and G. Winans. 1987. Interpreting genetic variation detected by electrophoresis, pp. 21-46. *In* N. Ryman and F.M. Utter, eds. **Population Genetics and Fishery Management**. University of Washington, Washington.

Vali, U., M. Brandstrom, M. Johansson and H. Ellegren. 2008. Insertion-deletion polymorphisms (indels) as genetic markers in natural populations. **BMC Genetics** 9: 8. DOI:10.1186/1471-2156-9-8.

van Herwerden, L., J. McIlwain, H. Al-Oufi, W. Al-Amry and A. Reyes. 2006. Development and application of microsatellite markers for *Scomberomorus commerson* (Perciformes; Teleostei) to a population genetic study of Arabian Peninsula stocks. **Fisheries Res.** 79: 258–266.

van Oosterhout, C., W.F. Hutchinson, D.P.M. Wills and P. Shipley. 2004. Micro-Checker: software for identifying and correcting genotyping error in microsatellite data. **Mol. Ecol. Notes** 4: 535-538.

van Treuren, R. 1998. Estimating null allele frequencies at a microsatellite locus in the oystercatcher (*Haematopus ostralegus*). **Mol. Ecol.** 7: 1413-1417.

Vidthayanon, C., J. Karnasuta and J. Nabhitabhata. 1997. **Diversity of Freshwater Fishes in Thailand**. Office of Environmental Policy and Planning, Bangkok.

Wade, M.J. 1988. Extinction and recolonization: their effects on the genetic differentiation of local population. **Evolution** 42: 995-1005.

Wahlund, S. 1928. Zusammensetzung von population und korrelationserscheinung vom standpunkt der vererbungslehre aus betrachtet. **Hereditas** 11: 65-106.

Waits, L.P., G. Luikart and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. **Mol. Ecol.** 10: 249-256.

Wang, D., X. Liao, L. Cheng, X. Yu and J. Tong. 2007. Development of novel EST-SSR markers in common carp by data mining from public EST sequences. **Aquaculture** 271: 558-574.

Wang, J. 2004. Sibship reconstruction from genetic data with typing errors. **Genetics** 166: 1963-1979.

Ward, R.D., M. Woodwark and D.O.F. Skibinski. 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. **J. Fish Biol.** 44: 213-232.

Weber, J.L. and P.E. May. 1990. Abundant class of human DNA polymorphisms with can be types using polymerase chain reaction. **Am. J. Hum. Genet.** 44: 388-396.

Weir, B. S. 1996. Intraspecific differentiation, pp. 385-406. *In* D. M. Hillis, C. Moritz and B. K. Mable, eds. **Molecular Systematics**, 2nd edition. Sinauer Associates, Massachusetts.

Weir, B.S. and C.C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. **Evolution** 38: 1358-1370.

Wilson, G.A. and B. Rannala. 2003. Bayesian inference of recent migration rates using multilocus genotypes. **Genetics** 163: 1177-1191.

Wright, J.M. 1993. **DNA Fingerprinting in Fishes**, pp. 58-91. *In* P.W. Hochachka and T. Mommsen, eds. **Biochemistry and Molecular Biology of Fishes**, Vol. 2. Elsevier, Amsterdam.

Wright, S. 1951. The genetical structure of populations. **Ann. Eugenics** 15: 323-354.

Wright, S. 1969. **Evolution and the Genetics of Populations**. Vol. 2. University of Chicago Press, Chicago.

Wu, X.C. 2005. The loss of genetic diversity in the Chinese paddlefish (*Psephurus gladius* Martens) as revealed by DNA fingerprinting. **J. Genet.** 84: 323-327.

Yang, C., X. Zhu and X. Sun. 2008. Development of microsatellite markers and their utilization in genetic diversity analysis of cultivated and wild populations of the mud carp (*Cirrhina molitorella*). **J. Genet. Genomics** 35: 201-206.

Yang, R. 1998. Estimating hierarchical *F*-statistics. **Evolution** 52: 950-956.

Yeh, F.C. and R. Yang. 1999. **Popgene Version 1.31 Microsoft Window Based Freeware for Population Genetic Analysis**. University of Alberta and Tim Boyle, Centre for International Forestry Research, Canada.

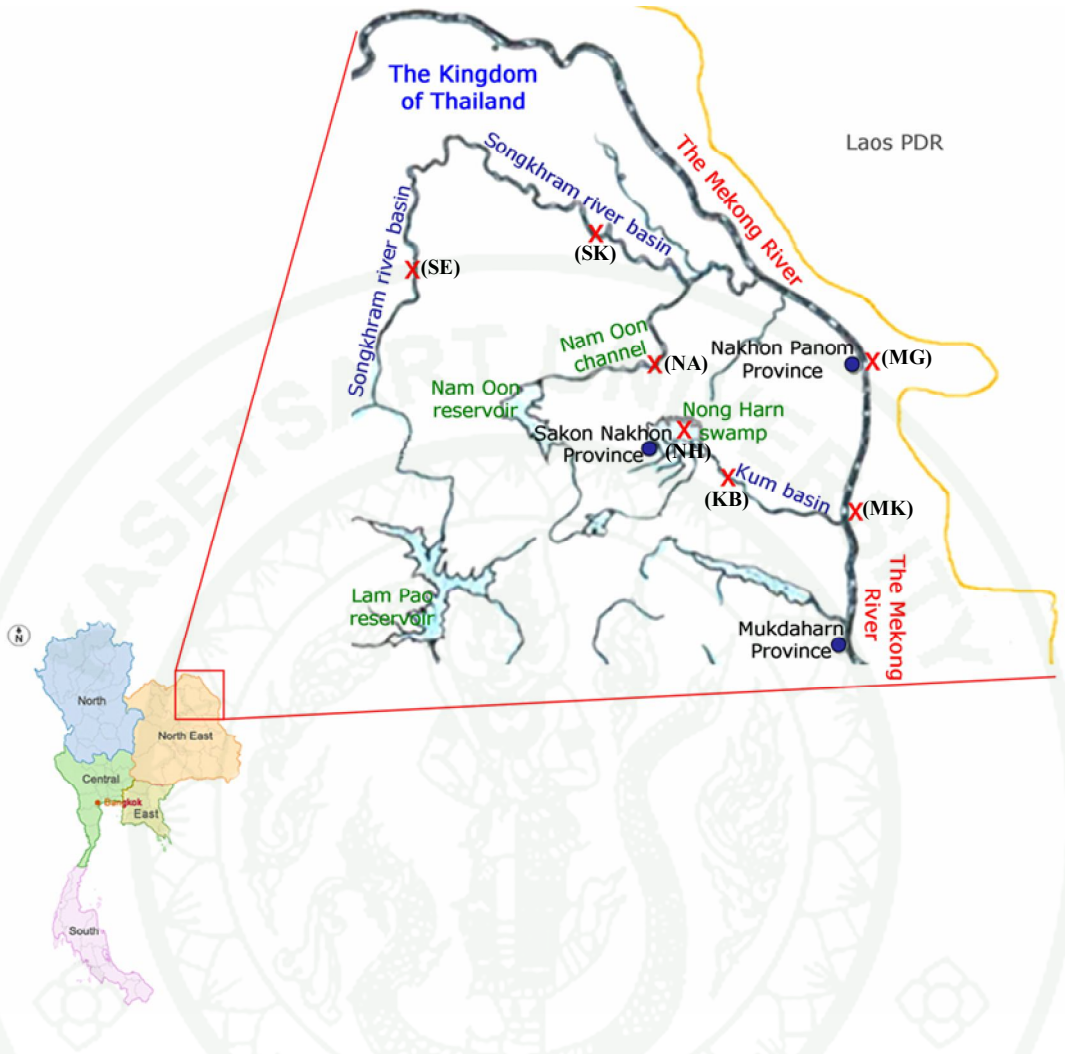
Yue, G. H., Y. Li, L. C. Lim and L. Orban. 2004. Monitoring the genetic diversity of three Asian arowana (*Scleropages formosus*) captive stocks using AFLP and microsatellites. **Aquaculture** 237: 89-102.

Zane, L., L. Bargelloni and T. Patarnello. 2002. Strategies for microsatellite isolation: a review. **Mol. Ecol**. 11: 1-16.

Zheng, X.D., M. Ikeda, A. Barinova and N. Taniguchi. 2007. Isolation and characterization of microsatellite DNA loci from the golden cuttlefish, *Sepia esculenta* Hoyle (Cephalopoda). **Mol. Ecol. Notes** 7: 40-42.

APPENDICES

Appendix A
Collecting Sites



Appendix Figure A1 Seven collection sites of whisker sheatfish *Micronema bleekeri* (Gunther, 1864) populations (the Nong-Han Lake, NH; Nam Kum basin, KB; the Mekong river at lower Nakhon Phanom province, MK; the Mekong river at upper Nakhon Phanom province, MG; Songkhram river basin at Srisongkharm district, SK; Songkhram river basin at Seka district, SE and Lam Nam Oon, NA) of whisker sheatfish *Micronema bleekeri* (Gunther, 1864) populations.

Appendix Table A1 Collection details of seven whisker sheatfish population.

Population	Number of sample	Collecting site
Nong Han Lake (NH)	30	17° 8' 31.5702"N 104° 14' 57.3648"E
Kum basin (KM)	30	16° 57' 25.4190"N 104° 30' 20.3184"E
The Mekong River (MK)	30	16° 55' 22.1592"N 104° 43' 51.4158"E
The Mekong River (MG)	30	17° 24' 9.8454"N 104° 47' 24.0036"E
Songkhram basin (SK)	30	17° 37' 30.1620"N 104° 26' 49.8942"E
Songkhram basin (SE)	30	17° 51' 52.4304"N 103° 46' 28.6464"E
Nam Oon channel (NA)	30	17° 22' 22.5156"N 103° 51' 23.4210"E

Appendix B
Reagent Preparation Protocol

1. Genomic DNA digestion with *Mse*I restriction enzyme

*Mse*I cut genomic DNA. The mixture consisted of 1 µl of 1µg genomic DNA, 5 µl of 10X buffer of *Mse*I, 1 µl of 1,000U/ml *Mse*I, 0.5 µl of 10 mg/ml BSA. Ultrapure water was added into the mixture up to 50 µl and gently mixed by inverting tube, then was incubated at 37º C overnight and check the digestion by gel electrophoresis.

2. *Mse*I adapter ligation

The digested DNA was ligated with *Mse*I adapter. One-hundred microliters consisted of 48 µl of digested genomic DNA, 8 µl of 12.5 pmol adapter, 10 µl of 10X T_4 DNA ligase buffer, 4 µl of 1U/ml T_4 DNA ligase and 30 µl ultrapure water were added. Gently mix by inverting tube. The mixture was incubated at 37º C for 1 h and then at room temperature for 2 h.

To increase yield, the PCR was performed. The mixture consists of 10 µl DNA with adapter, 2 µl of 10 pmol/µl *Mse*I primer, 3 µl dNTP, 3 µl of 10X PCR buffer, 0.3 µl $MgCl_2$, 0.2 µl *Taq* DNA polymerase and 11.5 µl ultrapure water. The PCR amplification products were checked by gel electrophoresis.

3. Hybridization with biotinylated oligo

Fragmented DNAs were hybridized with biotinylated oligo. The mixture consisted of 40 µl DNA with adapter, 4 µl biotinylated oligo, 56 µl hybridization solution (6X SSC + 0.1% SDS). Gently mix by inverting tube and incubated at 95º C for 10-15 minutes in heating block, then incubated at 65º C overnight in water bath and check the hybridization by gel electrophoresis.

4. Hybridization solutions

The stock of hybridization solution was prepared with 20X SSC and 20% SDS. For 20X SSC, dissolve 175.3 g of NaCl and 88.2 g of sodium citrate in 800 ml water. Adjust the pH to 7.0 with a few drops of 14 N solution of HCl. Adjust the volume to 1 liter with ddH_2O. Dispense into aliquots. Sterilize by autoclaving. The solution was kept at room temperature. For 20% SDS, dissolve 20g of SDS into 80 ml of ddH_2O by stirring. Add ddH_2O until final volume is 100 ml. Store at room temperature. SDS powder is hazardous. Prepare solution in a ventilated fume hood.

5. pGEM-T Easy vector system ligation

The PCR product containing microsatellite were ligated with the vector. The reaction consisted of 5 µl of 2X Rapid Ligation buffer, 1 µl 50 ng pGEM-T Easy vector, 3 µl PCR product, 1 µl T_4DNA ligase (3 unit/µl). Mix the reaction by pipetting. Incubate the reaction for 1 h at room temperature or 4º C overnight for maximum number of transforments.

6. PCR component for colony PCR

Colonies containing microsatellite were checked with PCR. The reaction consisted of 1 µl of 10X PCR buffer, 1 µl of 2mM dNTP, 0.4 µl of 50mMgCl , 0.5 µl of 5 pmol M13 forward primer, 0.5 µl of 5 pmol M13 reverse primer, 0.05 µl of *Taq* DNA polymerase, 5 µl colony containing microsatellite DNA, 1.91 µl of ultrapure water. Preparation was done on ice and gently mixing by inverting tube, then spin down and put the mixture into PCR machine. Check the PCR product by gel electrophoresis.

7. PCR component for microsatellite amplification

Microsatellite amplification were done by PCR, 12.5 µl PCR mixture consisted of 1.25 µl of 10X PCR buffer, 0.625 µl of dNTP, 0.375 µl of 50mMgCl , 0.5 µl of 10 pmol forward microsatellite primer, 0.5 µl of 10 pmol reverse microsatellite primer, 0.1 µl of *Taq* DNA polymerase, 2 µl of 25 ng genomic DNA, 7.15 µl of ultrapure water. Preparation was done on ice and gently mixing by inverting tube, then spin down and put the mixture into PCR machine. Check the PCR product by gel electrophoresis.

8. *Eco*RI cut for plasmid containing microsatellite checking

*Eco*RI cut product containing DNA fragment. Ten microliters consisted of 2 µl Plasmid or PCR product containing microsatellite, 1 µl 10X *Eco*RI buffer, 0.5 µl of 10u/ml *Eco*RI, 6.5 µl of ultrapure water. Mix the reaction by pipetting. Incubate the reaction for 1-3 h at 37º C and checking by gel electrophoresis.

9. 10 mg/ml Ampicillin

Ten mg/ml of ampicillin was prepared by mixing 1 g ampicillin and 10 ml ddH_2O. Sterilize solution through syringe filter (0.2 mM), keep 1.0 ml aliquots in individual eppendorf tubes, and store at -20º C.

10. 0.1 M IPTG

One milliliter of IPTG consisted of 0.0238 g IPTG, 1 ml of ddH_2O. The solution was kept at -20º C.

11. X-gal

The X-gal (5-bromo-4-chloro-3-indolyl-β-D- galactopyranoside) was prepared by mixing 100 mg of X-Gal, 5 ml of dimethylformamide. Dissolve 100mg of X-Gal into 5ml of dimethylformamide and divide into 1ml aliquots. Dark room preparation was recommended and solution aliquots were wrapped with aluminum foil. Store at -20°C (solution does not need to be filtered).

12. LB-amp agar

The LB containing ampicillin agar was prepared. One-hundred milliliters consisted of 3.7 g LB-agar powder and 100 ml of ddH_2O. The content was done by mixing to dissolve powder, and was autoclaved. When the solution has cooled to 50º C, 100 µl of ampicillin was added to 100 ml LB-agar solution, then let to harden for 1-2 hours and stored at room temperature or 4º C.

13. 6% PAGE

Six percentage of PAGE was prepared to 60 ml solution which consisted of 29.4 g Urea powder (7M), 6 ml of 10X TBE, 24 ml of ddH_2O, 9 ml of 40% acrylamide/bis-solution (29:1 ratio). Mix by swirling to dissolve powder. Then add 600 µl 10%APS (1 g APS + 10 ml ddH_2O) and 35 µl TEMED (N, N, N', N',-tetramethyl ethylenediamine). Mix by swirling to dissolve and pour into the gel cassette and let harden for 15 minutes. Wrap all cassettes with parafilm (food grade) and store at room temperature. Polymerization will be completed in 3 h.

14. Loading dye for agarose gel

The loading dye was prepared to 6 times (6X). The solution consisted of 0.15 g xylene cyanol, 0.15 g bromphenol blue and 50 ml of glycerol. Double distilled water was added up to 100 ml, and mixed. The working solution was prepared to 1X.

15. Loading dye for denaturing gel

The loading dye for denaturing gel was prepared by adding 9.8 ml formamide, 10 mg bromphenol blue, 10 mg xylene cyanol, 200 µl 0.5M EDTA (pH8.0), ddH_2O to 10 ml, and mixing was done to make the solution.

16. 1M Nitric acid

One molar of nitric acid was prepared by adding 62.9 ml 2M nitric acid with 437.1 ml of ddH_2O. The solution was mixed and stored at room temperature. The working solution was prepared from 45.32 ml 1M nitric acid and add ddH_2O to make 500 ml.

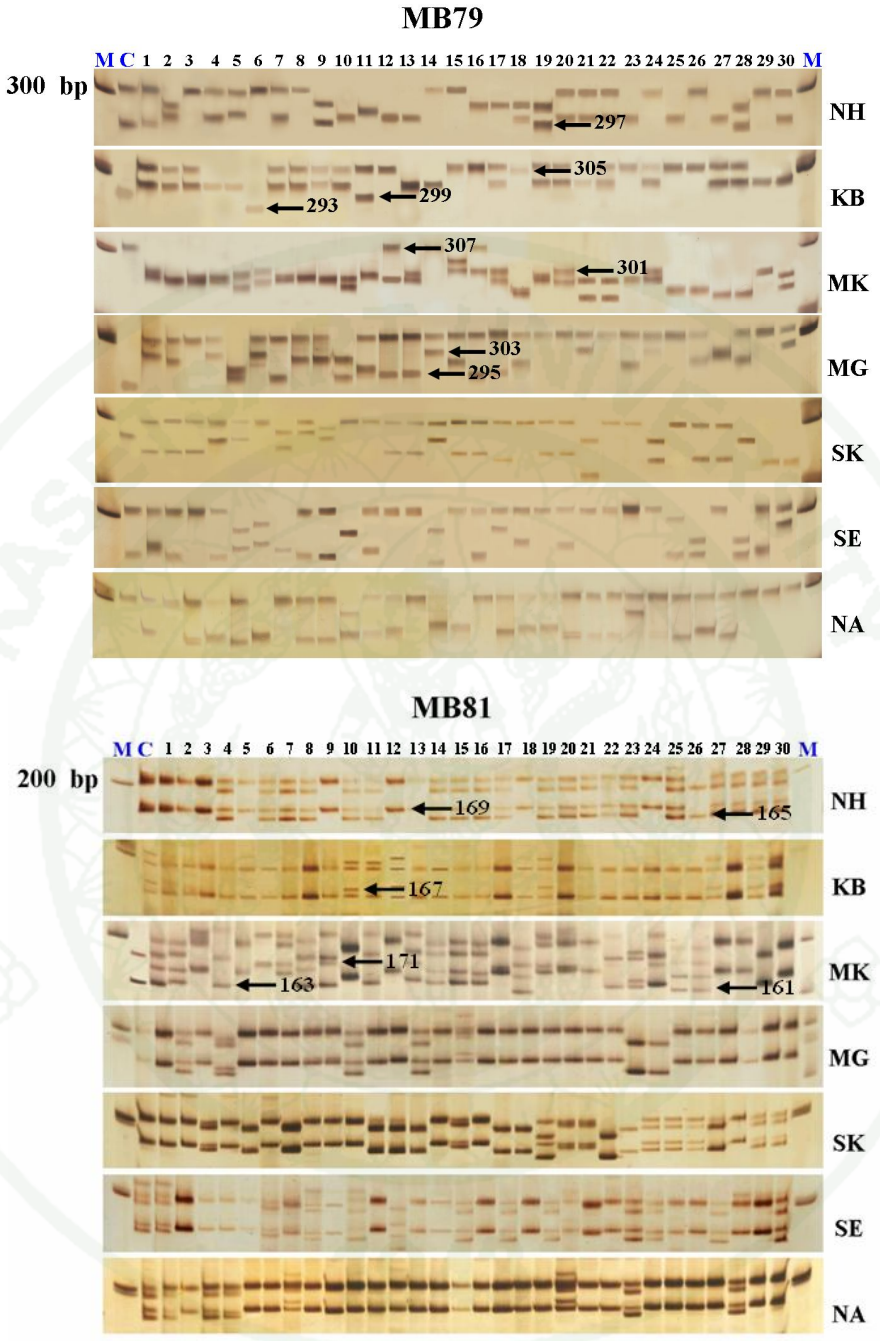
17. Developer (3% sodium carbonate + 40% formaldehyde) for silver staining

The developer was prepared by putting 24 g sodium carbonate, ddH_2O was added to make 800 ml, the solution was done by mixing to dissolve. Freshly prepare 3 % sodium carbonate and then added 100 µl 40% formaldehyde before use.

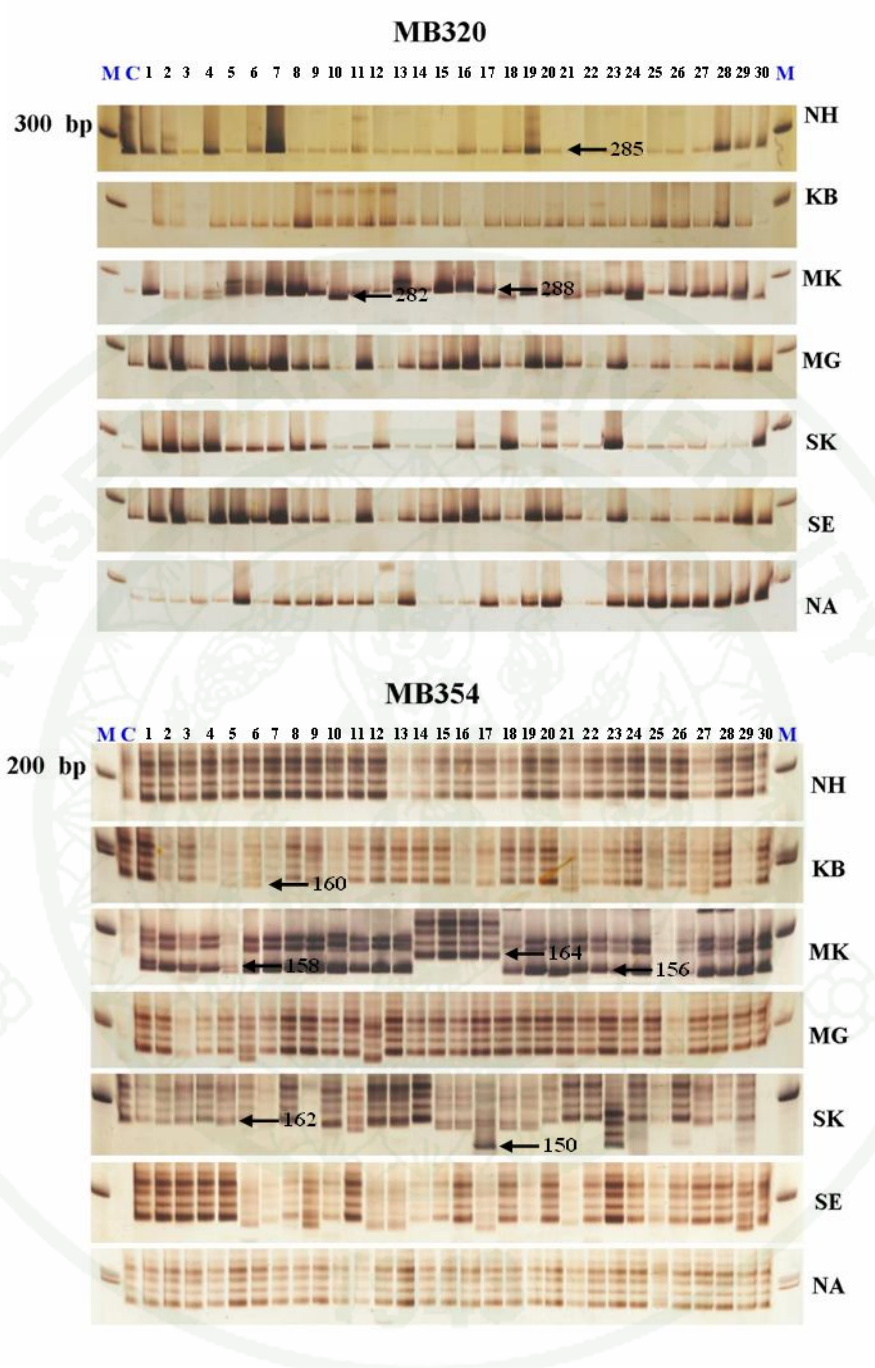
18. Stop solution for silver staining

The stop solution was prepared by mixing 20 ml glacial acetic acid, ddH_2O was added to make 1000 ml and gently mix. The solution was stored at room temperature.

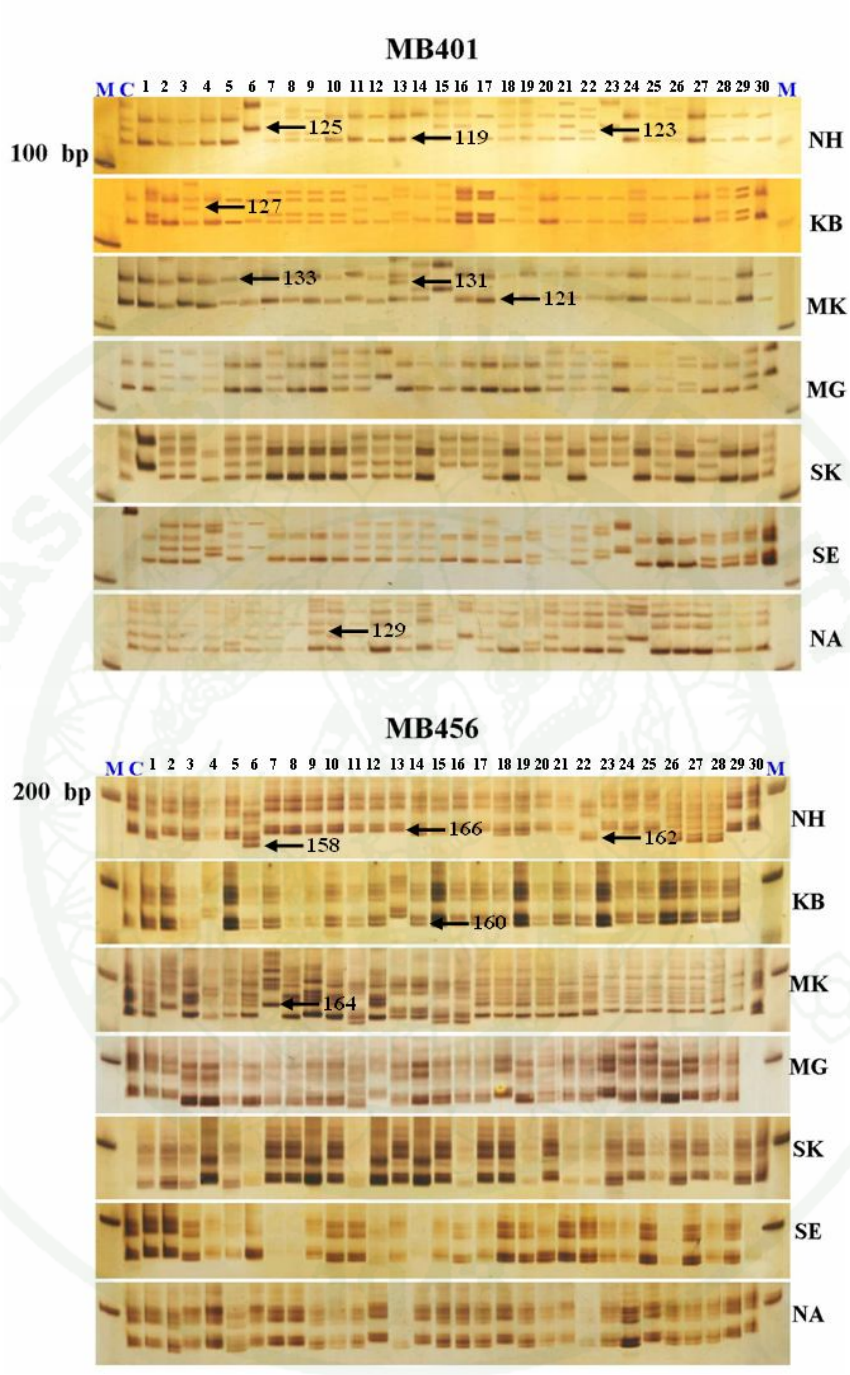
Appendix C
Allele Distribution Pattern



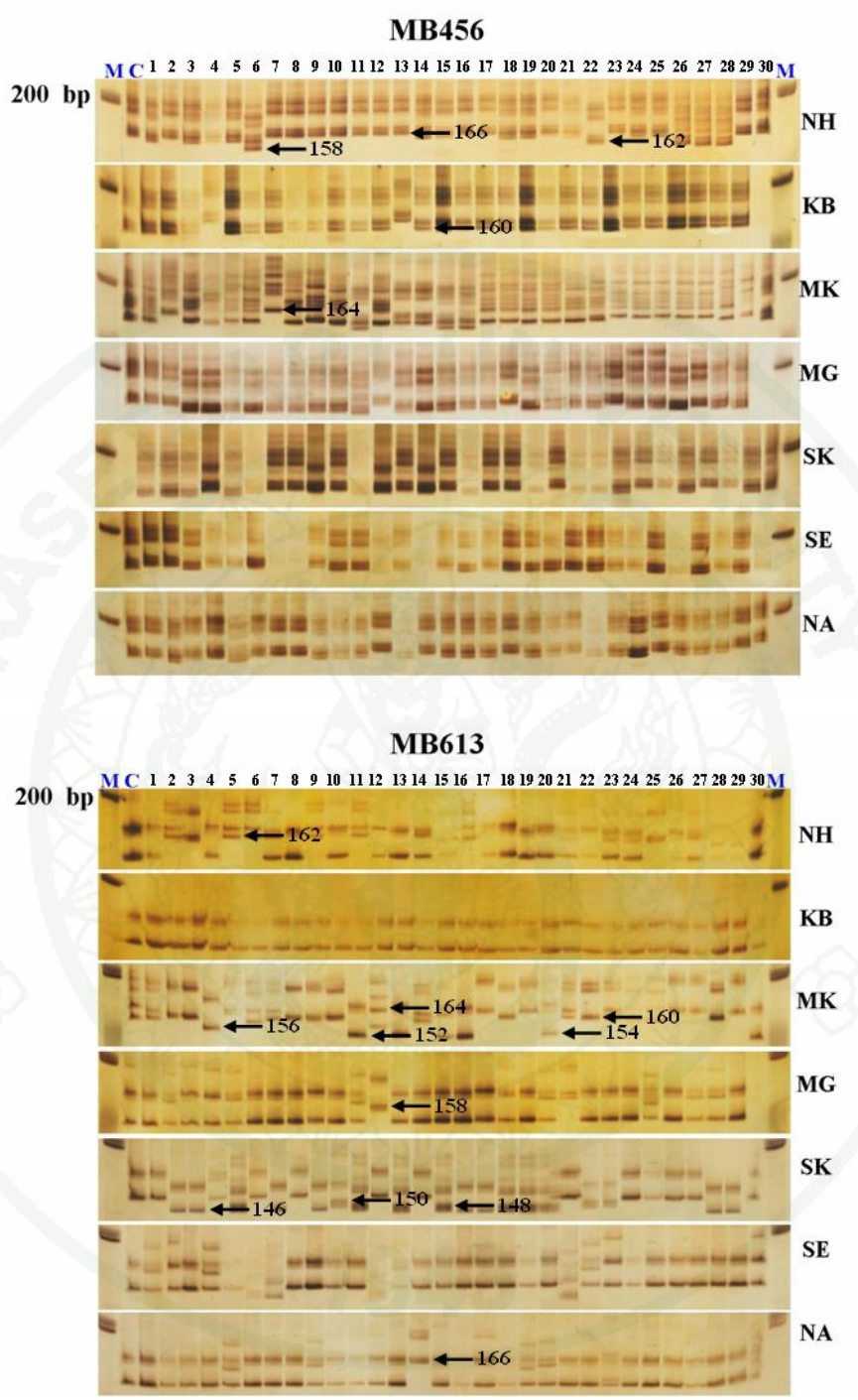
Appendix Figure C1 Allele distribution pattern of seven population at loci MB79 (293-307 bp) and MB81 (161-171 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.



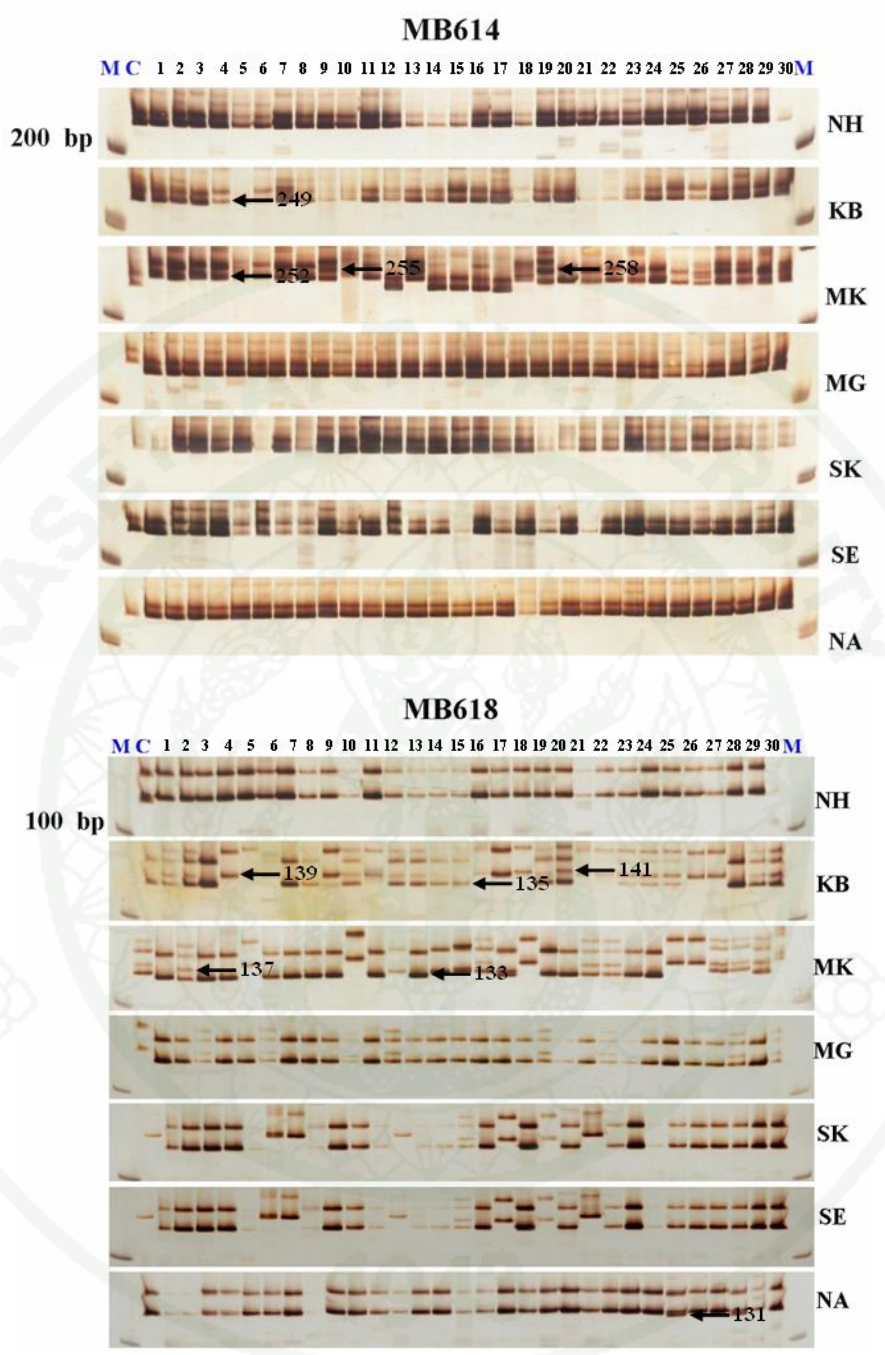
Appendix Figure C2 Allele distribution pattern of seven population at loci MB320 (282-288 bp) and MB354 (150-164 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.



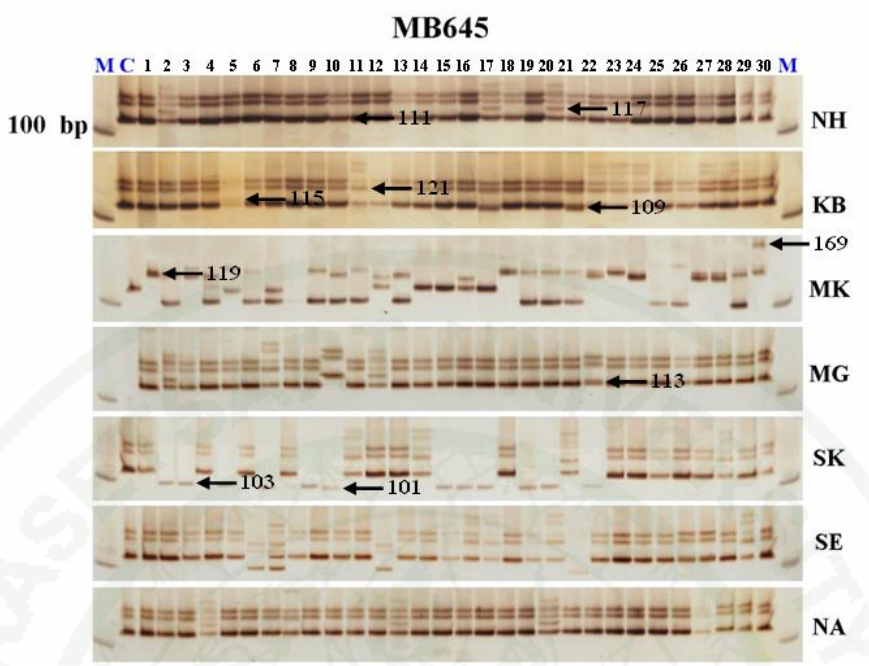
Appendix Figure C3 Allele distribution pattern of seven population at loci MB401 (119-133 bp) and MB456 (158-166 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.



Appendix Figure C4 Allele distribution pattern of seven population at loci MB515 (188-200 bp) and MB613 (146-166 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.



Appendix Figure C5 Allele distribution pattern of seven population at loci MB614 (249-258 bp) and MB618131-141 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.



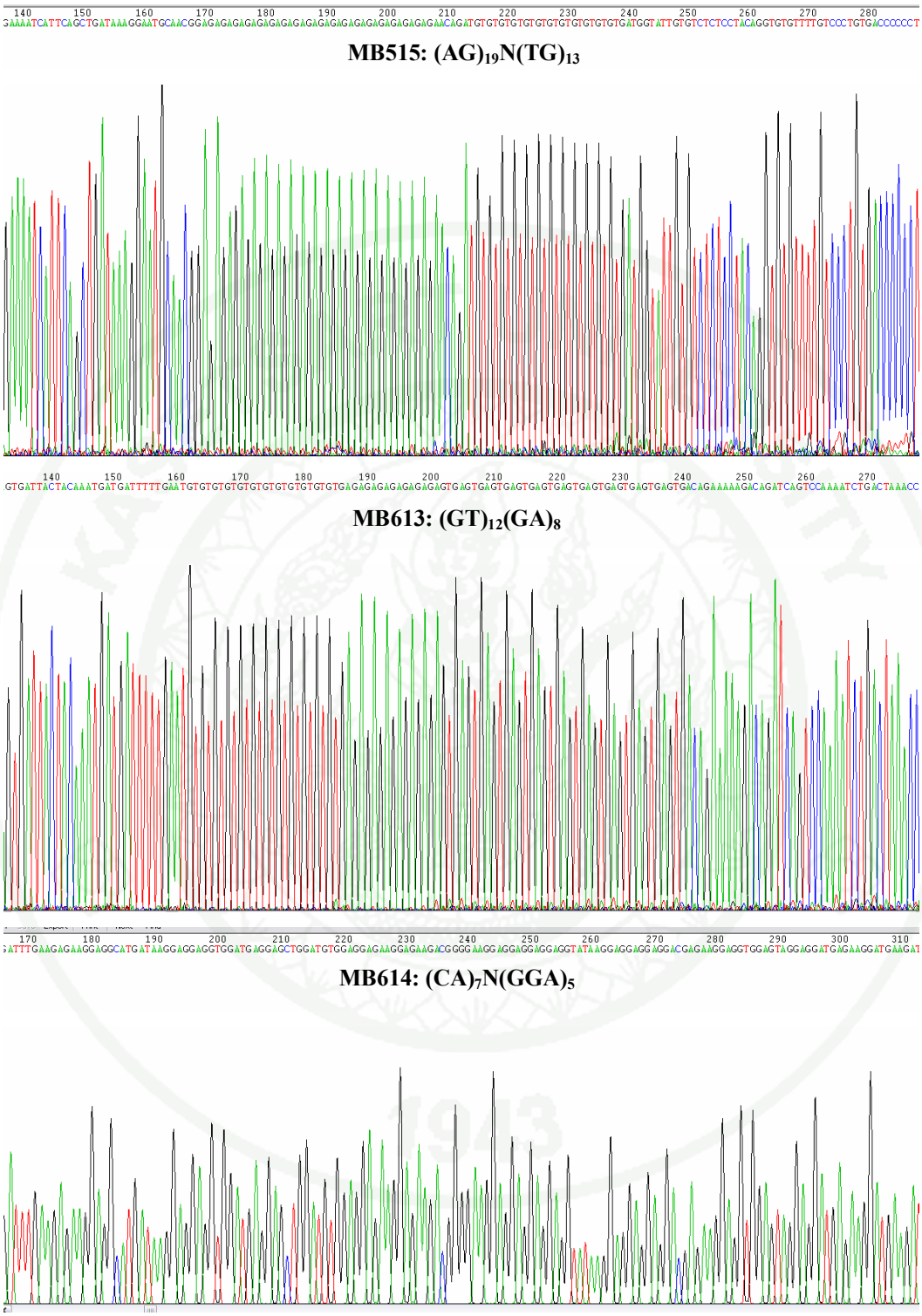
Appendix Figure C6 Allele distribution pattern of seven population at loci MB645 (101-123 bp). M is 100 bp ladder marker. C is alleles pattern of a sample of previous population; lanes 1-30 are alleles pattern of 30 individual for each population.

Appendix D

Chromatogram of Microsatellite Sequences

Appendix Figure D1 Chromatogram of microsatellite sequences (MB79, MB81and MB320).

Appendix Figure D2 Chromatogram of microsatellite sequences (MB354, MB401 and MB456).



Appendix Figure D3 Chromatogram of microsatellite sequences (MB515, MB613 and MB614).

Appendix Figure D4 Chromatogram of microsatellite sequences (MB618 and MB645).

CURRICULUM VITAE

NAME : Mrs. Piyamas Phongkaew

BIRTH DATE : October 23, 1974

BIRTH PLACE : Phranakorn Sri Ayuttaya, Thailand

EDUCATION :

YEAR	INSTITUTION	DEGREE
1996	Kasetsart Univ.	B.Sc. (Agriculture)
2000	Chulalongkorn Univ.	M.Sc. (Biotechnology)

POSITION : Lecturer

WORK PLACE : Faculty of Natural Resources and Agro-Industry Kasetsart University, Chalermphrakiat Sakon Nakhon Province Campus

SCHOLARSHIP : Office of the Higher Education Commission (CHE490140) (2007-2010)

: Kasetsart University Research and Development Institute (KURDI) (2007-2008)

: International Horizon Project, Faculty of Science (2009)

PUBLICATION : Phongkaew, P. V. Hongtrakul, N. Sangduen, A. Thongpan, U. Arunyawat, A. Swatdipong, S. Hatachote and P. Rukaewma. 2011. Isolation and characterization of eight novel microsatellite loci in whisker sheatfish *Micronema bleekeri* (Gunther, 1864). **Conserv. Genet. Resour.** 3: 425-427.