MOLECULAR GENETIC STUDIES ON COMMON CARP (*CYPRINUS CARPIO* L.)

by

Binh Thanh Thai (B. Sc., Grad. Dip)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Science and Primary Industries
Faculty of Education, Health and Science
Charles Darwin University, Casuarina, Northern Territory, Australia

August, 2006

CHARLES DARWIN UNIVERSITY

CANDIDATE DECLARATION

I certify that the thesis entitled:

Molecular genetic studies on common carp (*Cyprinus carpio* L.)

Submitted as a thesis for the degree of Doctor of philosophy of the Charles Darwin University, is the result of my own investigations, and all references to ideas and work of other researches have been specifically acknowledged. I hereby certify that the work embodied in this thesis has not already been accepted in substance for any degree, and is not being currently submitted in candidature for any other degree.

Full Name: Binh Thanh Thai

Signed...

Date...

ACKNOWLEDGEMENTS

I am very grateful to Australian Agency for International Development (AusAID) for financially supporting this research, and Charles Darwin University (CDU), Deakin University, Arafura Timor Research Facility (ATRF), Australia, and Research Institute for Aquaculture No 1 (RIA 1), Vietnam for providing facilities for my studies.

I would like to thank my principal supervisor, Professor Christopher Austin, for his support, guidance, enthusiasm and encouragement during the entire course of PhD study. I feel extremely lucky to have Dr. Chris Austin as my major professor.

I am grateful to Dr. Christopher Burridge (Associate Supervisor, Deakin University) for providing significant support for my research.

I also need to express my thanks to Dr. Sarah Smith and Mark Schultz for their valuable suggestions and for proof reading my thesis.

My research involved collecting common carp samples from many different countries around the world. I would like to take this opportunity to thank:

Prof. Li Sifa (China), Dr. Varadi Laszlo, Dr. Zsigmond Jeney (Hungary), Dr. Y. Basavaraju (India), Dr. Graham C. Mair (AIT-Thailand), Dr. Ilan Karplus, Dr. Uri Lavi (Israel), Dr. Estu Nugroho (Indonesia), Dr. Dunz D.N. (Vietnam), Ms. Yuko Lemura (Japan), Lukas Kalous (Czech public), Dr. Martin Kumar, Dr. Matthew Baranski, Vince Versace (Australia), Dr. Riho Gross (Estonia), Phuc Dinh Phan

(Research Institute for Aquaculture No III), Van Si Ngo, Hung Quang Le, Ninh Huu Nguyen, Tan Thi Nguyen, Anh The Bui (RIA1), staff of RIA1, Yen Thuy Duong (Can Tho University, Vietnam), and numerous fish farmers in Vietnam, all of whom helped me to collect the samples.

I would like to thank Dr. Tuan Anh Pham for his support and encouragement for my study.

Special thanks are given to Dr. Martin Kumar of South Australian Research and Development Institute (SARDI) for his support and advice throughout this study.

I give my sincere thanks to Dr. Kohji Mabuchi for his providing mitochondrial DNA data of Lake Biwa common carp from Japan.

I wish to thank my colleagues, both past and present, from the Molecular Ecology and Biodiversity Laboratory, who provided me with various assistance. In particular, Mark Schultz, Mike Truong, Leon Meggs, Sally Williamson, Claire McClusky, Amanda Peucker, Rebecca Overeem, Dr. Matthew Baranski, Celeste Jacq, Dr. Nicholas Murphy and Dr. Adam Miller have assisted me greatly with laboratory techniques.

I wish to thank Long Nguyen, Bac Nguyen, Hanh Nguyen, Nhu Ly and Luong Van Thing for their encouragement and support for my study while I studied in Australia.

Finally, I would like to thank my family in Vietnam for allowing me to be far away from home. Love and care from my parents (Ung Doan Thai, Ke Thi Nguyen), brothers (An Thanh Thai, Toan Thai), sisters (Phuc Hong Thi Thai, Ha Thanh Thai), sister in-law (Thoa Kim Nguyen), neice Linh Phuong Thi Thai, and nephew Nam Phuong Thai have been critical for the success of my studies.

CONTENTS

ACKNOWLEDGEMENTS iii
CONTENTS vi
LIST OF FIGURES ix
LIST OF TABLES xi
ABSTRACT xiii
CHAPTER 1
GENERAL INTRODUCTION 1
1.1 COMMON CARP 1
1.1.1 Distribution, taxonomy and translocation 1
1.1.2 Biology of common carp 3
1.1.3 Aquaculture 4
1.1.4 Common carp culture in Vietnam 5
1.2 MOLECULAR MARKERS 8
1.2.1 Molecular techniques 8
1.2.1.1 Mitochondrial DNA Sequencing 9
1.2.1.2 Single Strand Conformation Polymorphism (SSCP) 10
1.2.1.3 Microsatellites 11
1.2.2 Applications of molecular markers 12
1.2.2.1 Population genetics 12
1.2.2.2 Aquaculture 17
1.2.2.3 Conservation genetics 21
1.2.2.4 Taxonomy and evolution 22
1.2.2.5 Molecular genetic studies on common carp 29
1.3 RESEARCH OBJECTIVES AND THESIS FORMAT 32
CHAPTER 2
USING MITOCHONDRIAL NUCLEOTIDE SEQUENCES TO INVESTIGATE DIVERSITY AND GENEALOGICAL RELATIONSHIPS WITHIN COMMON CARP (*CYPRINUS CARPIO* L.)* 35
2.1 INTRODUCTION 35
2.2 MATERIALS AND METHODS 37
2.2.1 Common carp samples 37
2.2.2 DNA extraction and Amplification of Mitochondrial DNA 37
2.2.3 Data Analysis and Phylogenetic Reconstruction 38
2.3 RESULTS 39
2.4 DISCUSSION 41
CHAPTER 3
MOLECULAR TAXONOMY AND EVOLUTION OF COMMON CARP 47
3.1 INTRODUCTION 47
3.2 MATERIALS AND METHODS 51
3.2.1 Sample collection 51
3.2.2 DNA Extraction and Amplification 51
3.2.3 Phylogenetic Analyses 52

3.2.4 Phylogenetic hypothesis testing 53
3.3 RESULTS 54
3.3.1 Phylogenetic analysis 54
3.3.2 Test of taxonomic and evolutionary hypotheses 56
3.4 DISCUSSION 57
3.4.1 Sequence variation 57
3.4.2 Phylogeny and evolutionary history 58
3.4.3 Taxonomy of common carp 61
3.4.4 Conclusion 63
CHAPTER 4
GENETIC DIVERSITY OF COMMON CARP IN VIETNAM USING DIRECT SEQUENCING AND SSCP ANALYSIS OF THE MITOCHONDRIAL DNA CONTROL REGION* 73
4.1 INTRODUCTION 73
4.2 Materials and Methods 76
4.2.1 Sample collection 76
4.2.2 DNA extraction and sequencing of control region 76
4.2.3 Single Strand Conformation Polymorphism (SSCP) amplification 77
4.2.4 Data analysis 78
4.3 RESULTS 79
4.3.1 Control region sequences and SSCP variation 79
4.3.2 Genetic differentiation and relationships among populations 81
4.4 DISCUSSION 83
4.4.1 Control region sequence variation and utility of the SSCP techniques 83
4.4.2 Genetic diversity of common carp in Vietnam 85
4.4.3 Insights into the dissemination of cultured common carp in Vietnam 87
4.4.4 Conservation of wild carp stocks 89
4.4.5 Conclusion 90
CHAPTER 5
GENETIC DIVERSITY OF COMMON CARP (*CYPRINUS CARPIO* L.) IN VIETNAM USING FOUR MICROSATELLITE LOCI* 99
5.1 INTRODUCTION 99
5.2 MATERIALS AND METHODS 102
5.2.1 Sample collection and DNA extraction 102
5.2.2 Data collection 103
5.2.3 Genetic diversity analysis 103
5.2.4 Variation at microsatellite loci and tests from Hardy-Weinberg equilibrium 103
5.2.5 Genetic differentiation between populations 104
5.3 RESULTS 105
5.3.1 Within population variation 105
5.3.2 Genetic differentiation and relationships among populations 106
5.4 DISCUSSION 109
5.4.1 Genetic diversity within common carp populations 109
5.4.2 Differentiation between populations 111
5.4.3 Comparison of microsatellite and mitochondrial DNA data 113

CHAPTER 6
PHYLOGENETIC EVALUATION OF SUBFAMILY CLASSIFICATION OF THE CYPRINIDAE, FOCUSING ON VIETNAMESE SPECIES 125
6.1 INTRODUCTION 125
6.2 MATERIALS AND METHODS..128
6.2.1 Sample collection 128
6.2.2 DNA extraction and PCR amplification 128
6.2.3 Data analysis 129
6.2.4 Phylogenetic hypothesis testing 130
6.3 RESULTS 131
6.3.1 Sequence variation 131
6.3.2 Phylogenetic analysis 132
6.4 DISCUSSION..135
CHAPTER 7. 152
GENERAL DISCUSSION 152
7.1 GLOBAL GENEALOGY, TAXONOMY AND EVOLUTION OF COMMON CARP 153
7.2 GENETIC VARIATION AND DOMESTICATION OF VIETNAMESE COMMON CARP 154
7.3 UTILITY OF SSCPS AND MICROSATELLITES 157
7.4 UTILITY OF MICOCHONDRIAL DNA GENE REGIONS 158
7.5 CLASSIFICATION AND EVOLUTION OF THE CYPRINIDAE...159
7.6 FURTHER STUDY 160
APPENDICES..163
BIBLIOGRAPHY 169

LIST OF FIGURES

Figure. 2.1. Tree reconstruction derived from *ATPase6/ATPase8* and *CR*, using HKY+I+G model of evolution. Bootstrap values are based on 1,000 replicates, neighbour-joining (normal font) parsimony (*italic*), Bootstrap values are given for nodes with at least 50% or more support................. 46

Figure. 3.1. Phylogenetic hypotheses examined based on the species studied. Zhou *et al*. (2004b), Baruš *et al*. (2002) and Kohlmann *et al*. (2005) indicate representative topologies based on their more extensive studies............. 68

Figure. 3.2. Neighbour-joining estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations (- = <50% bootstrap support)... 69

Figure. 3.3. Maximum likelihood estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations (- = <50% bootstrap support)... 70

Figure. 3.4. Bayesian estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations. 71

Figure. 3.5. UPGMA clustering of common carp populations based on microsatellite variability at four loci derived from Kohlmann *et al*. (2005).................... 72

Figure. 4.1. Collection localities for *Cyprinus carpio* L. samples in Vietnam.......... 95

Figure. 4.2. Silver stained polyacrylamide gel showing the eight SSCP variants detected in common carp populations in Vietnam (a). Neighbour-joining tree reconstruction derived from *CR* sequences, using HKY+I+G model of evolution. Bootstrap values are based on 1,000 replicates. Bootstrap value is given for nodes with at least 50% or more support (b).......................... 96

Figure. 4.3. Relationships among common carp from wild and hatchery population in Vietnam using the unbiased genetic distance of Rogers (1972) and UPGMA joining method. A, B, and C are SSCP haplotypes which predominate in each cluster.. 97

Figure. 4.4. MDS plot of pairwise *F*st value among hatchery and wild populations of common carp in Vietnam. Population codes are given in Table 4.1......... 98

Figure. 5.1. Allele number and distribution of allele frequencies for four loci between Hungarian, Indonesian and Vietnamese common carp populations in experimental group.. 122

Figure. 5.2. UPGMA dendrogram of common carp populations in Vietnam based on matrixes of genetic distance (Nei *et al*., 1983)...................................... 123

Figure. 5.3. MDS plot of genetic distance (Nei *et al.*, 1983) among experimental, hatchery and wild populations of common carp in Vietnam. Population codes are given in Table 4.1. 124

Figure. 6.1. Phylogenetic tree resulting from neighbour-joining analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support . The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combined three region data sets. 144

Figure. 6.2. Phylogenetic tree resulting from maximum parsimony analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support. The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combined three region data sets. 145

Figure. 6.3. Phylogenetic tree resulting from maximum likelihood analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support. The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates sample used in combine three region data sets.. 146

Figure. 6.4. Phylogenetic tree resulting from Bayesian analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent Bayesian posterior probabilities. The subfamily group are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combine three region data sets. 147

Figure. 6.5. Neighbour-joining tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe.. 148

Figure. 6.6. Maximum parsimony tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe. 149

Figure. 6.7. Maximum likelihood tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe. 150

Figure. 6.8. Bayesian tree from combined *16S*, *Cyt b* and *CR* mtDNA data. The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe. 151

LIST OF TABLES

Table 1.1. Type of DNA marker, their characteristics, potential applications and cost* 34

Table 2.1. Sample code, collecting locality and number of individuals sequenced (N) for each strain or population. 45

Table 3.1. Sources, status, GenBank accession numbers of strains of common carp (n number of samples sequenced) and taxonomy considered by Baruš *et al*. (2002), Zhou *et al*. (2004b), Kohlmann *et al*. (2005). 64

Table 3.2. Tests of alternate phylogenetic hypotheses using Kishino-Hasegawa (KH), Shimodaira-Hasegawa (SH) test. (* significant difference between optimal and alternate topologies, P <0.05). 66

Table 3.3. Nucleotide divergence (%) of common carp and other fish in *Cyt b* gene and *CR*. 67

Table 4.1. Location, code and number of samples sequenced and analysed by the SSCP technique. 91

Table 4.2. Number of haplotypes and haplotype diversity in each common carp population. Population code given in Table 4.1 92

Table 4.3. Pair-wise estimate of variance of haplotype frequencies (*F*st) among of samples. Population codes given in Table 4.1 93

Table 4.4. AMOVA results for three groups (experimental, hatchery, wild) of 20 common carp populations base on SSCP data. (Intra = intrapopulation, Inter = interpopulation, values are %) 94

Table 5.1. Characteristics of *Cyprinus carpio* microsatellite loci tested 116

Table 5.2. Genetic variability of four microsatellite loci in 20 populations for common carp in Vietnam. Population codes given in Table 4.1 117

Table 5.3. Number of private alleles at four microsatellite loci in 20 common carp populations and group. The population codes are given in Table 4.1 118

Table 5.4. Pairwise *F*st values between 20 common carp populations in Vietnam based on four microsatellite loci. Population codes given Table 4.1. 119

Table 5.5. Results of assignment test (self-classification) of common carp individuals based on four microsatellite loci. Population code are given in Table 4.1. 120

Table 5.6. Genetic diversity of experimental, hatchery and wild common carp groups in Vietnam based on variation at four microsatellite loci (Intra = intrapopulation, Inter = interpopulation, values are %) 121

Table 6.1. Species, sampling localities, GenBank access numbers subfamily designation as proposed by previous studies.(* samples from this study) 139

Table 6.2. Major hypotheses for phylogenetic relationships of cyprinid species.... 141

Table 6.3. Summary of results of phylogenetic analysis in three mitochondrial DNA gene regions of cyprinid species. 142

Table 6.4. Tests of alternate phylogenetic hypotheses using combine *16S*, *CR*, *Cyt b* regions, Shimodaira-Hasegawa (SH) test. (* significant difference between optimal and alternate topologies, *P* <0.05... 143

ABSTRACT

Common carp (*Cyprinus carpio*. L) is the world's oldest domesticated and the most important aquaculture species. The overall objective of this thesis is to advance the understanding of genetics of common carp with a special focus on Vietnamese domesticated and wild populations.

The first part of the study demonstrated the potential of mitochondrial DNA for the analysis of genetic diversity, origin, divergence and genealogy of carp strains and populations. The second part, also using mitochondrial DNA sequences, examines taxonomic questions and evolutionary hypotheses focussing on samples originating from wild populations from across Eurasia. Overall, nucleotide divergence within common carp and genetic analyses were not consistent with contemporary views on the taxonomy of the species and, somewhat controversially, supports an Asian origin for European carp populations.

The third and fourth parts of the thesis present an examination of genetic diversity in introduced and indigenous domesticated stocks and wild populations of common carp from Vietnam using DNA direct sequencing, single strand conformation polymorphism analysis of the mitochondrial DNA control region and microsatellite markers. Both mitochondrial and nuclear markers were effective in distinguishing among experimental strains and showed high levels of diversity within most hatchery stocks, consistent with the crossbreeding and dissemination of experimental lines. The molecular genetic data sets showed a high level of concordance in the patterns of relationships among populations and stocks. The genetic data indicates that some

wild populations of Vietnamese common carp are genetically distinct indicating the need for management strategies to preserve these gene pools.

The last part of the thesis presents a phylogenetic study of the taxonomic classification and relationship within subfamilies of the family Cyprinidae, including the subfamily Cyprininae to which common carp belong, using data from three mitochondrial DNA gene regions.

Chapter 1.

General Introduction

1.1 Common carp

1.1.1 Distribution, taxonomy and translocation

Common carp (*Cyprinus carpio* L.) occur in water ways throughout Eurasia, from Western Europe through to China, Korea, Japan and South East Asia and from Siberia, south of latitude 60^0 N, to the Mediterranean and India (Kohlmann, Kersten, 1999). The actual origin and natural distribution of common carp is disputed by some authors due to its long history of domestication in both Europe and Asia which have led to many translocations over a significant period of time (Balon, 1995). Because of the long documented cultivation history of common carp in China, some scientists considered that the ancestor of European domestic carp were derived from Asian common carp stocks, during ancient Greek and Roman periods (Chiba *et al.*, 1966; Vooren, 1972). Others consider that the common carp is indigenous to Europe and in fact postulate a European origin for wild carp and subsequent dispersal east to Siberia and China (Balon, 1995; Kottelat, 1997).

Due to its popularity as an aquaculture and ornamental species, common carp has been widely translocated, outside its European and Asian distribution. As a result of these transfers and introductions, it is now perhaps the most widespread species of freshwater fish in the world, with naturally reproducing populations established in

many countries in both northern and southern hemispheres. FAO (1998) lists introductions of common carp into 124 countries, of which 81 are recorded as having established feral populations or viable aquaculture stocks.

Cyprinus carpio is taxonomically a most confusing species with over 30 synonyms listed on FishBase (http://www.fishbase.org/search.php). Based on just recent taxonomic treatments, opinions vary widely: Balon (2004) and Kirpitchnikov (1999) recognised three subspecies: European common carp (*Cyprinus carpio carpio* Linnaeus, 1758), Far Eastern common carp (*C. carpio haematopterus* Temminck & Schlegel, 1845), and South East Asian common carp (*C. carpio viridiviolaceus* Lacepede, 1803), whereas Kottelat (2001) recognised European common carp as *Cyprinus carpio* Linnaeus, 1758 and Asian common carp *Cyprinus rubrofuscus* Lacepede, 1803. These taxa are distinguished mainly by morphology, such as number of gill rakers and scales, shape and color characteristics. However, there are often overlaps in these traits which may be in part due to stock mixing and hybridization which may have blurred taxonomic boundaries. Moreover, the domestication of common carp has not only led to changes in body proportion, scalation and color, but also in physiological characteristics (Balon, 1995; Baruš *et al.*, 2002).

The taxonomic uncertainties, compounded by translocations, and the development of phenotypically distinctive domesticated lines, demands the use of molecular markers to elucidate the taxonomy of common carp and genetic relationships among stocks (Balon, 1995; Lever, 1996). There is also a clear need to examine the status of possible wild common carp gene pools to determine if there is a need to conserve genetic resources in this species. Lever (1996) considered that the wild ancestor of

the common carp may be a single species, *Cyprinus carpio*, widely distributed from the Danube to Amur Rivers. While the existence of several self-sustaining wild populations of carp have been documented (Kohlmann, Kersten, 1999; Paaver, Tammert, 1993), Komen (1990) believes that truly self-sustaining wild populations of common carp are probably rare.

1.1.2 Biology of common carp

Common carp is a largely benthic species that prefers shallow water habitats covered with aquatic weeds and grasses. It is an omnivorous fish that mostly feeds on the bottom but can exploit all levels in the water column. The natural diet of carp is dominated by chironomids, snails, young clams, shrimps and other benthic animals. This species also consumes aquatic plants, filamentous algae, seeds of plants and organic detritus. Under pond culture conditions, common carp takes soybean and peanut cakes, rice and wheat bran (Zhong, 1989).

Common carp may be sexually mature as early as the end of its first year; however it typically requires three to four years to reach sexual maturity (Cooper, 1987). According to Linhart *et al*. (1995) common carp have a high fecundity for a freshwater species producing 100,000 to 300,000 eggs per kg body weight with reports of as many as 360,000 to 599,000 eggs per female. The eggs are sticky in nature, and attach to aquatic weed and other material after spawning.

Common carp reach 0.6 to 1.0 kg body weight within one season in the polycultural fish ponds of subtropical/tropical areas (FAO, 2004). Currently the world record for the largest carp stands at 34.3 kg. In their natural range carp can live up to 15 years,

however they have been reported in some areas living to over 24 years of age with males often living longer than females (Balon, 1995).

1.1.3 Aquaculture

Common carp is the most extensively cultivated freshwater fish species in the world (Chiba *et al.*, 1966; Komen, 1990; Wohlfarth, 1984; Zhou *et al.*, 2004b). This fish has several advantages that make it popular for commercial culture: (1) very fast growth rate, (2) high environment tolerance, (3) ease of handling, (4) ability to be raised in high density, (5) ability to utilise artificial diet with relatively low protein content, and (6) occurrence of highly productive strains and breeds produced from long-term domestication and selective breeding (Kirpitchnikov, 1999).

Culturing and breeding of common carp has a long history dating back about 4,000 years in China and close to 2,000 years in Europe. Several special breeding centres have been developed in different regions of Europe, like the Czech Republic, Germany and Hungary, as well as Russia and Ukraine. China and Japan are the ancient culturing centers in Asia, but during the last decades India, Indonesia and Vietnam have started to culture common carp as a result of deliberate fish importation and acclimatization activities (Bakos, Gorda, 2001).

Annual production of common carp worldwide is over three million metric tons (FAO, 2003). In Asian countries, common carp contributed 17% to total carp production from aquaculture in 2001 (Gupta *et al.*, 2005). Li (2001) reported that production of common carp reached 2.05 million tons in 1999 and accounted for 20%

of total freshwater fish output in China. In Indonesia, common carp production reached 178,362 tons in 1996, and accounts for 54.3% of total cultured freshwater fish in that country (Hardjamulia *et al.*, 2001). In Europe, common carp is by far the most important freshwater fish with annual aquaculture yield of about 220,000 tons (Linhart *et al.*, 2002). Breeding and culture of common carp has been the backbone of fish farming in many European countries. For example, in 2001 common carp production reached 14,000, 10,500, 17,000 and 21,000 tons in Hungary, Germany, Czech, and Poland respectively. Common carp are cultured in a wide range of environments including ponds, cages, tanks, reservoirs and rice-fields, as monoculture and polyculture with a variety of other species (FAO, 2001).

In addition to production for food, common carp have been selectively bred for a variety of colors and color patterns for the ornamental fish market (Balon, 1995). Best known varieties are Koi carps which are called "swimming flowers" and are among the most expensive of ornamental fish species. Amano (1968) recorded an annual production of some 10 million fish amounting 1,000,000,000 yen in Japan. Although originally developed in Japan, Koi carps are now cultured in many parts of the world, including China, Europe and America (Balon, 1995).

1.1.4 Common carp culture in Vietnam

Common carp (*Cyprinus carpio* L) is thought to be indigenous to northern Vietnam, and translocated to southern Vietnam (Nguyen, Ngo, 2001). Eight local varieties have been recognised in Vietnam: white carp, Bac Kan carp, high body carp, Ho Tay carp, South Hai Van carp, red carp, violet carp, and reduced scale carp. These varieties differ in morphology, color, distribution, and some other biological

characteristics (Tran, 1983). The white carp is one of the most popular and important fish in aquaculture in Vietnam. In the study by Tran (1983) it was found that indigenous strains of common carp have poor growth rates and highly variable color and scale phenotypes. For example, the Bac Kan strain is usually elongated in shape and is morphologically distinct from other strains. It commonly attains a weight of only 70-80 g in first year in polyculture and 160-200 g in the second year when grown in low input rice-field. Nevertheless, it is an important strain as it is adapted to the rice-field environment and farmers can maintain broodstock that do not depend on wild seeds. The most important traits of the Bac Kan strain are that it can be cultured in shallow water and is tolerant of fluctuations of water temperature that characterises this environment. Another useful trait for culture in terraced rice-field, in which water flows from terraced field to the next below, is that the fish rarely leave the fields even during flooding when water spills across dikes (Edwards *et al.*, 2000). The local “rice field” common carp strains are called “resident fish” or “fix- home fish” because of this useful characteristic (Tran, 1983).

In general, Vietnamese common carp is considered a good aquaculture species because it exploits natural pond productivity, has good survival rate and the flesh has low fat content (Bakos, Gorda, 2001). There are however concerns about wild carp stocks which are thought to be in decline because of excessive harvesting and crossbreeding with introduced carp stocks (Nguyen, Ngo, 2001).

A government supported common carp selective breeding program has been in place in Vietnam for over 30 years with the aim of developing and disseminating strains with high growth and survival rates to fish farmers (Tran, Tran, 1995). For these

reasons, Hungarian and Indonesian carp were imported in Vietnam in the 1970s. Three hybrid lines of common carp were obtained by crossbreeding among three lines (Vietnamese white, Hungarian scale and Indonesian yellow carp) which were also used as the initial material for a common carp selective breeding program in Vietnam. Three different hybrid lines called “three blood” carp were made by crossing (1)✿Hungarian X ❝ (Indonesian X Vietnamese), (2) ✿Vietnamese X ❝ (Indonesian X Hungarian), (3)✿Indonesian X ❝ (Hungarian X Vietnamese).

Mass selection of three blood carp was carried out at the Research Institute for Aquaculture No. 1 (RIA1), Bac Ninh, Vietnam for six generations. In the first generation the number of fish obtained for mass selection ranged from 1,720, to 400. For the second generation the number of fish in each stock was reduced to 250, however the selection differential in the second generation was higher than that in the F1. In the F3 generation the total number of fish obtained for selection was reduced further, due to poaching of some of the fish stocks. Generally around 20% of fishes in each stock were selected and progressively the selection intensity and the selection differential declined. The fifth generation showed a 33% increase in growth rate in compared to the base population. However, the realized heritability (h^2) of the body weight of fish gradually declined from 0.20 in first generation to nearly zero in the sixth generation (Nguyen *et al.*, 2005; Tran, Nguyen, 1992). In addition, the distribution of seed stock to farmers was only partly successful due to mixing of stocks in hatcheries and on farms. Inbreeding depression and undesirable genetic effects may be a problem as many farmers report carp growth is slow and fish mature when they are a small size. Furthermore, pure stocks of experimental common carp

strains kept in separate ponds at RIA1 may have been mixed (A. T. Pham pers.comm).

There are a number of important issues that need to be addressed concerning Vietnamese common carp stocks including their genetic status and history, the extent of dissemination of selectively bred stocks and the conservation status of wild stocks. These issues can be addressed by obtaining molecular genetic information which can provide insights into genetic diversity and therefore, inbreeding and the extent of stock mixing in both cultured and wild stocks. In addition, molecular genetic data can contribute to the resolution of issues relating to the taxonomy, evolution, and biogeography of common carp and help identify genetically divergent wild stocks of conservation significance.

1.2 Molecular markers

1.2.1 Molecular techniques

Molecular genetic data can be generated using a variety of techniques. The data generated, representing a range of different loci, are now commonly referred to as molecular markers and are applied to a wide range of problems and issues in a variety of fields (Hillis *et al.*, 1996a; Liu, Cordes, 2004). The techniques used to generate molecular markers for the research documented in this thesis and the relevant fields of genetic research are outlined and discussed in the following sections.

1.2.1.1 Mitochondrial DNA Sequencing

Nucleotide sequence differences among individuals can be examined directly by DNA sequencing. The technique is now becoming a routine, albeit still relatively expensive procedure since the development of the dideoxy chain termination method (Sanger *et al.*, 1977). In combination with the Polymerase Chain Reaction (PCR) (Mullis *et al.*, 1986; Saiki *et al.*, 1988), this method provides a means for collecting precise data from short DNA fragments, and it has proven to be especially powerful when applied to the analysis of various regions of the animal mitochondrial genome (Parker *et al.*, 1998).

Techniques using mitochondrial DNA (mtDNA) sequence information have been widely employed for investigating the taxonomy, genealogical relationships, origins and diversity of domesticated animal species because this molecule has several useful characteristics. Firstly, mtDNA is maternally inherited. It is especially sensitive to the detection of reduced population size and recent population isolation (Avise, 1994). Secondly, the rate of mtDNA mutation appears to be as much as 10 times faster than that of nuclear DNA, and it is non-recombining making it effective for detecting recent population isolation and for establishing genealogical relationships between populations within species (Avise, 2000). Direct DNA sequencing has been successfully used to analyse relationship of domesticated animal species including, fish (Wang, Li, 2004; Watanabe *et al.*, 2005), shrimps (De Francisco, Galetti, 2005), pigs (Kim *et al.*, 2002), rabbits (Long *et al.*, 2003), buffalo (Kierstein *et al.*, 2004), goats (Manjunath *et al.*, 2004) and deer (Cronin, 2003; Kuznetsova *et al.*, 2005). The following sections review how these markers are generated and applied to population, aquaculture, conservation, taxonomic and evolutionary studies.

1.2.1.2 Single Strand Conformation Polymorphism (SSCP)

Single strand conformation polymorphism (SSCP) analysis allows the detection of nucleotide difference in short DNA fragments (100-300bp) (Thomas, Robert, 1996) due to mobility differences of single strand DNA molecules (ssDNA) with secondary structure. A DNA sample, usually a PCR product, is denatured by heat or chemical denaturants allowed to form a single strand conformation and electrophoresied in a non-denaturing gel. Intra-strand base pairings cause folding of the fragments into stable conformations and nucleotide differences between fragments result in different conformations. Mobility differences among folded strands can be detected under the appropriate electrophoretic condition (William *et al.*, 1998).

The use of the SSCP technique for examining genetic variation in mtDNA fragments in population and conservation genetic studies is becoming increasingly popular (Sunnucks *et al.*, 2000; Thomas, Robert, 1996). Indeed, it offers a sensitive but inexpensive and rapid method for determining which DNA samples in a set differ in sequence, and so can greatly reduce the amount of sequencing (Durand *et al.*, 2005; Hayashi, 1992; Hayashi, Yandell, 1993; Weder *et al.*, 2001). SSCP studies have been conducted to investigate mitochondrial *CR* sequence variations among population of a range of fish species (Bernal-Ramirez *et al.*, 2003; Hoarau *et al.*, 2004; Julie, Louis, 2003; Turgeon, Bernatchez, 2001), New Zealand greenshell mussel (Apte, Gardner, 2002; Apte *et al.*, 2003) and the donkey (Ivankovic *et al.*, 2002). It has also been used to investigate genetic divergence in inbred lines of horses (Kavar *et al.*, 1999; Marklund *et al.*, 1995; Mirol *et al.*, 2002a). The SSCP technique has also been used for identification of fish products in the market places (Rehbein, 2005; Rehbein *et al.*,

1999; Weder *et al.*, 2001). To date, no studies of common carp populations using the SSCP technique have been reported.

1.2.1.3 Microsatellites

Microsatellites, are simple sequence repeats of 1-6bp (SSRs) with great variability among individuals, and exhibiting large numbers of alleles (Cross, 2000). Once isolated and characterised, microsatellites can be amplified by PCR (Cross, 2000). Compared to other popular marker systems, microsatellites offer the advantages of codominant inheritance and high polymorphism (Li, 1997; Liu, Cordes, 2004).

Microsatellites have been extremely useful in fish, crustacean, and mollusc population genetic and aquaculture studies, and has become the marker of choice for a variety of applications (Baranski *et al.*, 2006; Karsi *et al.*, 2002; Liu, Cordes, 2004; Xu *et al.*, 1999) (see 1.2.2.2). However, disadvantages of microsatellites are that identifying these regions from a genome library for new species can be time-consuming and expensive. Known primers are not usually effective in amplifying the same locus across related taxa unless the flanked regions are conserved (Ellegren, 1992). Genotyping of microsatellite loci can be time consuming and expensive. However, multiplexing of loci using different fluorescent dyes coupled with the use of automated genetic analyser or sequencers can increase the efficiency microsatellite genotyping (Dzialuk *et al.*, 2005; Hailer *et al.*, 2005; Tommasini *et al.*, 2003).

1.2.2 Applications of molecular markers

1.2.2.1 Population genetics

The field of population genetics encompasses the description of the distribution of genetic variation within and between populations, thereby providing indirect information on population isolation or structure. Populations may exhibit different kinds of genetic structure, and determining the appropriate structure is of critical importance for understanding population biology and history (Baverstock, Moritz, 1996). More broadly, population genetics provides information on genetic diversity that can be used to measure levels of inbreeding, gene flow, population subdivision, and migration rates (Ward, Grewe, 1995; Weir, 1990). Information of this kind is essential for sustainable exploitation of fish species (Ward, Grewe, 1995).

With respect to population genetic structure studies, in the most extreme case of no genetic variation, a species may consist of only a single population unit (i.e. panmixia), or many isolated subpopulations but with no genetic differences. However, genetic subdivision of species into multiple populations is common and can be of several different kinds (Johnson, 2000). Thus, a species can consist of a series of subpopulations in which genes are likely to be exchanged with adjacent populations (the stepping-stone model), or each subpopulation is equally likely to exchange genes with any other subpopulation (the island model), or a series of isolated subpopulations within which individuals exchange genes but the more isolated the subpopulations are, the less the likelihood of gene flow between them (the isolation-by-distance model). These three different models of population structure result in different patterns of genetic differentiation within and between geographic localities (Richardson *et al.*, 1986).

Molecular genetics has proven useful in assessing the extent and patterns of population subdivision as well as for investigation of the forces that change population structure (Avise, 1994; Hillis *et al.*, 1996a). The general advantages of molecular markers for the study of population structure are their ready availability, genetic simplicity, comparability across taxa, and ease of use with population genetic models (Johnson, 2000). Various molecular genetic techniques have been employed to address population related issues, including allozyme electrophoresis (Avise, 1994; Richardson *et al.*, 1986) and DNA-based markers, e.g. nucleotide data (Hillis *et al.*, 1996a) and microsatellite DNA (Queller *et al.*, 1993) which provide a source of highly polymorphic nuclear genes for the study of fine-scale population structure.

Evaluation of the amount of genetic variation within and between populations is one of the main tasks of population genetics. Allele frequencies provide a very simple description of the amount of genetic variation in a populations in the case when there are only two alleles at a locus (Page, Holmes, 1998). However, because many techniques generate multiple-allelic data, dealing with their frequencies becomes cumbersome. Therefore, a more useful measure of genetic diversity is frequency of heterozygous individuals in the population, otherwise known as heterozygosity (Nei, 1978). It is also possible to calculate the average heterozygosity (H) across all loci, which can also be thought of as the average fraction of heterozygous individuals per locus.

Once allele frequencies and heterozygosities have been calculated, the next task commonly undertaken is to determine whether populations are under Hardy-

Weinberg equilibrium. This can be performed by testing the null hypothesis of equal observed and expected numbers of genotype within each population by using χ^2 goodness-of-fit tests. When sample size is small, G tests or Fisher's exact tests are more appropriate (Zar, 1984). If the null hypothesis is rejected, at least one of the following assumptions of Hardy-Weinberg equilibrium breaks down. These are: (i) diploid, (ii) sexual reproduction, (iii) random mating, (iv) infinitely large population, (v) no mutation, (vi) no natural selection and (vii) no migration (Richardson *et al.*, 1986). Looking for deviations from what is predicted by Hardy-Weinberg is therefore a good starting point for studying the evolution of genes in population. Several additional ways of assessing genetic diversity within populations are often used. The most commonly used parameters are: (i) average number of alleles per locus, (ii) number of polymorphic loci, (iii) Shannon index (Shannon, Weaver, 1949), and (iv) gene diversity of Nei (1973) (Yeh *et al.*, 1999).

The major question in population genetic studies is the extent of population subdivision and the significance of isolation, inbreeding and gene flow in shaping genetic structure of populations in the species under study. This can be done by assessing how the observed levels of heterozygosity differ from those expected under the Hardy-Weinberg theorem. The most common strategy is the *F*-statistic approach of Wright (1969). This measure θ, or the fixation index, is considered most important as it estimates the reduction of heterozygosity due to population subdivision. Several other estimation of Wright's (1969) *F*-statistics have been developed (e.g. *F*st, and *G*st in an analysis of molecular variance framework), where the notation "ST" refers to comparisons being made among subpopulations relative to the total population. Other *F*-statistics include f (*F*is), which used to measure the reduction in

heterozygosity of individual relative to their own subpopulation due to local inbreeding, and F (*F*it) which measures the effect of both inbreeding and population subdivision (Page, Holmes, 1998). The null hypothesis that the allele frequencies in different populations are not significantly different can be tested by measuring the significance of divergence of the observed data from that predicted by the null hypothesis. This procedure is undertaken by using χ^2 test of homogeneity. For small sample size, Fisher's exact probability test or Yate's correction for continuity can also be adopted (Richardson *et al.*, 1986).

A more general measure of genetic variation between populations is the use of similarity derived from allele frequencies. This reduces relationships among samples to pairwise distances derived from all sampled loci in the form of a single matrix. A number of algorithms have been developed for calculating genetic distance but the most commonly used are probably those that have been implemented in the most popular software packages (Page, Holmes, 1998).

Genetic relationships are most commonly inferred from a pairwise distance matrix subject to one of a number of clustering methods. However, these methods sometimes fail to represent relationships accurately, especially those involving samples that are genetically intermediate. Alternatively, the methods such as Principle Component Analysis (PCA) or Multidimentional Scaling (MDS) that allow the examination of the proximity of samples in multivariate space can be effective in this context (Kohlmann *et al.*, 2005; Winans *et al.*, 2004).

Mitochondrial DNA data have also been widely used to investigate genetic population structure (Apte *et al.*, 2003; Duran *et al.*, 2004; Gross *et al.*, 2002). The *F*st method is generally used to estimate differentiation among populations, while the number of haplotypes, halotype frequencies, haplotype diversity, and nucleotide diversity are used to evaluate within population genetic variation.

General applications of population genetics to aquaculture species are discussed in section (1.2.2.2). A special and increasingly important application concerns the study of the escape of domesticated fish species and their mixing and interaction with conspecific wild populations.

Hatchery or farmed raised and wild fish can differ in many traits. When fish are removed from the natural environment and placed in the culture environment gene frequencies are altered and a reduction of genetic variation occurs especially after a number of generations. The process of domestication reduces genetic variation in fish through both unintentional selection processes and random genetic drift (Koljonen, 1989; Primmer *et al.*, 1999). Breeding programs often aim to produce specific phenotypes though intensive artificial selection (Ferguson, 1995b) add to the differentiation of domestic stock. Translocation and dissemination of domesticated breeds adds further to the potential genetic differences between farmed and wild populations at a particular location.

Through flooding or other mechanisms domesticated fish often escape, leading to the mixing of domesticated and wild populations. The genetic impact of exogenous fish on wild populations are summarized by Hindar *et al.* (1991) and include:

interbreeding, introgression, and competition. In general, the consequences of the mixing of domestic and wild population is not well understood. If domesticated fish do not survive or reproduce or if their progeny do not survive in the wild, no effect on gene frequencies of the wild populations will be observed. Alternatively, if large numbers of domesticated fish survive and reproduce relative to wild fish, much of the genetic variation of wild population may be lost due to genetic swamping (Cross *et al.*, 2005; Koljonen, 1989). If both domestic and wild fish reproduce, genetic variation of the fish in the affected natural environment may increase, but it may be maladaptive (Dunham, 2004). Empirical studies of the interaction of hatchery raised or domesticated fish and wild populations include Atlantic salmon (Clifford *et al.*, 1998; Hansen *et al.*, 2000) and shrimp (Xu *et al.*, 2001).

1.2.2.2 Aquaculture

Molecular genetic approaches began to be used in fish genetic research in the 1950's (Ward, Grewe, 1995). The development of DNA-based genetic markers has had revolutionary impact on animal genetics in general, including a large number of commercial fish species (Liu, Cordes, 2004). The popular genetic markers in aquaculture genetic research include allozymes, mitochondrial DNA (mtDNA), Restriction Fragment Length Polymorphism (RFLPs), Random Amplified Polymorphism DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), microsatellites (SSR), Single Nucleotide Polymorphism (SNP) and Expressed Sequence Tag (ESTs) (Liu, Cordes, 2004) (Table 1.1). The application of DNA markers has allowed significant progress in aquaculture investigation of genetic variability, inbreeding, parentage assignments, species and strain identification and the construction of high resolution genetic linkage maps for aquaculture species

(Davis, Hetzel, 2000; Ferguson, Danzmann, 1998; Liu, Cordes, 2004; Ward, Grewe, 1995).

Genetic identification of species or strains is often required in an aquaculture setting. This is because it is important to determine if fish stocks are pure species (or strains) or are hybrids or mixed stocks, a problem seen in common carp (Lehoczky *et al.*, 2005), and tilapia (Bardakci, Skibinski, 1994). RAPD, AFLP, and microsatellite data can provide a rapid solution. For instance, microsatellites were used to detect pure stock of common carp in live gene bank in Hungary (Lehoczky *et al.*, 2005) and strain identification in channel catfish (Waldbieser, Wolters, 1999). The AFLP approach has been utilised to identify strains of common carp (David *et al.*, 2001) and channel catfish (Mickett *et al.*, 2003).

Molecular genetic markers are extremely useful for monitoring levels of genetic variation within and between populations of aquaculture species. In the past, allozyme and mtDNA have been most frequently used in fishes but the more recently developed microsatellite, AFLP and RAPD teachniques are now more commonly used (Liu, Cordes, 2004). The microsatellite method has been used to investigate genetic diversity of black bream (Jeong *et al.*, 2003), tiger shrimp (Xu *et al.*, 2001), tilapia (Rutten *et al.*, 2004), and oyster (Yu, Guo, 2004). The AFLP approach has been successfully used to evaluate genetic variation among channel catfish populations (Simmons *et al.*, 2006), while the RAPD approach has been used to estimate genetic structure of black tiger shrimp (Tassanakajon *et al.*, 1998).

With the advent of powerful genetic markers and an emerging mathematical framework to calculate parentage, it is now possible to analyse relatedness and inheritance in aquaculture species with a high degree of precision (Hastein *et al.*, 2001). Microsatellites provide the best result in parentage assignment tests due to high levels of polymorphism. For example, paternity and reproductive contribution in natural populations of bluegill sunfish were determined successfully by Neff (2001) using 11 microsatellite loci. The paternal and maternal genotype of squid have been reconstructed successfully using five microsatellite loci (Emery *et al.*, 2001).

Recently, one of the major benefits of molecular genetic techniques is their integration with traditional breeding methods for faster genetic improvement in aquaculture species. One very powerful application of new DNA-based technologies is to identify nuclear marker loci which are associated with loci that control economically important traits such as body weight (Reid *et al.*, 2005), disease resistance (Rodriguez *et al.*, 2004) cold tolerance (Sun, Liang, 2004), and salinity tolerance (Lee, 2003). The identification of quantitative trait loci and directly associated or linked molecular markers for them facilitate marker-assistance selection (MAS) in aquaculture species (Davis, Hetzel, 2000; Dunham, 2004; Lande, Thompson, 1990). Microsatellite and AFLP markers are the most reliable, efficient and abundant markers for detailed genetic linkage mapping in catfish (Liu *et al.*, 2003), and scallop (Li *et al.*, 2005). Currently, medium density framework linkage maps are available for salmon (Stein *et al.*, 2001), rainbow trout (Sakamoto *et al.*, 2000; Young *et al.*, 1998), brown trout (Gharbi *et al.*, 2006), catfish (Liu *et al.*, 2003), tilapia (Agresti *et al.*, 2000; Kocher *et al.*, 1998; McConnell *et al.*, 2000), oyster (Li, Guo, 2004), shrimp (Wilson *et al.*, 2002), and common carp (Sun, Liang, 2004).

Where fish breeding is targeting multiple traits together it is usually found that family selection (where the best performing families are selected as parents) or modifications thereof, is more effective than mass selection (where the best performing individuals are selected) (Cross *et al.*, 2003). Family selection, however, means that progeny must be identifiable as to family throughout life. Marking by physical methods is usually feasible in later stages, but is often impossible with fry. Thus large numbers of individual family tanks must be maintained until the fish are old enough to mark and mix. Usually 50-100 tanks per select strain are required, with a similar number of unselected control families. The reason for using so many selected families, is to allow sufficiently strong selection to achieve substantial genetic gain (the increase in the mean value of the target trait/s per generation) while, at the same time, minimizing inbreeding.

The use of molecular markers as genetic tags to achieve family identification allows for communal rearing throughout the life cycle including the juvenile stage. Moreover, these markers can be used to identify individual and family groups so that they can be reared together, thus simplifying experimental designs. This approach also reduces the environmental variation between tanks, which can reduce the veracity of the selection program. Furthermore, as less physical space is required, more families can be used allowing the application of stronger selection, while preventing inbreeding (Cross *et al.*, 2003).

1.2.2.3 Conservation genetics

Conservation genetics is the application of genetics to preserve species as dynamic entities that can evolve to cope with environmental changes and thus minimize their risk of extinction (Frankham, 2003). Many factors impact negatively upon local biodiversity including over-exploitation of natural resources, habitat destruction, pollution, and the introduction of exotic species or non-native strains causing competition or interbreeding with resident species (Sala *et al.*, 2000).

The loss of genetic variation is not only an issue with cultured stocks but is probably more significant for wild populations. Without retaining genetic diversity, the cost and effort expended in breeding programs for targeted species may be lost, as stocked individuals into the natural environment may fail to survive and reproduce due to reduced fitness (Waples, 1991). Utilising selection for genetic diversity, aquaculture breeding programs have and continue to be established for threatened and endangered species, with the aim of restocking the natural environment (Cross *et al.*, 2003).

A large amount of molecular data have been applied in recent years to address questions related to conservation genetics, including genetic diversity, population isolation, divergence and interaction between cultured and wild populations using a range of molecular markers (Bouza *et al.*, 1997; Hansen *et al.*, 2000; Mjolnerod *et al.*, 1997; Simmons *et al.*, 2006; Xu *et al.*, 2001). Molecular markers are also used to identify species and investigate cryptic speciation (see 1.2.2.2) essential for the management and conservation of species.

1.2.2.4 Taxonomy and evolution

Molecular genetic information is now increasingly being used to resolve taxonomic issues including the delimitation of species boundaries in morphologically variable or conservative groups of organisms (Hillis *et al.*, 1996a). In fact, some scientists have proposed that genetic information in the form of sequences of the mitochondrial DNA (mtDNA) COI gene region should be the primary information in taxonomic classification and identification. However, the DNA-Barcoding approach to biological classification has both strong advocates and opponents (Hebert *et al.*, 2003; Miya, Nishida, 2000; Schander, Willassen, 2005).

Irrespective of one's position on DNA barcoding it is widely agreed that taxonomic classification at all levels should reflect evolutionary relationships (Blaxter, 2004; Hebert *et al.*, 2003; Schuh, 2000). Further, it is considered essential that any taxonomically recognized group of organisms should be monophyletic on the basis of rigorous phylogenetic analyses. While this has been traditionally achieved though an analysis of morphological data, the availability of nucleotide sequences from a range of genes coupled with powerful methods of phylogenetic estimation using various software packages means that more and more classifications are constructed or tested using molecular data (Steinke *et al.*, 2005). The next section briefly reviews contemporary approaches to the recognition of species using molecular data and the principles of phylogenetic estimation.

One can not satisfactorily study a species or relationships or diversity within and among them without a conceptualisation of what species are (Cracraft, 2000). However, no single issue in biodiversity related studies has been as controversial as

the nature of species and how they should be defined and delimited (Avise, 2000; Claridge *et al.*, 1997; Mayr, 2000; Shaw, 1998; Wheeler, Meier, 2000). The species problem, as it is commonly known, has a large and complex literature dedicated to the topic. This problem is largely due to the inability of biological scientists (systematists and speciation biologists) to agree on what fundamentally constitutes a species, resulting in a proliferation of species concepts, of which there is no less that 24 modern definitions species according to Mayden (1997).

From a practical viewpoint the morphological species concept has been and continues to be the most commonly used approach to defining and identifying species based on phenotypic differences (Blackwelder, 1967). However, this approach has significant conceptual and operational problems when dealing with phenotypically plastic species or morphological conservatism (cryptic speciation) (Mayr, 2002).

Mayr (1963) developed the biological species concept to address these limitations. The biological definition of species, as reproductive communities genetically isolated from other assemblages, became the dominate species concept for most of the last part of the last century.

An advantage of the biological species concept is that for putative species in sympatry, reproductive isolation can be tested by using molecular data to determine genetic relationships. The finding of genetic differences generally provides strong indirect evidence for reproductive isolation and the existence of separate gene pools and the presence of two biological species.

While the biological species concept provided a more rigorous basis for determining species difference, especially when there is access to information from molecular markers, it also has a number of limitations. These include various operational aspects (Sokal, Crovello, 1970), hybridization (Wu, 2001) determination of the status of allopatric populations (Avise, 2000), asexual species and fossils. The development of the phylogenetic approach in biodiversity studies has resulted in a reconceptualisation of what species are and how they can be identified and delimited.

In fact, a confusing diversity of phylogenetic species concepts have now been developed (Avise, 2000; Nixon, Wheeler, 1990; Shaw, 1998; Shaw, 2001; Wheeler, Meier, 2000). Collectively they all emphasise the importance of shared historical relationships and agree that species are lineages with individuals united by genealogy. It is also widely agreed that species satisfying phylogenetic criteria are also reproductive communities (Templeton, 2001) consistent with the Biological Species Concept (BSC) as proposed by Mayr (2000). Because distinct species are reproductively incompatible it allows individuals within each species to establish independent sets of genealogical relationships. The processes of hybridization and gene flow between incipient species will disrupt the acquisition of reproductive isolation and the establishment of species specific gene trees.

When conducting an analysis of genetic relationships within and between closely related species using phylogenetic analysis of nucleotide sequences a phylogenetic or genealogical approach to the recognition of species is a necessity. Operationally, phylogenetic species are identified on the basis of reciprocal monophyly (Avise,

2000; Baum, Shaw, 1995). Importantly, the testing of species boundaries using nucleotide data under this criterion can be placed within a statistical framework using bootstrap and other methods of establishing reliability and hypothesis testing (Wiens, Penkrot, 2002). As a consequence the collection of nucleotide data, most commonly derived from mitochondrial gene regions, allows existing taxonomic hypotheses to be tested using explicit criteria which represents a major advance in taxonomic methodology (Sites, Crandall, 1997; Templeton, 2001; Wiens, Penkrot, 2002). Nevertheless, a potential problems associated with this approach is that individual gene genealogies do not always correspond to species genealogies, especially in the case of recently speciated taxa and the commonly used mt markers represent only one locus. Access to nuclear molecular makers, reliable morphological traits, information on ecological differences or reproductive compatiblity would provide important ancillary data on the biological validity of potential species defined on the basis of a geneaological analysis of one or more mitochondrial gene regions.

Molecular phylogenetics is the study of evolutionary relationships among organisms by using molecular data such as DNA and protein sequences or other molecular markers (Graur, Li, 2000). Phylogenetics is concerned with genealogy and reconstructing evolutionary history of species and higher taxonomic groupings. Relationships among species or groups are depicted in branching diagrams or evolutionary trees that represent historical speciation events, and are now considered to provide essential information for the establishment of reliable classifications (Hillis *et al.*, 1996c). Some of the earliest molecular data to be used in phylogenetic studies were derived from immunological techniques (Maxson, Maxson, 1990; Sarich,

Wilson, 1967) and protein sequencing (Zuckerkandl, Pauling, 1962). Allozyme electrophoresis, until relatively recently the predominant method for assaying genetic variation within and between species, has also been utilised for phylogenetic studies (Avise, 1974; Avise, 1983; Buth, 1984; Hillis *et al.*, 1996b; Richardson *et al.*, 1986). Emphasis, however, has now shifted to methods of DNA analysis, primarily through the use of restriction enzymes and direct sequencing, and a number of PCR-based techniques which allow increased direct access to the phylogenetic information content of DNA sequences from both nuclear and mitochondrial genes (Avise, 1994; Avise, 2000). As a result, an enormous amount of molecular data, especially in the form of nucleotide sequences, is now accumulating and being used to resolve systematic relationships of organisms at various levels, ranging from populations to kingdoms (Hillis *et al.*, 1996b).

The most commonly used methods useful for phylogenetic reconstruction are distance, maximum parsimony, maximum likelihood and Bayesian analysis. Distance methods involve the calculation of a matrix of genetic or evolutionary distance values between all pair of taxa or samples, and constructing phylogeny based on these values (Nei, Kumar, 2000; Page, Holmes, 1998). The calculation of genetic or evolutionary distances is a reflection of the mean number of nucleotide changes per site that have occurred between a pair of sequences since their divergence from a common ancestor. Distance methods are dependent upon the selection of an appropriate model of nucleotide evolution for the calculation of distance values (Posada, Crandall, 1998).

Maximum parsimony is based on the underlying principle that the simplest explanation, involving the least assumptions is always preferable. Parsimony methods were originally developed for the analysis of character evolution by morphological systematists and have since been modified for molecular systematic analyses (Nei, Kumar, 2000). The basic principle for the construction of maximum parsimony trees based on analysis of DNA sequence data is the identification of the phylogeny that requires the smallest number of nucleotide substitution (evolutionary steps) for a given set of data (Nei, Kumar, 2000).

The maximum likelihood approach involves finding the optimal tree that is the most likely to have occurred given the observed data and an assumed model of evolution. The statistical properties of maximum likelihood analysis make this an attractive method for phylogenetic analysis using nucleotide data, because it allows robust statistical tests to be performed (Steel, Penny, 2000). However, the computationally intense nature of likelihood calculations means that, especially for large data sets, it is difficult to ensure that the optimal tree is found.

More recently, Bayesian approaches to phylogenetic inference have been gaining popularity (Reed *et al.*, 2002; Reeder, 2003; Vicario *et al.*, 2003). Bayesian inference is based on the likelihood function and can use the same models of evolution, so therefore has similar statistical properties to the maximum likelihood method. However the Bayesian method differs from others in that it enables prior information about phylogeny to be specified. Bayesian phylogenetic inference utilize Monte Carlo Markov chain methods to explore possible phylogenetic trees (Huelsenbeck *et*

al., 2002). Baysian methods are attractive due to their relative speed, which allows complex and realistic evolutionary models to be examined (Buckley *et al.*, 2002).

It is now considered essential to assess the reliability of tree constructed from nucleotide data using any phylogenetic method (Graur, Li, 2000). Significant difficulties are associated with determining accuracy and reliability of phylogenetic trees because the true tree representing the actual evolutionary relationships between taxa, is rarely known (Page, Charleston, 1997). Bootstrap analysis, congruence, and testing of phylogenetic hypotheses are the most common approaches to reliability estimation in phylogenetic studies (Graur, Li, 2000; Hillis *et al.*, 1996c; Page, Holmes, 1998).

The bootstrap (Felsenstein, 1985) is a computational technique that is frequently used to estimate the confidence level of phylogenetic hypotheses. The technique belongs to a class of methods termed resampling and reuse the nucleotide data from the original data set. Pseudoreplicates (usually 100-1,000) are generated by resampling the nucleotide sites in the data with replacement. Each pseudoreplicate is then used to generate a phylogenetic tree. The frequency at which pseudoreplicate data sets suggest a given relationship is termed “bootstrap support”.

Congruence is the agreement between estimates of phylogeny based on different data sets. Because there is only one true set of evolutionary relationships for a particular group of taxa, congruence among independent data sets provides support that the true set of relationships has been inferred. The probability of recovering the same phylogeny from two independent data sets by chance is extremely small (Page,

Holmes, 1998). Consequently, the use of multiple data set is considered highly advantageous for phylogenetic studies (Hillis *et al.*, 1996a).

Another way of gaining confidence in any resultant trees is to test the probability that they provide a better explanation of the relationship than other possible tree topologies. Often referred to as phylogenetic hypothesis testing (Goldman *et al.*, 2000; Huelsenbeck, Crandall, 1997), this process generally involves testing the optimal tree (derived from phylogenetic analyses) against alternative hypothesised topologies.

There are a number of statistical procedures for testing phylogenetic hypotheses. Non-parametric methods such as Templeton's (1983) test and the KH test (Kishino, Hasegawa, 1989) are not valid for test of a posteriori hypotheses, and should only be used for testing a priori hypotheses where there is no reason to think that one tree is better than the others. These two procedures have been generally criticised due to their conservativeness and have been recently superceded by the SH-test (Shimodaira, Hasegawa, 1999), which is appropriate for testing a posteriori hypotheses.

1.2.2.5 Molecular genetic studies on common carp

Common carp have been the subject of numerous molecular genetic studies due to its aquaculture importance. The foci of these studies have been questions relating to phylogenetics, population structure, taxonomy, aquaculture, and the impact of domestication.

In the past, traditional protein markers had been used to study the genetics of common carp populations on regional levels. For instance, allozyme information is available for domesticated and wild population culture and samples in Hungary (Csizmadia *et al.*, 1995), Czech Republic (Desvignes *et al.*, 2001; Slectova *et al.*, 2002), Uzbekistan (Murakaeva *et al.*, 2003), Italy (Cataudella *et al.*, 1987), Japan (Macaranas *et al.*, 1986), Indonesia (Sumantadinata, Taniguchi, 1990), Israel (Ben-Dom *et al.*, 2000), Estonia (Paaver, Gross, 1991), Poland (Anjum, 1995), Germany (Kohlmann, Kersten, 1999) and Australia (Davis *et al.*, 1999). More recently mitochondrial DNA (Davis *et al.*, 1999; Froufe *et al.*, 2002; Mabuchi *et al.*, 2005; Zhou *et al.*, 2003), microsatellite (Bartfai *et al.*, 2003; David *et al.*, 2001; Desvignes *et al.*, 2001; Lehoczky *et al.*, 2005; Tanck *et al.*, 2000), RADP (Bartfai *et al.*, 2003; Wang, Li, 2004), and AFLP data (David *et al.*, 2001) have been used to examine genetic variation in common carp populations.

To date, there are only few studies of common carp populations across its full geographical range in Eurasia. These studies was not thorough, but nevertheless suggest that carp may be divided into European and Asian groups (Brody *et al.*, 1979; Kohlmann *et al.*, 2003; Kohlmann *et al.*, 2005); although this interpretation is not consistent with the actual data presented in some studies (Kohlmann *et al.*, 2005).

The complete mitochondrial DNA nucleotide sequence of common carp was described by Chang *et al.*(1994) (GenBank accession number: X61010). This allowed the design of primers for sequencing mitochondrial DNA genes and fragments to investigate relationships of common carp strains or populations. For example, the molecular phylogeny of three subspecies of common carp in China was

analysed by Zhou *et al*. (2004b) using sequence of Cytochrome *b* (*Cyt b*) and control (*CR*) gene regions. Mabuchi *et al*. (2005) discovered a genetically divergent form of common carps from Lake Biwa, using mitochondrial DNA *CR* sequence data. Phylogenetic relationships of ornamental (koi) carp, Oujiang color carp and Long-fin carp were analysed by Wang and Li (2004) using mitochondrial DNA Cytochrome c oxidase subunit II (*COII*) gene sequences.

Currently there are several published studies of common carp involved the use of microsatellites (Crooijmans *et al.*, 1997; Sun, Liang, 2004; Yu, Guo, 2004). Thirty two microsatellite markers of poly (CA) type in common carp were described by Crooijmans *et al*. (1997). While these authors stated that these loci will be valuable as genetic markers for use in population genetic, breeding and evolution studies, they did not present any population genetic analyses (Kohlmann *et al.*, 2005; Lehoczky *et al.*, 2005; Zhou *et al.*, 2004a). Kohlmann *et al.* (2003; 2005) have presented the only population genetic studies of common carp using four microsatellite loci and found significantly greater variation than was apparent from allozyme studies.

Studies of QTLs in common carp are limited. There is only one gene linkage map of common carp which mapped loci associated with cold tolerance (Sun, Liang, 2004), using the segregation of 272 markers, including 105 gene markers, 110 microsatellites, and 57 RAPD markers. Despite the large number of genetic studies of common carp, none have addressed genetic diversity and genealogical relationships on a global basis. Taxonomy and origin of carp are questionable and require detailed investigation. In addition, the study of common carp from certain regions have been neglected, including Vietnam (south East Asia), which is home to indigenous

common carp that have not been examined using modern molecular genetic methods to any significant extent.

1.3 Research objectives and thesis format

The overall objective of this thesis is to advance the understanding of the genetics of common carp with a special focus on Vietnamese domesticated and wild populations. Following from this objective there are three major components to this study. First, genetic diversity and relationships among common carp stocks and strains are examined on a global scale using a phylogenetic analysis of mitochondrial DNA sequences. By using samples representing major common carp strains obtained throughout the species geographical range a global genealogy of common carp is presented. A phylogenetic representation of nucleotide variation amongst populations of common carp provides a new way of describing genetic variation within carp stocks from different countries, establishing relationships among strains and suggesting their possible origin. This phylogenetic study is then extended by using more conserved gene regions to address questions relating to the taxonomy and evolution of common carp. Second, this study quantifies genetic variation within and between populations of Vietnamese common carp in detail using samples collected from both the wild and hatcheries. The degree of strain mixing and the identification of inbred populations is investigated using the Single Strand Conformation Polymorphisms (SSCP) and microsatellite techniques. Lastly, an investigation of phylogenetic relationships among a range of cyprinid species is undertaken to assess the higher level taxonomic classification of this family using little known Vietnamese species. This study also allowed an examination of the level of diversity within a range of cyprinid genera for comparison with *Cyprinus*.

More specifically in relation to the format for this thesis, the first research chapter of this thesis (Chapter 2) investigates diversity and genealogical relationships within common carp, using sequences from the mitochondrial *CR* and *ATPase6/ATPase8* gene regions for samples of common carp representing a diversity of domesticated strains from around the world. Chapter 3 extends this study using mitochondrial DNA *CR* and *Cyt b* gene region sequences to examine taxonomic questions and evolutionary hypotheses for common carp. This chapter analysed samples from Vietnam, China and Europe and utilised common carp sequences available on GenBank, focusing wherever possible, on specimens originating from wild populations.

Chapter 4 presents an examination of genetic diversity in domesticated and wild population common carp from throughout Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA *CR*. In Chapter 5, the examination of common carp is extended by examining variation in the same populations using four highly variable microsatellite loci.

Finally, Chapter 6 extends the taxonomic and evolutionary question in Chapter 3 by presenting a phylogenetic based testing of the taxonomic classification and relationships within subfamilies of the family Cyprinidae, including the cyprininae, the subfamily of common carp, using three gene regions (*CR*, *16S*, *Cyt b*). This is the first time that the relationships and taxonomic classification of Vietnamese cyprinids has been investigated using molecular genetic data.

Table 1.1. Type of DNA marker, their characteristics, potential applications and cost*

Marker type	Mode of inheritance	Locus under	Polymorphism	Major applications	Relative cost
Allozyme	Mendelian, codominant	Single	Low	Linkage mapping, population studies	Low
Mitochondrial DNA (mtDNA)	Maternal inheritance	Single	High	Maternal lineage	Moderate-high
Restriction fragment length polymorphism (RFLP)	Mendelian, codominant	Single	Low	Linkage mapping	Low
Microsatellites (SSR)	Mendelian, codominant	Single	High	Linkage mapping, population studies, paternity analysis	High
Random amplified polymorphism DNA (RAPD)	Mendelian, codominant	Multiple	Intermediate	Linkage mapping	Moderate
Amplified fragment length polymorphisms (AFLP)	Mendelian, codominant	Multiple	High	Linkage mapping, population studies	Moderate
Expressed sequence tags (EST)	Mendelian, codominant	Single	Low	Linkage mapping, physical map	High
Single nucleotide polymorphism (SNP)	Mendelian, codominant	Single	High	Linkage mapping	High

* The table was modified from Liu and Cordes (2004)

Chapter 2.

Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.)*

2.1 Introduction

Direct DNA sequencing is being increasingly used to investigate the taxonomy, genealogical relationships, origins and diversity of domesticated animal species including: rabbits, pigs, goats and buffaloes (Kierstein *et al.*, 2004; Kim *et al.*, 2002; Long *et al.*, 2003; Manjunath *et al.*, 2004). In contrast, similar studies of aquaculture species are rare (Nguyen *et al.*, 2004), perhaps reflecting their more limited relevance due to the short history of domestication for such species. A notable exception is the common carp (*Cyprinus carpio*. L), the world's oldest domesticated and the most important aquaculture species (FAO, 2003). The domestication of common carp commenced over 4,000 years ago in China (Hollebecq, Haffray, 1999), and now nearly three million tonnes are produced annually worldwide (FAO, 2003).

---

* Two peer reviewed publications have been derivered from this chapter: Thai, T. B., C. P. Burridge, T. A. Pham and C.M. Austin (2005). Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.). *Animal Genetics* 36: 23-28, and Thai, T. B., T. A. Pham, U. D. Thai, C.M. Austin (2006) Progress towards a global genealogy of common carp (*Cyprinus carpio* L.) strains using mitochondrial nucleotide sequence data. *NAGA* 29 (3&4): 55-61.

The species is cultured in many parts of the world both within and outside its natural range as a result of wide-scale translocations. The species is phenotypically plastic, which has led to an extensive and confusing taxonomic nomenclature (Hollebecq, Haffray, 1999). Considering only relatively recent taxonomic literature, Kirpichnikov (1967) recognised four subspecies of carp: *Cyprinus carpio carpio* (Europe), *C. c. aralensis* (Central Asia), *C. c. haematopterus* (East Asia), and *C. c. viridiviolaceus*. (South East Asia). In contrast, Balon (1995) considered only two subspecies were worth recognising: *C. c. carpio* (Europe) and *C. c. haematopterus* (East Asia). Subsequently, Kirpichnikov (1999) questioned the validly of *C. c. viridiviolaceus* where as Li *et al*. (2001) recognised four morphologically distinctive red carp strains from China, which he refers to as *C. c. xingguonensis, C. c. wannanensis, C. c. wuyanensis,* and *C. c. color*. In one of the most recent taxonomic treatments, Kottelat (2001) considers the common cultured carp in southeast Asia to be a distinct species, *C. rubrofuscus*, although this is disputed by Nguyen and Ngo (2001) who consider this species to be quite rare.

A wide range of molecular marker systems have been used for the study of carp including microsatellites (Crooijmans et al., 1997), RAPDs (Bartfai *et al.*, 2003), AFLPs (David *et al.*, 2001), allozymes (Kohlmann, Kersten, 1999), RFLPs (Kohlmann *et al.*, 2003; Zhou *et al.*, 2003) and direct sequencing of mtDNA fragments (Froufe *et al.*, 2002; Wang, Li, 2004). Techniques using mtDNA have been widely employed for aquaculture and fisheries related genetic studies because this molecule has several useful characteristics including a rapid rate of mutation, making it effective for detecting recent population isolation (Ward,

Grewe, 1995) and for establishing genealogical relationships among populations within species (Avise, 2000).

This study was initiated to investigate genetic diversity and genealogical relationships of common carp strains and the usefulness of two rapidly evolving mtDNA regions, the *CR* region and the *ATPase6/ATPase8* genes, for these purposes. This paper presents mtDNA nucleotide sequences exceeding 1,500 bp for each of 87 individual carp representing samples obtained from throughout the world making it the most globally comprehensive set of DNA sequences accumulated so far for this species.

2.2 Materials and Methods

2.2.1 Common carp samples

Common carp tissue samples were acquired from populations or strains in Vietnam, Czech Republic, Hungary, Israel, China, India, Indonesia, Japan and Australia. Tissue samples were taken as fin clips and preserved in 90 % ethanol. Details of sampling locations and codes are provided in Table 2.1

2.2.2 DNA extraction and Amplification of Mitochondrial DNA

Total DNA was extracted as described by Crandall *et al*. (1999). Between two and six individuals were analysed from each population or strain. The mitochondrial *ATPase6/ATPase8* gene fragment was amplified using primers L8331 (5’ AAA GCR TTR GCC TTT TAA GC 3’) and H9236 (5’ GTT AGT GGT CAK GGG

CTT GGR TC 3') (Lovette *et al.*, 1998). Primers for the *CR* were designed from common carp sequences on GenBank (AC: X61010): Carp-Pro (5' AAC TCT CAC CCC TGG CTA CCA AAG 3'), and Carp-Phe (5' CTA GGA CTC ATC TTA GCA TCT TCA GTG 3'). PCR was carried out in 50 µl reaction volumes (1 X reaction buffer, 2 mM dNTP, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 0.5 units *Taq* polymerase, and approximately 200 ng DNA template). Thermal cycling comprised 95 ^{o}C for 3 min, followed by 34 cycles of 95 ^{o}C for 30s, annealing at 55 ^{o}C for 30s, and an extension temperature of 72 ^{o}C for 1 min. This was then followed by a final extension of 72 ^{o}C for 3 min. PCR products were purified using the QIAGEN QIAquick PCR purification kit, sequencing followed ABI PRISM BigDye Terminator protocols. For each individual, sequencing reactions were performed using both primers.

2.2.3 Data Analysis and Phylogenetic Reconstruction

Sequences of carp from GenBank (Chang *et al.*, 1994) were included in the data set for comparative purposes, and *Carassius auratus* (GenBank Accession NC_002079) was used as the outgroup.

Control DNA sequences were aligned using the program Clustal X (Thompson *et al.*, 1997), while the *ATPase6/ATPase8* sequences were aligned by eye. Information content of each gene region was estimated using the method of Hillis and Hillis (1992) based on the *g1* statistic. Sequences were imported into PAUP* 4.0b.10 (Swofford, 2000) for phylogenetic analysis. The most suitable model of evolution for the gene regions was obtained using Modeltest 3.06 (Posada, Crandall, 1998). This model was used to calculate pairwise sequence distances for

the construction of a neighbour-joining tree (NJ). Unweighted parsimony analyses were carried out using the heuristic search option. Confidence levels in the resulting relationship were assessed using the bootstrap procedure with 1,000 replications for NJ and maximum parsimony (MP). Nucleotide (π) and Haplotype diversity (h) were calculated using DNASP 3.51 (Rozas et al., 2003).

2.3 Results

All sequences obtained in this study have been submitted to GenBank (accession numbers: AY 597942-AY 597985 and AY 600150-AY600241). A total of 857 bp of the *ATPase6/ATPase8* genes and 745 bp of the *CR* were obtained for all individuals. The level of nucleotide variation at both gene regions was low, although the *CR* exhibited more variation than the *ATPase6/ATPase8*. Excluding the GenBank *Cyprinus carpio* which is of unknown origin, the two most divergent *ATPase6/ATPase8* haplotypes differed by 10 bp and the two most divergent *CR* haplotypes differed by 14 bp. The *CR* and *ATPase6/ATPase8* fragments yielded 29 and 17 parsimony informative sites respectively, and despite this relatively limited information, the two gene regions possessed significant phylogenetic signal (P <0.01) (Hillis, Huelsenbeck, 1992). The two fragments were largely consistent in the patterns of variability within and relationships among samples.

Nucleotide diversity is low overall (0.000-0.009) but differences are apparent between carp from Asian countries (mean $\pi = 0.007$) and carp of European origin ($\pi = 0.000$). These differences were most apparent in terms of haplotype diversity

(h), which exceeded 0.80 for all carp from South East Asian countries but was 0.00 for European carp.

The genealogical analysis (Fig. 2.1) does not support a fundamental dichotomy between European and Asian carp but instead reveals there are several divergent haplotypes in carp from China and Indonesia. While the European carp are genetically distinct they are placed in a relatively derived position as the sister lineage to a lineage containing native white and red Vietnamese, Koi, Chinese Color and Indonesian yellow carps.

Koi carp samples were invariant and identical to the Chinese Color (CL) carp sample and this haplotype shows minimal divergence from Vietnamese carp (Fig. 2.1). Two of the Indonesian carp strains are very similar (Widan- WI and Rajadanu- RJ) and the third strain (Majalaya- MA1&MA2) shows two divergent haplotypes; one is similar to the Widan/Rajadanu strains, while the other is not related to any of the other haplotypes. Furthermore, the two closely related haplotypes found in the Indonesian strain maintained at the RIA1, Vietnam, are quite divergent from all haplotypes obtained from the other three Indonesian carp strains. While haplotype diversity was high among Vietnamese carp, nucleotide diversity was low with all indigenous samples forming a single lineage with Koi and Chinese Color carp.

Chinese carp showed considerable haplotype diversity and divergence. While no differences were detected between the Xingguonensis (XI) and Wananensis (WN) strains, their haplotype is highly differentiated from Chinese Color carp (CL) and

Wuyuanensis (WU) carp; the latter possesses the most distinctive common carp haplotype found so far. Lastly, the sample of Amur River carp (WA1 and WA2 maintained in India) is interesting as it contains two haplotypes, one close to Chinese Color carp and therefore also Vietnamese carp, and one haplotype that is identical to European carp.

2.4 Discussion

This study demonstrated that direct sequencing of variable mtDNA fragments has the potential to provide useful insights into the genetic diversity, origin, divergence and genealogy of carp strains and populations. A surprising result was the low levels of divergence observed for what are generally considered to be the most rapidly evolving vertebrate mtDNA fragments, the *ATPase6/ATPase8* genes and the control region (*CR*) (Verspoor, 1998), especially for a freshwater species with such an extensive distribution across Eurasia. A second surprising result was the complete absence of variation within two major strains of carp, the European common carp and Koi carp. These findings are consistent with the long history of domestication of this species which has undoubtedly involved significant founder and population bottleneck events leading to localised loss of genetic variation (Balon, 1995).

The origin and relationships of a number of forms of common carp is uncertain. This is the case for Koi carp (Balon, 1995) with speculation that they have been developed directly from carp stocks in Japan, Chinese Color carp or German pond carp (Balon, 1995; Wang, Li, 2004). The results of this study suggest that Koi carp have originated from Chinese Color carp, which have a history of

domestication that can be traced back over 1,200 years (Wang, Li, 2004). Our finding is consistent with the results of Wang and Li. (2004) based on mtDNA *COII* sequence data, but contrary to those of Froufe *et al*. (2002) who sequenced 540 bp of the *CR*, and found that their sample of Japanese Koi (N = 4) had a haplotype identical to European carp. This anomaly could be due to mixing of stocks so that their line of Koi carp had acquired the European haplotype through crossbreeding. Despite their substantial color polymorphisms, only a single haplotype was detected among the 12 individual Koi carp (representing 6 strains) sequenced. This result is consistent with Koi - Chinese Color carp having passed through a substantial bottleneck during the domestication process.

The genetic makeup of Amur River carp represent an intriguing anomaly in that while they occur over 5,000 km from northern Vietnam and 10,000 km from Central Europe, they possess one haplotype that differs by only two bases from a haplotype found at Tuyen Quang in northern Vietnam, and one haplotype identical to the European carp. As the Amur River carp examined were originally derived from a live gene bank line in the Fish Culture Research Institute Szarvas, Hungary, presently maintained at Karnataka in India, this finding might be due to accidental stock mixing in captivity. However, Froufe *et al*. (2002) found a similar result for a sample of Amur River carp obtained from the wild, with some haplotypes being very close to European carp and others to Asian carp. They interpreted this finding to indicate an Asian origin for European carp. However the converse possibility, that European carp have been introduced to China, has also been proposed (Balon, 1995) and is also consistent with the data.

Carp from Europe all possessed the same haplotype and were divergent from all other haplotypes detected. Nevertheless the degree of divergence from their most closely related Asian strains is remarkably small (1%) given that the *CR* and *ATPase6/ATPase8* genes are the two most highly variable regions within the mtDNA genome (Hurt *et al.*, 2001). A remarkable feature of European carp samples is the complete absence of any variation in *CR* and *ATPase6/ATPase8* regions, which like the Koi carp, suggests a history of founder effects and small effective population size associated with translocation and domestication.

Molecular genetic information, including mtDNA sequence data is very useful for clarifying species boundaries (Avise, 1994; Avise, 2000) and so has considerable potential for resolving the confusion and controversy over the taxonomy of common carp (Kirpitchnikov, 1967; Kottelat, 2001; Nguyen, Ngo, 2001). Studies using allozyme, microsatellite, and RFLP analyses of the mtDNA *ND-3/4* and *ND-5/6* regions by Kohlmann *et al*. (2003) were interpreted as providing support for Balon's (1995) position for the existence of two subspecies, *C. c. carpio* and *C. c. haematopterus*, as the Asian and European samples formed two distinct clades. While the results of this study are largely consistent with the results of the Kohlmann *et al*. (2003) study in regard to the distinctiveness of the European carp from Vietnamese/Amur carp/Koi carp. The finding that samples of carp from China or of Chinese origin to be highly divergent and polyphyletic does not support an independent origin for western and eastern common carp lineages and their recognition as distinct subspecies (or species). It is clear that further taxonomic studies on common carp are required and these would benefit from the

parallel investigation of morphological and molecular variation and reproductive relationships.

The findings of this study indicate that the establishment of genealogical relationships among carp strains based on mtDNA sequencing can provide new and useful information regarding the history and development of carp strains. For example, this study has shown that Koi, Chinese, Indonesian yellow and Vietnamese white carp are all closely related and that the other Indonesian and some of the Chinese common carp are highly divergent (e.g Wuyuanensis). Further, this approach can reveal carp stocks of mixed origin (e.g. Majalaya) and others that appear to have been through genetic bottlenecks or subject to low effective population size (e.g. Koi and European carp). Information on relationships among strains can be used as a guide to the effective mixing of strains as part of genetic improvement programs (Ferguson, 1995b). It can also provide information on the historical origins of strains and if strains represent a composite of different strains and also if strains have lost variation during the process of domestication.

Table 2.1. Sample code, collecting locality and number of individuals sequenced (N) for each strain or population.

Strains or population	Sample	Locality	N
Xingguonensis	XI	Jaing xi China	3
Wananensis	WN	Jaing xi China	3
Wuyuanensis	WU	Jaing xi China	3
Color	CL	Jaing xi China	3
Dor 70	D70	Gan-Shmuel, Israel	4
Nasice	NA	Gan-Shmuel Israel	4
White Koi	WK	Gan-Shmuel, Israel	2
Black Koi	BAK	Gan-Shmuel, Israel	2
Kohaku Koi	KK	Gan-Shmuel, Israel	2
Sanke Koi	SK	Gan-Shmuel, Israel	2
Showa Koi	SHK	Gan-Shmuel, Israel	2
Red Koi	KO	Komaki-shi, Japan	3
Hajduboszomeny	HA	Szarvas, Hungary	5
Szeged	SZ	Szarvas, Hungary	5
Tata	TAT	Szarvas, Hungary	5
Wild Amur	WA	Karnataka, India	3
Hungarian P3	P3	Karnataka, India	2
Blatna	CZ	Bohemia, Czech Republic	4
German X Bangkok	LF	Karnataka, India	2
Bhadra River	LB	Karnataka, India	2
Majalaya	MA	Sukamandi, Indonesia	2
Rajadanu	RJ	Sukamandi, Indonesia	2
Widan	WI	Sukamandi, Indonesia	2
Yen Bai	YB	Yen Bai,Vietnam	2
Tuyen Quang	TQ	Tuyen Quang, Vietnam	4
Bac Kan	BK	Bac Kan, Vietnam	4
Vietnamese (RIA 1)	VN	RIA 1, Vietnam	2
Indonesian yellow	I	RIA1, Vietnam	4
Hungarian scale	H	RIA1, Vietnam	2
Rockland Reservoir	AU	Victoria, Australia	2
GenBank (X61010)	GB	Taiwan [1]	1
Goldfish (outgroup)	Goldfish	Unknown	1

[1] Origin of sample not provide

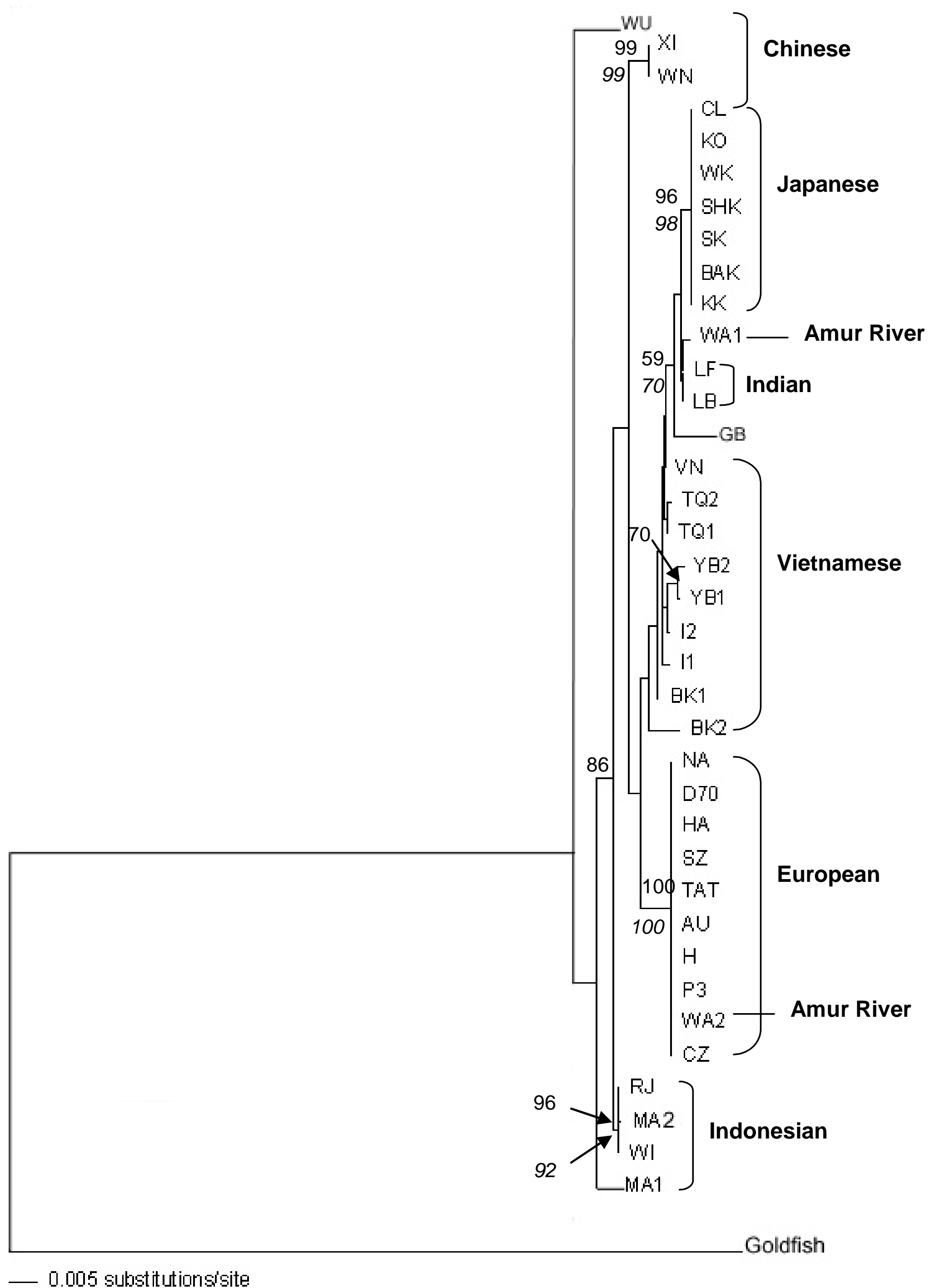

Figure. 2.1. Tree reconstruction derived from *ATPase6/ATPase8* and *CR*, using HKY+I+G model of evolution. Bootstrap values are based on 1,000 replicates, neighbour-joining (normal font) parsimony (*italic*), Bootstrap values are given for nodes with at least 50% or more support.

Chapter 3.

Molecular taxonomy and evolution of common carp

3.1 Introduction

The common carp [*Cyprinus carpio* Linnaeus (1758)] is the oldest domesticated species of fish. Culture of carp in China dates back to at least the 5th century BC, although domestication began later. Wild common carp range from Europe throughout Eurasia to China, Japan and South East Asia (Baruš *et al.*, 2002). However, due to its popularity as an aquaculture and ornamental species, common carp has also been widely translocated, both within and outside its natural distribution. As a result of many transfers and introductions, it is now one of the most widespread species of freshwater fish, with naturally reproducing populations established in many countries. FAO (1998) listed introductions of common carp into 124 countries, of which 81 are recorded as having established populations.

Cyprinus carpio is taxonomically one of the most confusing species (Balon, 2004). Most of the confusion comes from giving quasi taxonomic names to phenotypically unusual specimens or populations or domesticated strains. Nevertheless, some authors consider *Cyprinus* to be a monotypic genus (Lever, 1996), while others consider it to be represented by multiple genera (Baruš *et al.*, 2002; Kottelat, 1997; Nguyen, Ngo, 2001; Zhou, Chu, 1986). An extreme

example being Zhou and Chu (1986) who recorded 12 species from Yanan Province, China, alone. Over 30 synonyms and 10 subspecies of carp were documented by Baruš *et al.* (2002); Kottelat (1997) gave names to 15 subspecies and eight varieties and morphs. While Kottelat (2001), Nguyen and Ngo (2001) reported five different *Cyprinus* species in Vietnam including: *Cyprinus melanes* (Yen), *Cyprinus quidatensis* (Tu), *Cyprinus multitaeniata* (Pellegrin and Chevey), *Cyprinus exophthalmus* (Yen), and *Cyprinus hyperdorsalis* (Hao).

Early studies based on an analysis of morphological differences between European and Amur River carp led to recognition of two subspecies: *Cyprinus carpio carpio*, found throughout Europe, the Caucasus and central Asia, and a far eastern subspecies, *Cyprinus carpio haematopterus*, from the Amur basin and the lakes and rivers of South-Eastern China (Nikolyukin, 1956). It is assumed that the divergence of these subspecies occurred towards the end of the Pleistocene, or more probably during one of the glacial periods of the Pleistocene. Subsequently Kirpitchnikov (1967) suggested four subspecies: European-Transcaucasian common carp (*C. c. carpio*), Central Asian common carp (*C. c. aralensis*), East Asian common carp (*C. c. haematopterus*) and South East Asian common carp (*C. c. viridiviolaceus*). According to Kirpitchnikov (1967) *C. c. viridiviolaceus* is distributed throughout South East China, north Vietnam, Laos and Burma. Next Wu *et al*. (1977) classified *Cyprinus carpio* into three subspecies: *Cyprinus carpio rubrofuscus* (Lacepede) mainly restricted in South of Nanling Mountains in China and Vietnam; *C. c. haematopterus* (Temminck and Schlegel) ranging from Heilongjiang (Amur) River to Nanling Mountainous in China and Japan and *C. c. carpio* (Linnaeus) was mainly distributed from the Danube River to Volga River

and Xinjiang Vygur Autonomous Region of China. Kirpitchnikov (1999) and Baruš *et al*. (2002) asserted that only three subspecies of common carp should be recognized but their classification differed from that of Wu (1977) and consisted of: (1) the European and central Asian common carp, *C. c. carpio*, (2) east Asian common carp, *C. c. haematopterus*, (3) the South East Asian common carp, *C. c. viridiviolaceus*.

In addition, most domesticated varieties of common carp are distinguished and named on the basis of scale pattern, color and body forms, such as scale carp, mirror carp and leather carp. This categorization is done for aquaculture purpose and has no taxonomical basis although Latin names are often used. For example, there are numerous varieties and strains of common carp recognized in China such as the Xingguo Red carp (*C. capio.* var. *xingguonensis*), and "red purse" common carp (*C. capio.* var. *wuyanensis*) (Li, 1999).

Recently, a number of molecular studies of common carp have been conducted and have drawn taxonomic conclusions (Gross *et al.*, 2002; Kohlmann *et al.*, 2005; Mabuchi *et al.*, 2005; Zhou *et al.*, 2004b; Zhou *et al.*, 2003). Kohmann *et al.* (Kohlmann *et al.*, 2003; Kohlmann *et al.*, 2005) using allozyme, RFLP, and microsatellite data, and Gross *et al*. (2002) using RFLP, suggested that only two subspecies should be recognized. In contrast, Zhou *et al*. (2004), using sequence data from the *CR* and *Cyt b* genes region claimed support for morphological subdivision of common carp into three subspecies, following Wu (1977). Thus, despite the increasing application of molecular techniques to the understanding of the diversity of common carp, the resolution of taxonomic issues is no closer.

Further, of the molecular studies so far conducted, none have included comprehensive sampling, with south East Asian common carp being especially poorly represented.

The original geographical distribution and evolutionany origin of common carp is also disputed. Because of the long documented cultivation history of common carp in China, some scientists considered that the ancestor of European domestic carp was derived from Asian common carp stocks, during ancient Greek and Roman periods (Chiba *et al.*, 1966; Vooren, 1972). Others consider that the common carp is indigenous to Europe and in fact postulate a European origin for wild carp and subsequent dispersal east to Siberia and China (Balon, 1995; Kottelat, 1997).

To further examine the taxonomy and evolution of *Cyprinus carpio*, partial sequencing of the mitochondrial DNA (mtDNA) *CR* and *Cyt b* gene region was undertaken. MtDNA sequences, especially from the *CR* and *Cyt b* regions are frequently used for population genetic and molecular phylogenetic studies in fishes (Liu, Chen, 2003; Peng *et al.*, 2004; Perdices *et al.*, 2004). The *CR* is a fast evolving highly variable region in vertebrates and is the most appropriate segment for studies of intraspecific variation in fish (Ferguson, 1995a; Lee *et al.*, 1995), while the *Cyt b* gene is protein coding and evolves more slowly. *Cyt b* sequences are widely used for phylogenetic analyses and is considered to be one the most reliable mitochondrial markers for evolutionary and taxonomic studies (Zardoya, Meyer, 1996).

3.2 Materials and Methods

3.2.1 Sample collection

A total of 17 domesticated and wild common carp samples were obtained as fin clips from China, Vietnam, Hungary, India, Indonesia, Czech Republic and Israel between 2003 and 2004. Tissues samples were preserved in 90% ethanol. The sample origins, codes and taxonomic classifications following different authors are given in Table 3.1.

3.2.2 DNA Extraction and Amplification

DNA was extracted from fin tissue using a high salt precipitation method (Crandall *et al.*, 1999). A fragment of the *Cyt b* mitochondrial gene was amplified by PCR using primers CytF:(5' CGC ATT CCA CTT CCT ACT ACC 3') and CytR: (5' CTA ACC ATC CTG CTA GTC GC 3') which were designed from common carp sequences from GenBank (accession number: X61010). Partial sequences of the *CR* were obtained by PCR amplification using the Carp-Pro and Carp-Phe primers as described by Thai *et al*. (2004). PCR amplification was carried out using the following temperature regime: an initial denaturation step of 95 ^{o}C for 3 minutes, following by 34 cycles of 95 ^{o}C for 30 s, annealing at 55 ^{o}C for 30 s, and extension temperature of 72 ^{o}C for 1 min, followed by an additional extension of 72 ^{o}C for 3 minutes. PCR products were purified using an UltraClean PCRClean-Up DNA Purification Kit (Geneworks). Sequencing reactions were performed using ABI Big Dye Terminator Chemistry with 6 pmol of each primer and 30-50 ng of PCR product. The resultant sequencing reaction products were sent to the Institute of Medical and Veterinary Science (IMVS), South Australia

for sequencing on an ABI 3700 sequencer. Both strands of each fragment were sequenced.

3.2.3 Phylogenetic Analyses

The 34 sequences obtained in this study were supplemented by 50 sequences of common carp obtained from GenBank (Chang *et al.*, 1994; Mabuchi *et al.*, 2005; Thai *et al.*, 2004; Zhou *et al.*, 2004b) for the same mtDNA regions (Table 3.1). *Carassius auratus* (Murakami *et al.*, 1998), (GenBank accession number NC_002079) was used as outgroup species (Table 3.1). Multiple alignments for *CR* and *Cyt b* sequences were performed using Clustal X (Thompson *et al.*, 1997). To test for phylogenetic signal, a *g*1 statistic was calculated using 100,000 random trees as described by Hillis and Huesenbeck (1992).

Three methods of tree building were used to reconstruct phylogenetic relationships: maximum-likelihood (ML) and neighbour-joining (NJ) were implemented using PAUP* 4.0b10 (Swofford, 2000), and Bayesian methods were implemented using MrBayes 3.0 (Huelsenbeck, Ronquist, 2001). The appropriated model of evolution for ML and NJ analyses was obtained via testing alternatives models of evolution using Modeltest (Posada, Crandall, 1998). Heuristic searches were employed for ML analyses with 100 replicates of random sequence additions, whilst non-parametric bootstrapping consisted of 100 replications with 10 random sequence additions. The NJ tree was constructed with distances calculated under the same model of evolution as the ML analysis, with bootstrapping performed using 1,000 replicates.

Bayesian analyses were performed using the same general models identified by Modeltest. Analyses were initiated with random starting trees and run for 1.0×10^6 generations sampling four Markov chains every 100 generations resulting in 10,000 sampled trees. The likelihood scores of the sampled trees were plotted against generation time to ensure that stationarity was reached, trees generated prior to stationarity being reached were discarded as “burn-in” (1,000 trees in this case). Bayesian posterior probabilities of each bipartition, representing the percentage of time each node was recovered were calculated from a 50% majority rule consensus tree from the remaining trees.

3.2.4 Phylogenetic hypothesis testing

To test hypotheses of common carp taxonomic relationships, comparisons were made between optimal trees and those developed when enforcing topological constraints (e.g. monophyly of Asian common carp) using the KH (Kishino, Hasegawa, 1989) and SH tests (Shimodaira, Hasegawa, 1999). A parametric bootstrap was implemented by calculating the likelihood differences between the optimal tree created by PAUP* and each of the null hypothesis trees were based on various hypotheses proposed for the evolution or classification of common carp (constrained trees). This was then compared with likelihood differences of optimal and null hypothesis trees from simulated data sets (created essentially via Monte Carlo simulation of DNA sequence evolution, Rambaut and Grassly, 1997). One hundred replicate data sets, with the same tree topology as the null hypothesis trees, were generated using the same parameter and model of sequence evolution as estimated for the observed data and with the same number of sites.

3.3 Results

3.3.1 Phylogenetic analysis

All common carp sequences have been submitted to GenBank as outlined in Table 3.2. A total of 745 bp of the *CR* and 532 bp of *Cyt b* were obtained from each individual and these sequences were combined with the corresponding sequences available from GenBank for the same regions to complete the dataset for phylogenetic analyses. The *CR* showed more variation than *Cyt b* region and yielded 38 parsimony informative sites compared to 16. Intraspecific sequence divergence for the *CR* ranged from 0.00% to 3.41%, while for the *Cyt b* gene divergence ranged from 0.00 to 2.49%.

The partition homogeneity test did not reject phylogenetic congruence between the mitochondrial genes (P = 0.46, P <0.05), allowing combination of genes for phylogenetic analyses. Tree length frequency distributions were significantly skewed for all taxa ($g1$ = -0.57, P <0.01), supporting the presence of phylogenetic signal. The model selected for the NJ and ML analysis was HKY + G + I which accommodates differing transition/ transversion mutation rates. The mean nucleotide composition was A = 32%; T = 30%; C = 24%; G = 14% for the combined nucleotide fragments, indicating that the *CR* and *Cyt b* is adenine and thymine rich in common carp.

Neighbour-joining, maximum likelihood and Bayesian procedures produced similar trees. The neighbour joining tree is showed in Fig. 3.2. These trees did not recover relationships consistent with any of the existing taxonomic or

evolutionary hypotheses put forward for common carp. Further, *Cyprinus melanes* was almost identical to other Vietnamese wild common carp with the average genetic divergence of this species from other common carp being low (*Cyt b* = 0.31%, *CR* = 0.91%). The most divergent carp haplotype was from Lake Biwa, Japan, which forms the most basal position in the trees, but has not been taxonomically recognized as a distinct form of common carp.

In general, divergence levels were very small, with the exception of comparison with the Lake Biwa sample, giving only partially resolved relationships and trees with very short branches. Nevertheless, some surprising relationships were apparent and well supported. The best defined clade (A) contained all the European carps, with the exception of Russian scattered mirror carp, and included an Amur River haplotype. The Amur River haplotype was identical to the haplotype of Hungarian, Czech Republic and Russian carps from Volga River despite originating in far eastern Asia. The Russian scattered mirror carp were identical to carp from China and Japan that formed another relatively well defined clade (B).

Overall, the Asian carp samples were much more diverse than those of European origin. The Chinese carp samples were distributed over five clades, the Japanese carp over four clades, including the highly divergent lake Biwa clade and the Vietnamese carp were divided over three clades. Other than the highly divergent Lake Biwa clade, the Japanese carp are not distinct and are represented in the same clades as the Chinese carps with the exception of the European clade (A). In contrast, the Vietnamese carp are distinct in that most samples from this country

form a single clade (C). Two Vietnamese carp samples are atypical, one of these is from Bac Kan, a distinct highland environment, and does not show clear affinities with any other carp sample. The other distinct Vietnamese sample is from the Bang Giang River, in the extreme north of the country, which shows a closer relationship with a group of Chinese carp than it does to any of the other Vietnamese samples.

3.3.2 Test of taxonomic and evolutionary hypotheses

The maximum likelihood procedures, using the combined mtDNA data from the two regions, was used to assess three phylogenetic or taxonomic hypotheses in relation to the common carp samples studied. The hypotheses tested were that *Cyprinus carpio*: (a) consists of three subspecies *C. c. rubrofuscus*, *C. c. haematopterus* and *C. c. carpio*, and that *C. c. haematopterus* and *C. c. rubrofuscus* are more closely related to each other than to *C. c. carpio* (Zhou *et al.*, 2004b); (b) consists of three subspecies, the European and central Asian carp (*C. c. carpio*), the East Asian carp (*C. c. haematopterus*) and the south East Asian carp (*C. c. viridiviolaceus*) and that *C. c. haematopterus* and *C. c. viridiviolaceus* are more closely related to each other than to *C. c. carpio* (Baruš *et al.*, 2002); (c) consist of two subspecies, Asian common carp (*C. c. haematopterus*) and European and central Asian carp (*C. c. carpio*) (Kohlmann *et al.*, 2005) each of which form monophyletic lineages. The assignment of samples under each taxonomic hypothesis is given in Table 3.1 and the phylogenetic hypotheses are depicted in Fig. 3.1. The results of KH and SH tests are presented in Table 3.2, from which it can be seen that the hypotheses of Zhou *et al*. (2004b) and Baruš *et al*. (2002), postulating three lineages or taxonomic grouping is rejected. The third

hypothesis of Kohlmann *et al*. (2005), suggesting separate Eastern and Western lineages, has only a slightly larger log-likelihood than the best ML tree and so cannot be rejected using this dataset.

3.4 Discussion

3.4.1 Sequence variation

The results of this study indicate that mtDNA sequence data can provide useful insight into common carp taxonomy and evolution. A surprising outcome was the very limited nucleotide variation, given the wide geographic distribution of common carp and the dataset contained samples from populations that have been recognized as separate species.

Cytochrome *b* sequences have been frequently found to be useful for investigating taxonomic and phylogenetic relationship in fish species (Table 3.3). From Table 3.3, it can be seen that interspecific genetic divergence ranges from 9.4 to 19.0% for this gene which mean the maximum value for common carp sample (2.49%) is typical of intraspecific comparisons. The same pattern is repeated for *CR* sequences. A summary of variation in the *CR* sequence for a range of fish groups is also given in Table 3.3. Interspecific comparisons range from 11.30% to 22.00%, which are much higher than the maximum value in this study (3.4%). Moreover, my early studies also revealed only limited variation in *ATPase6*/*ATPase8* gene regions (Chapter 2) which includes the fastest evolving protein coding mitochondrial gene region in carp (Mabuchi *et al.*, 2006). The finding of low levels of sequence divergence in *CR* and *Cyt b* is consistent with

other studies on *C. carpio* (Froufe *et al.*, 2002; Mabuchi *et al.*, 2006; Mabuchi *et al.*, 2005).

Low divergence in a range of mtDNA gene regions could reflect a shallow evolutionary history or slower rate of mtDNA evolution in carp (Mabuchi *et al.*, 2006). Reproductive compatibility between European and Asian stocks favours a shallow evolutionary history as the most likely explanation; however, data from fast evolving nuclear gene regions would help to distinguish between these alternatives.

3.4.2 Phylogeny and evolutionary history

Despite using mtDNA regions of known phylogenetic utility (*Cyt b*), and the resolution of fine scale relationships at the intraspecific level (*CR*), the phylogenetic trees did not provide well resolved evolutionary relationships. Nevertheless, there is sufficient information content to reject the taxonomic classifications and evolutionary relationships suggested by Baruš *et al*. (2002) and Zhou *et al*. (2004b), but not of Kohlmann *et al*. (2005). While Kohlmann *et al*'s (2005) hypothesis of independent European and Asian lineages cannot be rejected, the geographic distribution of haplotypes among his samples indicated that East Asia is the centre of carp diversity, and that European carp most likely originated by translocation. By far the greatest numbers of haplotypes occur in Asia (26) with only four being found in Europe. Two of the European haplotypes are also found in Asia, 10,000km away. Haplotypes unique to Europe differed by only 1-2 bps from the widespread haplotypes. Kohlmann *et al*'s. (2005) data actually indicated that the highest diversity occurs in Asia. The three most basal branches

in their dendrogram (reproduced in Fig. 3.5) are all Asian samples, a result inconsistent with an independent origin of Asian carp as claimed by the authors. The occurrence of the highest genetic diversity in Asia (China, Japan, and Vietnam) based on both the mtDNA data and Kohlmann *et al*'s. (2005) microsatellite study suggests that Asia is the center of origin of common carp. Additional evidence supporting an Asian origin for common carp is the occurrence of its closest relative *Carassius auratus* in south eastern Asia (Li, 1999).

These two arguments are commonly used to deduce the geographic origin of species, with the most well known examples being the "out of Africa" hypothesis for *Homo sapiens* (Templeton, 2002) based on the geographic location of the highest diversity a nearest relative. The occurrence of two divergent lineages in European carp that are not closely related suggests that European carp were established from translocations. The data is consistent with the translocation of a mixed stock, or two independent translocations. However, it should be noted that the "Chinese" haplotype found in Russian scattle mirror carp may have been acquired though hybridization as the samples used by Zhou *et al*. (2003; 2004b) were obtained from stocks that have been imported and maintained in China. Significantly reduced variation in most European stocks (both wild and domesticated) is consistent with this hypothesis as translocation of domesticated stocks will lead to bottlenecks and founder events. An Asian origin for European carp supports the views of Chiba *et al*. (1966), Zhou *et al*. (2003) and Froufe *et al*. (2002) but contradicts the view of Kohlmann *et al*. (2005) and Balon (1995). However, there is very little data that support the view of Kolhmann *et al*. (2005)

and the results presented in this study are derived from the largest mtDNA nucleotide data set (1277 bp) and the most comprehensive set of samples so far assembled to address this question.

The results of this study also provide new insights into the evolutionary origin of Japanese carp. The ornamental Koi carp has been thought to have originated from domesticated carp. Thai *et al*. (2004) and Wang and Li (2004) presented data suggesting that ornamental Koi most likely originated from Chinese color carp. The more comprehensive sampling used in this study indicate that Koi carp have either been derived multiple times from different stocks or have subsequently been crossbred with a range of domesticated or wild stocks. A number of distinct haplotypes were found from Koi carp but all were very similar to Chinese haplotypes.

The most striking feature of the data set was the finding that an unusual and highly divergent haplotype from Lake Biwa in north Japan is the sister group to all other carp samples. This result is supported by recent research by Mabuchi *et al*. (2006). This finding indicates this is a distinct and relatively ancient form of carp that is of evolutionary and conservation significance, but there has been no suggestion that it should be described as a new taxonomic form, even though it co-exits with common forms of domesticated carp which are stocked routinely into this lake (Mabuchi *et al.*, 2005).

3.4.3 Taxonomy of common carp

While the phylogenetic information content of DNA sequence data used in this study is somewhat limited, certain conclusions relating to the taxonomy of common carp can be drawn. There is clear evidence that taxonomies of common carp based on morphological data are inadequate. The hypotheses of Zhou *et al*. (2004b) supporting morphological data and Baruš *et al*. (2002) are rejected and *Cyprinus melanes* described from the Son River, Vietnam is almost indistinguishable from other Vietnamese wild carp. Conversely the most distinctive form of carp so far discovered is from Lake Biwa, Japan and is not taxonomically recognized or considered morphologically particularly distinct (Mabuchi *et al.*, 2005). These findings are consistent with carp being one of the most phenotypically plastic of freshwater fish species (Hollebecq, Haffray, 1999). Thus it is apparent morphological and molecular variation is uncoupled in carp and it is unwise to undertake taxonomic studies based on morphology without access to biological (reproductive) or molecular information.

The taxonomic status of Lake Biwa carp needs to be established as a matter of urgency, as they may represent a new taxon of common carp and one that is under threat of extinction (Mabuchi *et al*., 2005). This form of carp is only known from this single ancient lake and is threatened by both pollution and crossbreeding with domesticated carp that have been introduced to the lake. If the Lake Biwa carp is a distinct species, the collection of nuclear DNA data would provide evidence of genetic isolation.

While there have been many diverse views on the taxonomy of common carp, a consistent recent trend has been the recognition of discrete Asian and European lineages that are recognised as separate subspecies or even species (Balon, 1995; Kohlmann *et al.*, 2003; Kottelat, 2001; Zhou *et al.*, 2004b). The results of Zhou *et al*. (2004b), for example, using *Cyt b* and *CR* sequences supported a fundamental east-west dichotomy in common carp, however they omitted sequences of South East Asian common carp and so did not have a comprehensive data set (Mabuchi *et al.*, 2006). Further, Kohlmann *et al*. (2005) also claimed that their microsatellite data support an east-west taxonomic split; however, an inspection of their data analyses (reproduced in Fig. 3.5) contradicts this interpretation as the East Asian samples do not actually cluster together as required by this interpretation.

While the data set used in this study did not provide any "prima facie" evidence supporting an east-west taxonomic dichotomy between European and Asian common carp it must be noted that this hypothesis cannot be statistically rejected using the mitochondrial data set used in this study. However it should also be noted that the power to reject the hypothesis is limited by low level of variation in the mtDNA data (Mabuchi *et al.*, 2006), the wide scale translocation of stocks, low sample numbers and possible incomplete sorting of mtDNA lineages in ancestral common carp. Nevertheless, the very low level of haplotype divergence over very substantial geographic distances and the fact the "European clade" is nested within the Asian clade strongly supports an Asian origin for European carp. Further, two independent studies have now found the "European" haplotype in Amur River wild carp (Froufe *et al.*, 2002) and another recent molecular genetic

study argues strongly for an Asian origin of carp based on the finding of a divergent carp haplotype in an ancient Japanese lake (Mabuchi *et al.*, 2006).

3.4.4 Conclusion

The taxonomy of common carp suggested by the results of this study using molecular genetic data is inconsistent with current taxonomic classifications based on morphological data. While the results of this study support an Asian origin of European common carp, the limited variation in the mtDNA gene fragments means that either more mtDNA sequence data are needed or sequences from a rapidly evolving nuclear DNA (e.g. introns) need to be obtained so the taxonomic and evolutionary relationships of Asian and European common carp can be finally resolved with some degree of confidence.

Table 3.1. Sources, status, GenBank accession numbers of strains of common carp (n number of samples sequenced) and taxonomy considered by Baruš *et al*. (2002), Zhou *et al*. (2004b), Kohlmann *et al*. (2005).

Strain or population	Code	Country	Accession number	Reference	Baruš *et al*. (2002) subspecies	Zhou *et al.* (2004b) subspecies	Kohlmann *et al.* (2005) subspecies
Red river common carp	RER (n = 1)	Vietnam	AY597951(CR)+DQ532100	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Son river common carp	SOR (n = 1)	Vietnam	DQ354146(CR)+DQ532101	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Lo river common carp	LOR (n =1)	Vietnam	AY597952(CR)+DQ532102	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Lam river common carp	LAR (n = 1)	Vietnam	DQ354145(CR)+DQ532103	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Bang Giang River carp	BGR (n = 1)	Vietnam	AY597948 (CR)+DQ532104	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Bac Kan carp	BAK (n = 1)	Vietnam	AY597957 (CR)+DQ532105	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Dak Lak carp	DAL (n = 1)	Vietnam	DQ354149(CR)+DQ532106	This study	*C. c. viridiolaceus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Indonesian yellow common carp	IDY (n = 1)	Vietnam	AY597969 (CR)+DQ532107	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Majalaya common carp	MAJ (n = 1)	Indonesia	AY597975 (CR)+DQ532108	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Rajadanu common carp	RAJ (n = 1)	Indonesia	AY597971 (CR)+DQ532109	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Amur River carp	WAR1 (n = 1)	India	AY597946 (CR)+DQ532111	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Amur River carp	WAR2 (n = 1)	India	AY597947 (CR)+DQ532112	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Xingguonensis common carp	XIN (n = 1)	China	AY597942 (CR)+DQ532110	This study	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Big belly carp	BBC1(n =1)	China	AY347303(CR), 347276 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Big belly carp	BBC2 (n = 1)	China	AY347304(CR), 347277 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Purse red carp	PRC1(n = 1)	China	AY347300 (CR), 347278 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Purse red carp	PRC2 (n = 1)	China	AY347301 (CR), 347279 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Qingtian carp	QTC1 (n = 1)	China	AY347297 (CR), 347288 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Qingtian carp	QTC2 (n = 1)	China	AY347296 (CR), 347286 (cyt b)	Zhou *et al* . (2003)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*

Table 3.1. continued..

Strain or population	Code	Country	Accession number	Reference	Baruš *et al*. (2002) subspecies	Zhou *et al.* (2004) subspecies	Kohlmann *et al.* (2002) subspecies
Yangte River wild common carp	YWC1(n = 1)	China	AY345331 (CR), 347291 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Yangte River wild common carp	YWC2 (n = 1)	China	AY345334 (CR), 347281 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Xinggu red carp	XRC1 (n = 1)	China	AY345332 (CR), 347284 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Xinggu red carp	XRC2 (n = 1)	China	AY345335 (CR), 347282 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Yuanjiang River wild carp	SWC1	China	AY347302 (CR), 347280 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Yuanjiang River wild carp	SWC2	China	AY347305 (CR), 347290 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Japanese food carp	JFD (n = 1)	Japan	AB158811 (CR), 158806 (cyt b)	Mabuchi *et al*. (2005)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Japanese koi carp	JKC1 (n = 1)	Japan	AY347298 (CR), 347289 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Japanese koi carp	JKC2 (n = 1)	Japan	AY347299 (CR), 347285 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Japanese koi carp	JKD3 (n = 1)	Japan	AB158811 (CR), 158806 (cyt b)	Zhou *et al*. (2003)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Japanese koi carp	JKD4 (n = 1)	Japan	AB158812 (CR), 158807 (cyt b)	Mabuchi *et al*. (2005)	*C. c. haematopteus*	*C. c. rubrofuscus*	*C. c. haematopteus*
Lake Biwa wild common carp	LBW1 (n = 1)	Japan	AB158808 (CR), 158803 (cyt b)	Mabuchi *et al*. (2005)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Volga River wild common carp	VWC1(n = 1)	Russia	AY345340 (CR), 347294 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Volga River wild common carp	VWC2 (n = 1)	Russia	AY345339 (CR), 347295 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Russian scattered scaled mirror carp	RMC1 (n = 1)	Russia	AY345336 (CR), 347287 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. haematopteus*	*C. c. carpio*
Russian scattered scaled mirror carp	RMC2 (n = 1)	Russia	AY345333 (CR), 347283 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. haematopteus*	*C. c. carpio*
German mirror carp	GMC1 (n = 1)	Germany	AY345337 (CR), 347293 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
German mirror carp	GMC2 (n = 1)	Germany	AY345338 (CR), 347292 (cyt b)	Zhou *et al*. (2003)	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Balatna common carp	CZE (n = 1)	Czech Republich	AY597980 (CR)+DQ532113	This study	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Hungarian common carp	HUS (n = 1)	Hungary	AY597976 (CR)+DQ532114	This study	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Dor70 common carp	D70 (n = 1)	Israel	AY597981 (CR)+DQ532115	This study	*C. c. carpio*	*C. c. carpio*	*C. c. carpio*
Cyprinus melanes .Y*	CYM (n = 1)	Vietnam	DQ464943 (CR), 464970 (cyt b)	This study	_	_	_
Genbank	GEB (n = 1)	Unknow	X6100	Chang *et al*. (1994)	*C. c. haematopteus*	*C. c. haematopteus*	*C. c. haematopteus*
Goldfish	CRU	Japan	NC002079	Murakami *et al*. (1998)			

* see Nguyen and Ngo (2001)

Table 3.2. Tests of alternate phylogenetic hypotheses using Kishino-Hasegawa (KH), Shimodaira-Hasegawa (SH) test. (* significant difference between optimal and alternate topologies, P <0.05).

Tree	1nL	Diff- 1nL	KH	SH
optimal	3197.840	Best		
Zhou *et al*. (2004b)	3230.308	32.468	0.030*	0.030*
Baruš *et al*. (2002)	3243.448	45.607	0.020*	0.010*
Kohlmann *et al*. (2005)	3199.554	1.714	0.910	0.670

Table 3.3. Nucleotide divergence (%) of common carp and other fish in *Cyt b* gene and *CR*.

Genus or species	Cytochrome b	Control	Sourse
Cyprinus carpio			This study
Intraspecific	0.0 - 2.49	0.0 - 3.41	
Distoechodon			
Intraspecific	0.50 - 3.20	-	(Liu 2002)
Interspecific	9.60 - 10.70	-	(Liu 2002)
Galaxias maculatus			
Intraspecific	0.30 - 14.60	-	(Waters and Burridge 1999)
Rainbow fish			
Intraspecific	< 1.00	-	(Zhu *et al* . 1994)
Interspecific	15.00 - 19.00	-	(Zhu *et al* . 1994)
Rhinogobinus			
Intraspecific	-	0.80 - 1.80	(Chen *et al* . 1998)
Interspecific	-	11.30 - 11.70	(Chen *et al* . 1998)
Macquaria noremaculeata			
Intraspecific	-	0.30 - 3.50	(Jarry and Baverstock 1998)
Acanthopagrus			
Interspecific	-	15 - 29	(Hurt *et al* . 2001)
Anguilla			
Intraspecific	-	1.00 - 3.00	(Sang *et al* . 1994)
Interspecific	-	22	(Sang *et al* . 1994)

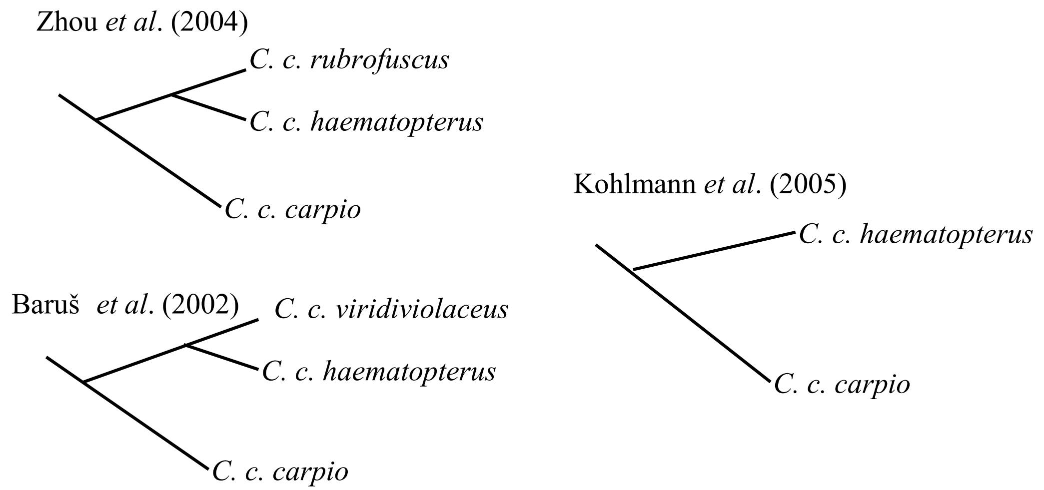

Figure. 3.1. Phylogenetic hypotheses examined based on the species studied. Zhou *et al*. (2004b), Baruš *et al*. (2002) and Kohlmann *et al*. (2005) indicate representative topologies based on their more extensive studies.

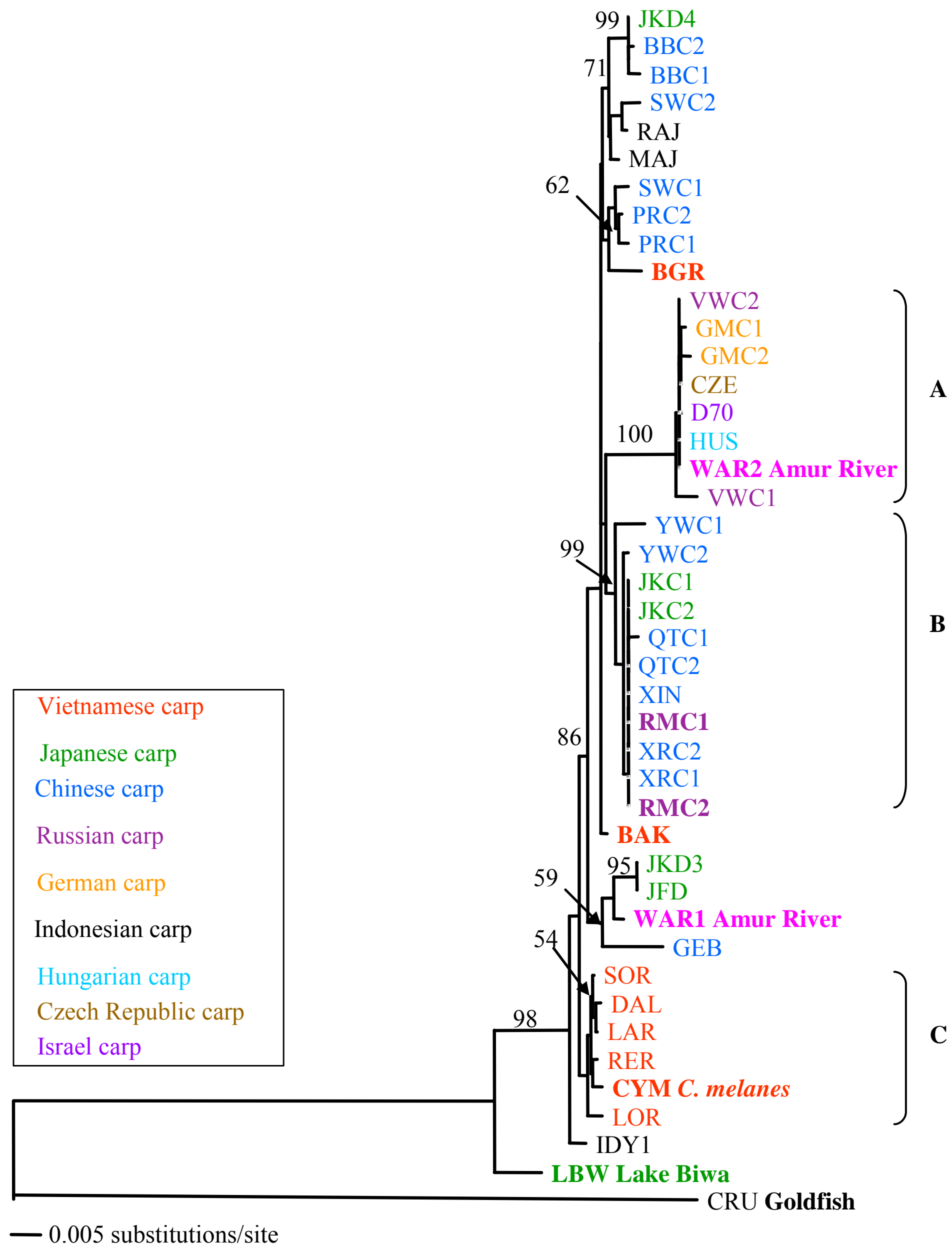

Figure. 3.2. Neighbour-joining estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations (- = <50% bootstrap support).

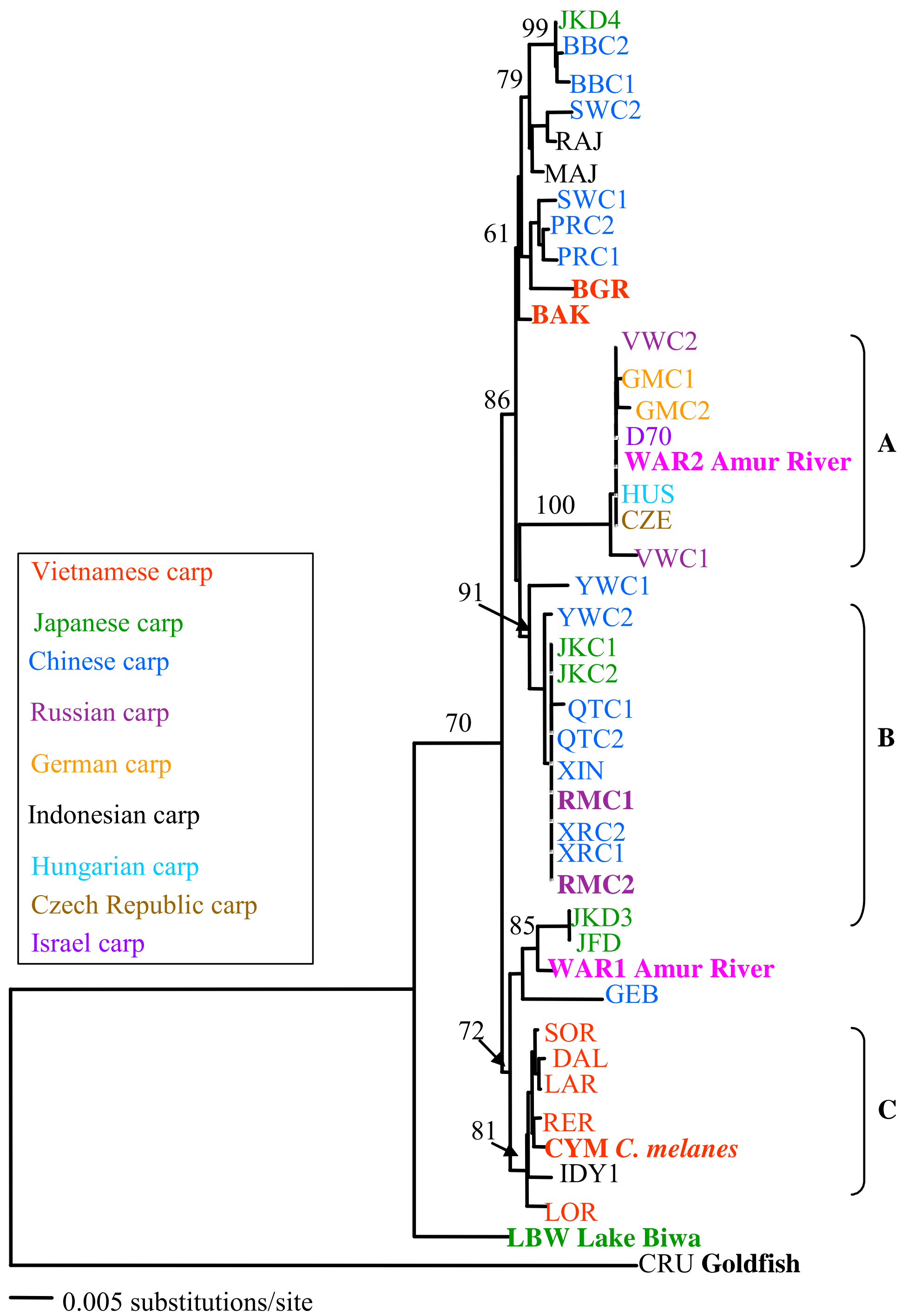

Figure. 3.3. Maximum likelihood estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations (- = <50% bootstrap support).

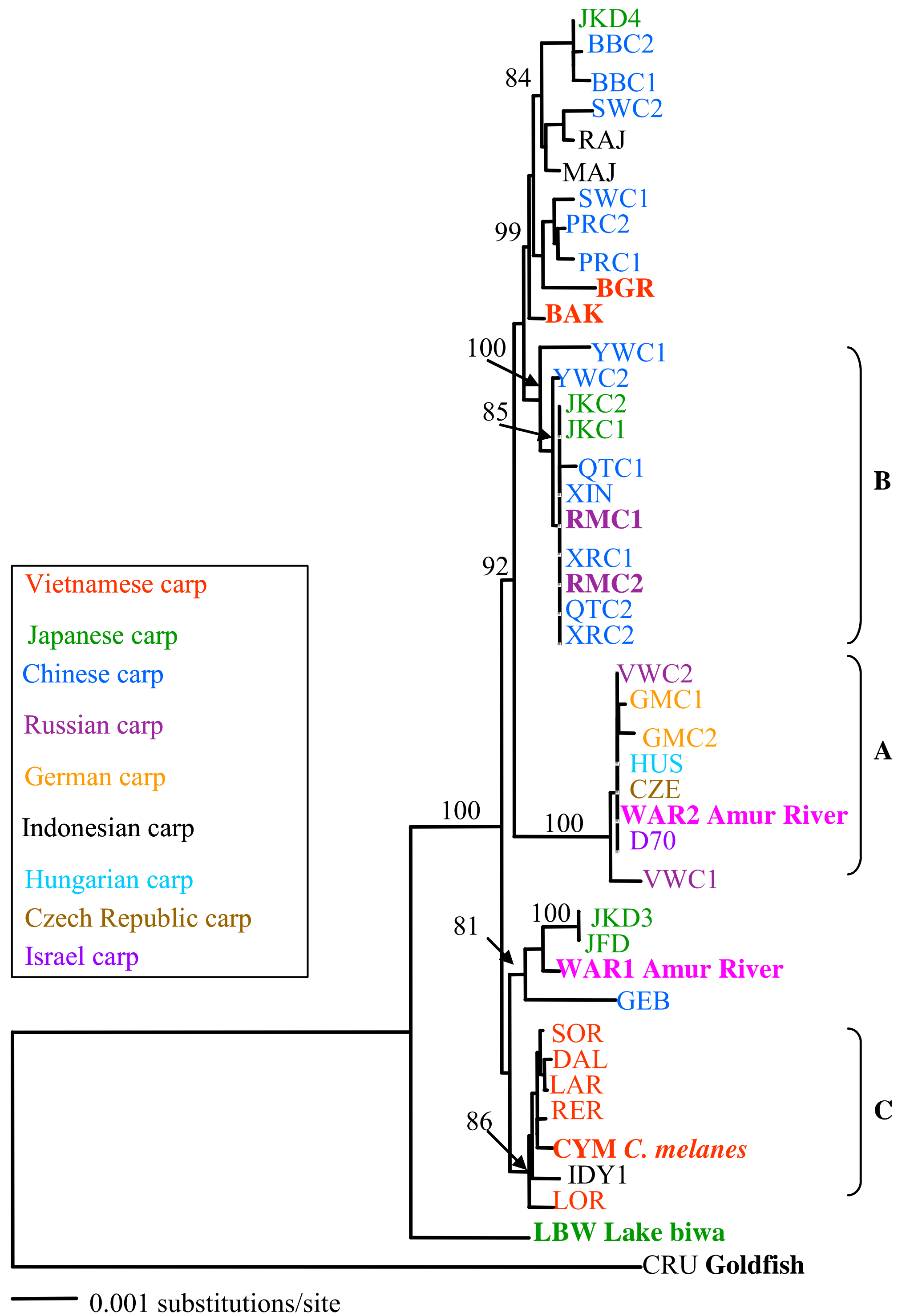

Figure. 3.4. Bayesian estimate of phylogenetic relationships based on *CR* and *Cyt b* mitochondrial DNA gene regions among common carp strains or populations.

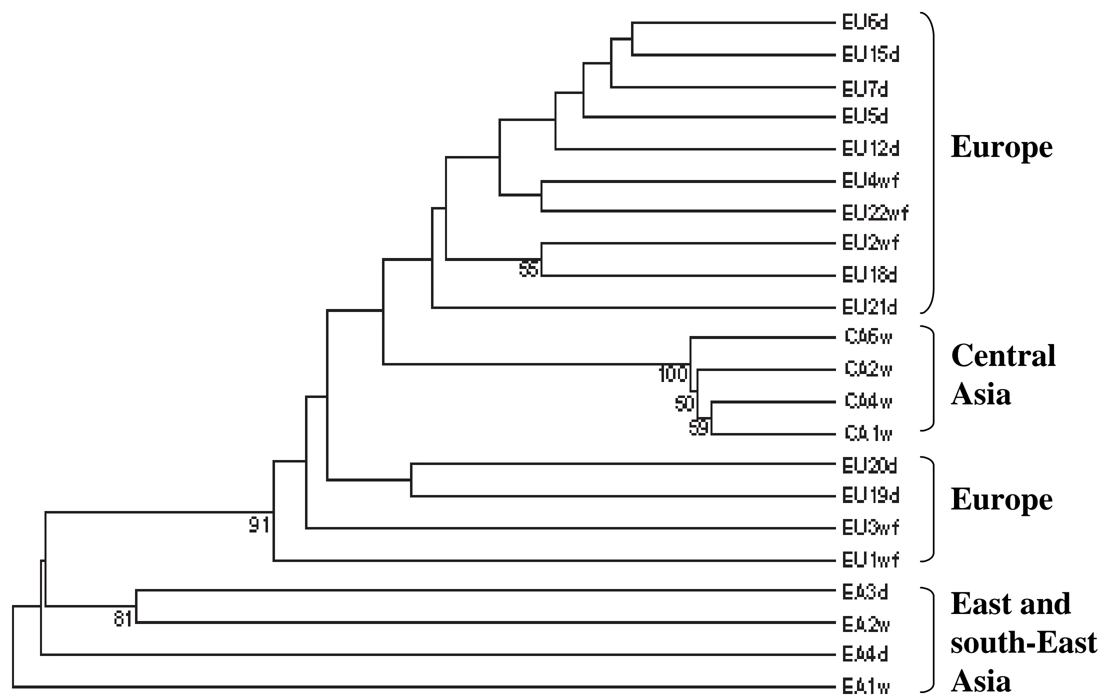

Figure. 3.5. UPGMA clustering of common carp populations based on microsatellite variability at four loci derived from Kohlmann *et al*. (2005).

Chapter 4.

Genetic diversity of common carp in Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA control region*

4.1 Introduction

The domestication of common carp (*Cyprinus carpio* L.) commenced over 4,000 years ago in China and at least 600 years ago in Europe and now over three million tonnes are produced annually (FAO, 2003). The species is cultured in many parts of the world both within and outside its natural range as a result of widescale translocations. Common carp is one of the main aquaculture species in Vietnam (Nguyen, Ngo, 2001) and is cultivated along with Chinese carp and Indian carp in polyculture ponds and rice-fields. These species together makes up one of the most important food and income resources in rural communities in Vietnam (Edwards *et al.*, 2000). The enhancement of common carp production in Vietnam has focused substantially upon the development of genetically improved strains (Tran, Tran, 1995). For this purpose, Hungarian and Indonesian carp strains were imported into Vietnam almost 30 years ago for crossbreeding and mass selection programs with local Vietnamese carp.

---

* A peer reviewed publication was derived from this chapter: Thai, T. B., T. A. Pham and C.M. Austin (2006) Genetic diversity of common carp in Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA control region. *Aquaculture* 258: 228-240.

The genetic improvement strategy adopted for common carp in Vietnam involved the development of hybrid common carp by crossbreeding among three genetic lines (Vietnamese white, Hungarian scale and Indonesian yellow common carp) coupled with mass selection.

The ongoing development of farmed common carp stocks in Vietnam requires that several important issues are addressed. These included the possibility that: (1) stocks of experimental or genetically improved lines have become mixed with other stocks in ponds at Research Institute for Aquaculture No 1 (RIA1) or provincial hatcheries; (2) distribution of these lines as seed stock or broodstock to regional hatcheries or farms has been ineffective in increasing genetic diversity, and (3) reduction in effective population sizes (*N*e) in genetically improved or hatchery maintained stocks has occurred, leading to loss of genetic variation, gene frequency changes and possibly inbreeding depression. This last possibility is consistent with reports by many farmers of slow growth and early maturity in their cultured stocks of common carp (T. A. Pham, personal communication).

In addition to the genetically improved lines of common carp produced at RIA1, eight local varieties have been recognised in Vietnam with reported useful features for culture and for marketing (Bakos, Gorda, 2001). There is now the possibility that distinct wild common carp stocks are in decline because of excessive harvesting and interbreeding with introduced common carp strains (Nguyen, Ngo, 2001). Therefore, it is important to identify and characterise genetically distinct local varieties of cultured fish for conservation purposes in general and specifically for common carp populations (Mabuchi *et al.*, 2005; Murakaeva *et al.*, 2003).

A wide range of molecular marker systems have been used to study genetic diversity in aquaculture species (Chen *et al.*, 2004; Diniz *et al.*, 2005; Garoia *et al.*, 2004; Sato *et al.*, 2005; Sotka *et al.*, 2005; Valles-Jimenez *et al.*, 2004; Yu, Guo, 2004). As one of the most important aquaculture species in the world, genetic variation in carp has been examined in many parts of the world using microsatellites (Desvignes *et al.*, 2001; Kohlmann *et al.*, 2005), RAPDs (Bartfai *et al.*, 2003), AFLPs (David *et al.*, 2001), allozymes (Kohlmann, Kersten, 1999), RFLPs (Gross *et al.*, 2002; Zhou *et al.*, 2003) and direct sequencing of mtDNA fragments (Froufe *et al.*, 2002; Zhou *et al.*, 2004b). While these have included studies of both European and Asian common carp so far there have been no detailed studies of molecular genetic variation in common carp in Vietnam nor has the technique of Single Strand Conformation Polymorphism (SSCPs), which is being increasingly used to assay for genetic variation in fish species (Aurelle, Berrebi, 2001; Liu, Cordes, 2004), been applied to common carp.

In this study, I investigated genetic diversity of Vietnamese common carp strains and populations by using direct DNA sequencing and SSCP analysis of the mtDNA *CR* . This region of the mtDNA molecule was chosen because it has a very high nucleotide substitution rate, making it particularly useful for estimating the genetic population structure of closely related animal populations (Sivasundar *et al.*, 2001; Vigilant *et al.*, 1991). These data are then used to examine a range of questions relating to genetic diversity and management of wild and domesticated populations of common carp in Vietnam.

4.2 Materials and Methods

4.2.1 Sample collection

Common carp were collected in 2003 and 2004 from broodstock populations maintained by eleven hatcheries, three experimental common carp lines maintained at RIA1, and by sampling six wild populations, using seine and lift nets. Samples of common carp strains from China, Indonesia, Japan, Hungary and India were obtained for comparative purposes. Tissue samples were taken as fin clips and preserved in 90% ethanol. Locality and sample size details are provided in Table 4.1 and Fig. 4.1.

4.2.2 DNA extraction and sequencing of control region

The total genomic DNA was extracted using the protocol from Crandall *et al*. (1999). Between three and six individuals from each populations or strain were first analysed by direct sequencing. The mitochondrial *CR* was amplified using primers Carp-Pro (5' AAC TCT CAC CCC TGG CTA CCA AAG 3') and Carp-Phe (5' CTA GGA CTC ATC TTA GCA TCT TCA GTG 3') designed from the common carp whole mitochondrial genome sequence (GenBank, AC: X61010). PCR was carried out in 50 μl reactions volumes (1 X reaction buffer, 2 mM dNTP, 1.5 mM $MgCl_2$, 0.5 μM of each primer, 0.5 units *Taq* DNA polymerase, and approximately 200 ng DNA template). Thermal cycling comprised 95 ^oC for 3 min, followed by 34 cycles of 95 ^oC for 30 s, annealing at 55 ^oC for 30 s, and an extension temperature of 72 ^oC for 1 min. This was then followed by a final extension of 72 ^oC for 3 min. PCR products were purified using the QIA quick PCR purification kit (Qiagen Hiden Germany), following ABI PRISM BigDye Terminator (Foster city, CA, USA) sequencing protocols. For each individual, sequencing reactions were performed

using both forward and reverse primers, resulting in a consensus fragment of 745 bp in length.

4.2.3 Single Strand Conformation Polymorphism (SSCP) amplification

A short highly variable fragment of the *CR* was selected from the common carp sequences obtained as described above. This fragment was approximately 230 bp and was amplified by the primers F1 (5' GCA GGT ACA TAA TAT TAA 3') and R1 (5' CAG ATG CCA GTA ATA ATT 3'). PCR was carried out in 10 µl reaction volumes (1 X reaction buffer, 2mM dNTPs, 1.5mM $MgCl_2$, 0.5µM of each primer, 0.1 units *Taq* DNA polymerase, and approximately 10ng DNA template). Thermal cycling comprised 94 ^oC for 3 min followed by 34 cycles 94 ^oC for 20 s, annealing at 55 ^oC for 20 s, and an extension temperature of 72 ^oC for 50 s, followed by a final extension of 72 ^oC for 3 min. PCR product size was estimated using the Promega DNA/Hae III marker.

To obtain SSCP phenotypes, PCR product (1µl) was added to 7 µl of loading buffer (99% formamide, 1mM NaOH, 0.2% w/v bromophenol blue and xylene cyanol), denatured for 2 min at 94 ^oC and placed directly on ice. Samples were loaded onto an 8% polyacrylamide (37.5: 1 acrylamide-bisacrylamide) gel (16 cm long, 1.5 mm thick) containing 5% glycerol and 0.5 X TBE, and run at 5W for 12h at 4 ^oC. SSCP products were visualised by silver staining using the method of Mirol *et al*. (2002b). The resultant bands were scored by comparison with five standard carp haplotypes (representing the five most frequently encountered haplotypes), which were included on each gel for reference. Rare haplotypes were subsequently run side by side to ensure they were correctly scored.

4.2.4 Data analysis

Nucleotide (π) and haplotype diversity (h) were calculated from the sequence data using DNASP 4.10 (Rozas *et al.*, 2003). Sequences were aligned using the program Clustal X (Thompson *et al.*, 1997). The most suitable model of evolution was obtained using Modeltest 3.06 (Posada, Crandall, 1998). This model was used to calculate pairwise sequence distances between haplotypes and for the construction of a Neighbour-Joining dendrogram (NJ) using PAUP*4.0b.10 (Swofford, 2000). Confidence levels in the resulting relationships were assessed using the bootstrap procedure with 1,000 pseudoreplicates (Swofford, 2000). The corresponding *CR* region sequence for common carp from GenBank (Chang *et al*., 1994) was included in the data set for comparative purposes, and the *Carassius auratus CR* sequence (GenBank accession NC_002079) was used as the outgroup.

For the SSCP data set, haplotype diversity for each population was calculated using ARLEQUIN (Schneider *et al.*, 2000). Genetic divergence between samples was estimated by pairwise *F*st analysis using the same program. Patterns of overall genetic relationships among populations were summarised using UPGMA clustering of Roger's (1972) genetic distance as implemented in TFPGA (Miller, 1997). Multidimension scaling (MDS) was performed on matrices of genetic distance to test for the presence of nonhierarchical patterns of relationships among populations using SPSS 10.0. The partitioning of genetic diversity within and between populations and populations grouped according to origin (experimental line, hatchery and wild), was undertaken by an analysis of molecular variance (AMOVA) for binary data, using the GenALEX add-in for Microsoft Excel (Peakall, Smouse, 2001). Multiple

comparisons were subjected to sequential Bonferroni correction to control the Type I error rate (Rice, 1989).

4.3 Results

4.3.1 Control region sequences and SSCP variation

Sequences for a 745 bp fragment of the mtDNA *CR* were obtained from 111 fish representing 41 populations or strains. A total of 19 haplotypes with 78 variable and 30 phylogenetically informative sites were identified. The nucleotide composition was A + T rich (A= 31%; T= 32%), and variation consisted predominantly of transitions (Ti : Tv = 2.56). All sequences have been deposited in GenBank (AC: AY597942-AY597976; DQ354144-DQ354149).

Vietnamese carp populations have high haplotype diversity (mean = 0.92±0.02), but low nucleotide diversity (mean = 0.01±0.00). The most divergent Vietnamese haplotypes differed by only 9 base pairs. Diversity and relationships among haplotypes are depicted in Fig. 4.2 together with the 14 corresponding SSCP phenotypes determined from the shorter (230 bp) fragment. From this figure it can be seen that the SSCP technique was successful in resolving a significant proportion of the nucleotide variation detected in sequencing the longer *CR* fragment. It is noteworthy that four of these haplotypes allow the discrimination of Vietnamese white (haplotype C), Hungarian (haplotype A) and Indonesian yellow (haplotypes B & D) carp strains from RIA 1. In addition, common carp samples from China (haplotypes I & R) and Indonesian (haplotypes B, D &J) and Koi carp (haplotype L) were all distinguishable from Vietnamese carp.

The summary of relationships among *CR* haplotypes (Fig. 4.2) shows that, apart from two of the Chinese strains, which share the same haplotype, and which are quite divergent from the other carp samples, there are a large number of haplotypes that are closely related to each other. Minor exceptions are the Hungarian carp haplotype (A) and a haplotype found in Bak Kan and Dak Lak (haplotype E), the Bang Giang River (haplotype F) and the Lo River (haplotype G) and Koi carp (haplotype L).

A total of 968 individuals from both wild and hatchery populations were scored for SSCP variation. In addition to the non-Vietnamese strains that had seven distinguishable SSCP phenotypes, five SSCP haplotypes were distinguishable among Vietnamese common carp samples. Comparison with the nucleotide sequences revealed that these SSCP haplotype differ by 3-8 bp. Haplotype frequencies and diversity estimates are summarized in Table 4.2 for 20 common carp populations. Three haplotypes, Hungarian (A), Indonesian (B) and Vietnamese (C), predominated in common carp samples and five (D, E, F, G and H) were relatively rare or occurred only at low frequencies. Intra-population diversity varies widely among the populations ranging from populations with a single haplotype (h =0) to six haplotypes (h = 0.55). The experimental strains from RIA1 have the lowest diversity (h= 0-0.28), the hatchery stocks with the exception of Thai Nguyen, have high diversity (h = 0.49-0.64) and the wild stocks have an intermediate level of diversity (h = 0.26-0.41) (Table 4.2).

The three experimental strains from RIA1 are also highly differentiated from each other. The Hungarian scale strain is fixed for haplotype A, the Indonesian yellow strain is dominated by haplotype B (84%), while the Vietnamese white strain in

dominated by haplotype C (94%). All three of these haplotypes are found in almost all hatchery and wild carp populations, however in the Vietnamese white strain haplotype C dominate (55%) followed by the predominate Indonesian haplotype B at 25% and the Hungarian haplotype A at 12%.

The six wild common carp populations (RER, LOR, LAR, SOR, DAL, and BGR) have generally similar haplotype profiles and like the Vietnamese experimental strain, haplotype C predominates. Six of the hatchery stocks have haplotype profiles largely similar to the wild populations (haplotype C = 0.52-0.94). The five other hatchery samples, in contrast, have haplotype profiles dominated by the Indonesian haplotype B (0.50-0.84), although it is noteworthy that they also all possess the Vietnamese haplotype C, albeit at a lower frequency (0.04-0.39). Interestingly, almost all hatchery and wild samples have the Hungarian haplotype (A), although it mostly occurs at a relatively low frequency (0.02-0.20).

4.3.2 Genetic differentiation and relationships among populations

Pairwise Fst analyses indicates significant genetic heterogeneity among populations with the majority of pairwise comparisons yielding signification differences (Table 4.3). All three experimental strains are highly differentiated from each other (Fst = 0.78-0.94; $P<0.05$). The Hungarian strain was the most divergent and is significantly differentiated from all other common carp populations (Fst = 0.58-0.94; $P<0.05$). The Indonesian strain is also highly distinct, and is significantly different from all other samples except for four hatchery populations. The Vietnamese strain is also divergent from most other samples with the exception of four of the wild populations.

The extent of the difference between the three experimental lines and their relationships to the hatchery and wild samples are clearly evident from the UPGMA dendrogram (Fig. 4.3) where it can be seen that the wild samples cluster with the Vietnamese white experimental line (group C1 in Fig. 4.3) and the hatchery populations cluster either with the Indonesian strain (group B in Fig. 4.3) or form a cluster (group C2 in Fig. 4.3) linked to the wild populations, which together form the Vietnamese cluster (group C in Fig. 4.3).

Multidimensional scaling (Fig. 4.4) also emphasises the clear differentiation of the three experimental strains and largely re-inforces the findings of the preceding UPGMA analysis. From Figure. 4.4, it is also clear that hatchery and wild stocks are well differentiated from the Hungarian experimental line and that all hatchery stocks are genetically intermediate between the Vietnamese white strain (or closely related populations) and the Indonesian strain. In contrast to the relationship depicted by the UPGMA dendrogram (Fig. 4.3), it is apparent that the hatchery stocks did not so much fall into two distinct groups associated with either the Vietnamese or Indonesian strains, but represented more of a continuum between these two stocks. For example population Tuyen Quang (TUQ) placed in the Vietnamese cluster and populations Can Tho (CAT) and Sai Gon (SAG), placed in the Indonesian cluster in Fig. 4.3 are actually quite similar genetically, and fall into an intermediate position on MDS axis 1 between the Vietnamese and Indonesian strains.

The AMOVA analysis indicates that the genetic variation is partitioned very differently within and between populations for the experimental strains, hatchery stocks and wild populations (Table 4.4). For the experimental group, the variation is

predominately between populations (86.30%) with very low levels of within populations (13.70%). This is in contrast to the wild populations for which the pattern of variation is reversed with 96.2% of the variation within populations and only 3.8% between populations. The hatchery populations have intermediate values, with within population variation significantly enhanced (80.47%) compared with the experiment lines and the between population variation substantially elevated compared with the wild populations (19.53%).

4.4 Discussion

4.4.1 Control region sequence variation and utility of the SSCP techniques

The sequencing of the *CR* revealed significant haplotype diversity among Vietnamese common carp samples and also distinguished non-Vietnamese strains, which is consistent with findings of other studies that have found this region to be highly variable in fish (Sivasundar *et al.*, 2001). Nevertheless, the degree of haplotype divergence was relatively low with the most divergent haplotypes differing by only 14 bp. Finding low levels of sequence divergence in the *CR* in carp is consistent with other studies on this species using this mitochondrial region and protein coding gene fragments (Froufe *et al.*, 2002; Mabuchi *et al.*, 2005; Thai *et al.*, 2004). This low level of divergence could reflect a shallow evolutionary history or it could be due to a slower rate of evolution for the mtDNA region in carp. The finding of only limited variation in the *CR* of brown trout has been attributed to an unusually slow rate of evolution for this fragment in this species (Aurelle, Berrebi, 2001; Giuffra *et al.*, 1996). However, this is unlikely to explain the low levels of variation

in this region in common carp as sequences from other high to moderately variable mitochondrial gene regions in fish including *COI*, *Cytb* and *ATPase*, also show low levels of variation in this species (Thai *et al.*, 2004; Wang, Li, 2004).

The SSCP procedure revealed only a portion of the haplotype diversity revealed by sequencing (14 out of 25 haplotypes). This loss of information is a result of the need to compromise between the competing needs of higher sensitivity (i.e. ability to distinguish fragments that differ 1-2 nucleotide positions) achieved through the use of shorter fragments versus the concomitant reduction in the number of variable nucleotide positions within such fragments. The SSCP procedure in this study able to detect phenotypes that differed by as few as 3 base pairs, which is similar to other studies on fish (Aurelle, Berrebi, 2001).

Despite the SSCP procedure only resolving a portion of the haplotype variation detected by direct sequencing, it importantly allowed all the non-Vietnamese strains to be distinguished from Vietnamese samples and resolved significant haplotype diversity among the latter samples. Thus, the variation detected using the SSCP technique was sufficient to allow effective comparisons among Vietnamese samples, which revealed significant and contrasting patterns of variation within and between experimental strains, hatchery and wild populations of carp. These findings and the insights they provided into the history of carp domestication in Vietnam are discussed below.

4.4.2 Genetic diversity of common carp in Vietnam

The overall diversity of common carp in Vietnam is considerably less than the diversity at the global level in this species. Several genetically differentiated carp strains found in China, Indonesia and Japan, identified in this study and by others (Froufe *et al.*, 2002; Mabuchi *et al.*, 2005; Zhou *et al.*, 2004b) are not present within Vietnamese wild and hatchery populations. This indicates that there have been few, if any, undocumented common carp introductions into Vietnam and that there are additional genetic resources outside the country, including relatively divergent lines that could be potentially utilized for future breeding programs in Vietnam.

The genetic variation detected with the mtDNA *CR* in Vietnamese carp, comprising both wild and domesticated lines, is consistent with the known history of this species in this country (Tran, Tran, 1995). The history includes the introduction of carp from Indonesia and Hungary and the dissemination of progeny of genetically improved strains to hatcheries.

Comparison between the results of our study and other studies on carp are difficult because no previous studies have been published on SSCP variation in carp. Surprisingly, given the importance of carp as a food fish worldwide, detailed genetic knowledge is lacking on domesticated and wild populations, especially within countries or regions other than using allozyme information (Kohlmann *et al*., 2003). Those studies that have addressed issues relating to genetic variation have usually used restricted sampling in terms of populations or sample sizes (Bartfai *et al.*, 2003; Lehoczky *et al.*, 2005), or have confounded the effects of geographic variation and domestication due to unequal representation of stocks and uneven sample sizes

(Kohlmann *et al.*, 2005). Nevertheless, our study is consistent with a number of studies using a variety of markers, including microsatellite, RFLPs and sequencing, that consistently revealed low levels of variation within European carp populations and an elevated diversity within and between Asian common carp populations (Kohlmann *et al.*, 2005; Thai *et al.*, 2004).

In addition, our findings of higher haplotype diversity within wild stocks compared with the RIA1 experimental lines is consistent with genetic studies of carp that have generally reported reduced genetic diversity in domesticated carp (Kohlmann, Kersten, 1999; Kohlmann *et al.*, 2005). For example, Kohlmann *et al*., (2005) found allelic richness based on microsatellite loci to vary from 4.44 in domesticated stocks to 8.22 in wild stocks using four microsatellite loci. Kohlmann *et al*., (2003) attributed these kinds of differences to effects of bottlenecks and losses of variation due to inbreeding and genetic drift in small hatchery populations similar to many other studies on domesticated fish species (Ward, Grewe, 1995). Such phenomena would most likely account for the low levels of variation found in the Hungarian, Indonesian and Vietnamese common carp experimental lines maintained at RIA1. These lines have been maintained for six generations and while an attempt has been made to maintain variation by using 50-100 families for the propagation of each generation, it is most likely that variation was either limited within the founding stocks or variation has been subsequently lost due to low *N*e (Vandeputte, 2003). The high degree of differentiation among the three experimental stocks indicates that there has been little, if any, accidental mixing of stocks and that the strategy of cross breeding among these stocks to enhance genetic diversity within cultured Vietnamese common carp stocks has been justified.

Genetic variation is often found to be lower in domesticated lines of cultured fish species. For example, genetic diversity in channel catfish has been found to be lower in domestic lines than in wild stocks (Mickett *et al.*, 2003; Simmons *et al.*, 2006). The main reason for these observations is generally thought to be the small number of founders. Thus, in contrast to theses kinds of observations and those of Kohlmann *et al*. (2003; 2005) for carp, our finding that all but one of the hatchery stocks of common carp (Thai Nguyen) examined in Vietnam had elevated diversity compared to both the experimental lines and the wild populations is somewhat unusual. Findings of elevated diversity in domesticated stocks are usually attributed to mixing of stocks during founding and subsequent propagation (Ferguson, 1995b; Thompson, 1985). In the case of common carp in Vietnam, the elevated diversity in hatchery stocks can be attributed to the successful dissemination of genetically improved carp partially derived from imported stocks. The effectiveness and extent of this program of dissemination is examined in further detail below.

4.4.3 Insights into the dissemination of cultured common carp in Vietnam

Research into the genetic improvement of carp conducted over 15 years in Vietnam has led to the development of three genetically improved lines through a combination of crossbreeding and mass selection which are referred to as “three blood” carp (Tran, Nguyen, 1992). These strains are referred to as the “Indonesian”, “Hungarian” and “Vietnamese” strains on the basis of the female parent used in the second of two generations to construct each cross. It is however, uncertain the extent to which these strains have been disseminated to regional hatcheries and small scale farmers and the relative success or performance of these strains (Edwards *et al.*, 2000). The effectiveness of the dissemination program of these three “three blood” strains can be

assessed by examining the distribution of SSCP haplotypes in hatchery stocks because of the maternal inheritance of the mitochondrial genome.

The analysis of the SSCP haplotype frequencies among the hatchery stocks indicates that the Hungarian strain has either not been disseminated to any significant extent or has a high mortality. The much higher proportion of the Indonesian haplotype suggests that the three blood Indonesian strain has been more extensively disseminated or has better survival. Although it should be noted that the high similarity between the southern samples (CAT and SAG) and the Indonesian line in RIA 1 most likely reflects the original introduction of Indonesian carp into southern Vietnam, where no indigenous common carp stocks occur.

A number of hatchery stocks show intermediate levels of divergence between the Vietnamese and Indonesian strains as is apparent from the MDS axis 1 (Fig. 4.4), and thus most likely represent various admixtures of domesticated Vietnamese stocks (either domesticated or from the wild) and the Indonesian three blood line. It is not possible to determine if this represents the dissemination of a mixture of stocks to the original hatcheries or subsequent stock mixing or a combination of both. A by product of the mixing of stocks has been the elevation of genetic diversity levels and it is apparent that those stocks which have scores equidistant between the Vietnamese and Indonesian experimental lines on MDS axis one have the highest diversity levels (Fig. 4.3). Thus, finding the hatchery populations to have elevated diversity compared to wild and experimental lines indicates that the Vietnamese Government's program of crossbreeding and dissemination has been effective in not just preventing the erosion of genetic variation but enhancing diversity of

domesticated stocks. However, genetic diversity is to be maintained and inbreeding minimised in common carp over the long term, additional stocks should be introduced to hatcheries on a regular basis. These stocks should be preferably from genetic improved lines rather than from unselected or wild stocks so that gains in performance (ie growth rate) are not diluted for the sake of the overall maintenance of genetic diversity.

4.4.4 Conservation of wild carp stocks

Wild common carp stocks, like those of many fish species, are under threat from a number of processes including environmental degradation (Murakaeva *et al.*, 2003) and the introduction of "exotic" genotypes into natural populations as a result of domestication and translocation (Mabuchi *et al.*, 2005). The results of this study indicate that wild Vietnamese common carp are relatively homogeneous genetically, but distinct from other strains of both Europe and Asia. The genetic distinctiveness of Vietnamese common carp revealed in this study, based on the analysis of the mtDNA *CR* is supported by Kohlmann *et al*'s (2003) analyses of genetic variation in common carp. These authors examined allozymes, mtDNA and microsatellite variation, throughout the species distributional range although their study only included a single sample of Vietnamese carp from the Red River.

It is suggested that the indigenous Vietnamese strains potentially represent a unique genetic resource for common carp and therefore needs to be conserved. Some of the distinct Vietnamese haplotypes are associated with forms that also have distinct phenotypes and behaviours, several of which are considered important for aquaculture. For example, although growth of Bac Kan common carp is

comparatively slow, it is an important strain as it is adapted to the rice-field environment and farmers can maintain broodstock and do not depend on wild seed (Edwards *et al.*, 2000). Thus, the distribution of genetically improved carp may need to be done with care in Vietnam and it should acknowledged that indigenous local populations may be threatened by the translocations of domesticated varieties. The diversity of environments and culture methods in Vietnam makes it unlikely that a single improved variety of common carp will be optimal throughout the country.

4.4.5 Conclusion

This study has demonstrated that the SSCP method, using a fragment of the mtDNA *CR*, is an effective and efficient method for assaying molecular genetic variation within and between populations of common carp. A survey of genetic variation amongst experimental lines, hatchery stocks and wild populations of common carp in Vietnam revealed contrasting patterns in the distribution of molecular genetic variation that contributes important insights into the domestication process and provides information that is essential for the effective management of domesticated and wild carp stocks in this country. While the data presented in this study represents one of the most comprehensive DNA-based data sets so far collected at a regional level for common carp, it will be important to extend this study to the sampling of other gene loci (e.g. microsatellites) and populations, especially those from small-scale farmers, private hatcheries and from markets, in order to obtain a more complete picture of the genetics of common carp in Vietnam.

Table 4.1. Location, code and number of samples sequenced and analysed by the SSCP technique.

Population	Code	Location	Type	Population size (n)	
				Sequencing	SSCPs
Hungarian scale-RIA1[1]	HUS	Tu Son, Bac Ninh, Vietnam	E	4	50
Indonesian yellow-RIA1	IDY	Tu Son, Bac Ninh, Vietnam	E	6	50
Vietnamese white-RIA1	VNW	Tu Son, Bac Ninh, Vietnam	E	4	50
Vinh Phuc	VIP	Me Linh, Vinh Phuc, Vietnam	H	4	50
Thai Nguyen	THN	Cu Van, Thai Nguyen, Vietnam	H	3	50
Son La	SOL	Son La town, Son La, Vietnam	H	4	50
Bac Kan	BAK	Bach Thong, Bac Kan, Vietnam	H	6	50
Tuyen Quang	TUQ	Hoang Khai, Tuyen Quang, Vietnam	H	4	50
Yen Bai	YEB	Van Chan, Yen Bai, Vietnam	H	3	50
Hoa Binh	HOB	Hoa Binh town, Hoa Binh, Vietnam	H	6	50
Ha Tinh	HAT	Duc Long, Ha Tinh, Vietnam	H	4	50
Can Tho	CAT	Cai Rang, Can Tho, Vietnam	H	4	36
Sai Gon	SAG	Binh Chanh, Sai Gon, Vietnam	H	4	35
Thac Ba Reservoir	TBR	Yen Binh, Yen Bai, Vietnam	H	3	50
Bang Giang River	BGR	Cao Bang town, Cao Bang, Vietnam	W	6	50
Lo River	LOR	Yen Son, Tuyen Quang, Vietnam	W	4	50
Red River	RER	Van Giang, Hai Hung, Vietnam	W	4	50
Lam River	LAR	Nam Dan, Nghe An, Vietnam	W	3	50
Son River	SOR	Bo Trach, Quang Binh, Vietnam	W	4	47
Dak Lak	DAL	Ea Kao, Dak Lak, Vietnam	W	4	50
Xingguonensis	XIG	Jaing xi China		3	5
Wananensis	WAN	Jaing xi China		3	5
Wuyuanensis	WUY	Jaing xi China		3	5
Color	COL	Jaing xi China		3	5
Red Koi	REK	Komaki Japan		3	21
Wild Amur	WAR	Karnataka, India		3	5
Majadanu	MAJ	Sukamandi, Indonesia		3	5
Rajadanu	RAJ	Sukamandi, Indonesia		3	5
Widan	WID	Sukamandi, Indonesia		3	5
GenBank	GBK	Taiwan[2]		1	
Goldfish	GOF	Unknown		1	

[1] Research Institute for Aquaculture No 1; [2] Origin of sample not provided; E: Experimental group; H: Hatchery group; W: Wild group.

Table 4.2. Number of haplotypes and haplotype diversity in each common carp population. Population code given in Table 4.1

Haplotypes	Experimental			Hatchery											Wild						Mean
	H	I	VN	VP	TN	SL	BK	TQ	YB	HB	HT	CT	SG	TBR	BGR	LR	RR	LA	SR	DL	
A	1.00			0.12	0.08	0.16	0.08	0.24	0.06	0.08	0.20			0.20	0.02	0.04	0.08	0.02			0.12
B		0.84	0.04	0.66	0.80	0.62	0.08	0.20	0.22	0.22	0.04	0.50	0.54	0.02	0.02	0.04	0.04	0.06		0.14	0.25
C		0.04	0.94	0.14	0.10	0.18	0.66	0.52	0.62	0.68	0.68	0.39	0.31	0.74	0.78	0.84	0.86	0.90	0.76	0.82	0.55
D		0.12	0.02	0.08		0.04	0.02	0.02	0.04	0.02	0.08	0.11	0.14		0.02	0.02	0.02	0.02			0.04
E					0.02		0.12		0.06										0.13	0.04	0.02
F							0.04	0.02							0.16						0.01
G														0.04		0.06					0.01
H																			0.19		0.01
No of haplotypes	1	3	3	4	4	4	6	5	5	4	4	3	3	4	5	5	4	4	3	3	3.85
Haplotype diversity	0.00	0.28	0.12	0.53	0.35	0.57	0.55	0.64	0.57	0.49	0.50	0.60	0.60	0.42	0.37	0.29	0.26	0.19	0.41	0.31	0.40

Table 4.3. Pair-wise estimate of variance of haplotype frequencies (*F*st) among of samples. Population codes given in Table 4.1.

	HUS	IDY	VNW	VIP	THN	SOL	BAK	TUQ	YEB	HOB	HAT	CAT	SAG	TBR	BGR	LOR	RER	LAR	SOR	DAL
HUS																				
IDY	0.86*																			
VNW	0.94*	0.78*																		
VIP	0.70*	0.05	0.61*																	
THN	0.81*	0.02	0.73*	0.01																
SOL	0.66*	0.09*	0.58*	0.01	0.03															
BAK	0.70*	0.54*	0.12*	0.36*	0.48*	0.32*														
TUQ	0.58*	0.43*	0.24*	0.23*	0.35*	0.18*	0.04													
YEB	0.70*	0.46*	0.16*	0.27*	0.39*	0.23*	0.00	0.02												
HOB	0.73*	0.51*	0.13*	0.31*	0.44*	0.27*	0.01	0.02	0.01											
HAT	0.69*	0.58*	0.14*	0.39*	0.52*	0.35*	0.01	0.03	0.04	0.03										
CAT	0.74*	0.21	0.44*	0.06	0.16*	0.05	0.18*	0.10	0.09	0.12*	0.22*									
SAG	0.74*	0.15	0.50*	0.03	0.11	0.03	0.23*	0.14*	0.14	0.18*	0.27*	0.01								
TBR	0.74*	0.63*	0.12*	0.45*	0.57*	0.40*	0.02	0.06	0.06	0.04	0.01	0.28*	0.33*							
BGR	0.81*	0.65*	0.08*	0.48*	0.60*	0.44*	0.03	0.13*	0.08	0.07	0.06	0.30*	0.35*	0.05*						
LOR	0.86*	0.71*	0.01	0.53*	0.65*	0.49*	0.05	0.15*	0.09*	0.07	0.06	0.34*	0.41*	0.04	0.04					
RER	0.86*	0.71*	0.01	0.53*	0.65*	0.49*	0.05	0.14*	0.09	0.06	0.04	0.35*	0.41*	0.03	0.04	0.01				
LAR	0.90*	0.74*	0.01	0.56*	0.69*	0.53*	0.08*	0.19*	0.11*	0.09	0.09*	0.38*	0.44*	0.08*	0.05	0.01	0.01			
SOR	0.81*	0.65*	0.09*	0.47*	0.59*	0.44*	0.02	0.13*	0.07*	0.07*	0.07	0.29*	0.35*	0.06*	0.05	0.05	0.05*	0.07		
DAL	0.84*	0.65*	0.04*	0.47*	0.59*	0.43*	0.03	0.12*	0.04	0.02	0.07*	0.26*	0.33*	0.06	0.04	0.01	0.01	0.01	0.04	

* P<0.05 following sequential Bonferroni correction.

Table 4.4. AMOVA results for three groups (experimental, hatchery, wild) of 20 common carp populations base on SSCP data. (Intra = intrapopulation, Inter = interpopulation, values are %)

Group	Intra	Inter	*P*-value	*F* st
Experiment	13.70	86.30	0.01	0.86
Hatchery	80.47	19.53	0.01	0.20
Wild	96.16	3.84	0.04	0.04
All	63.02	36.98	0.01	0.37

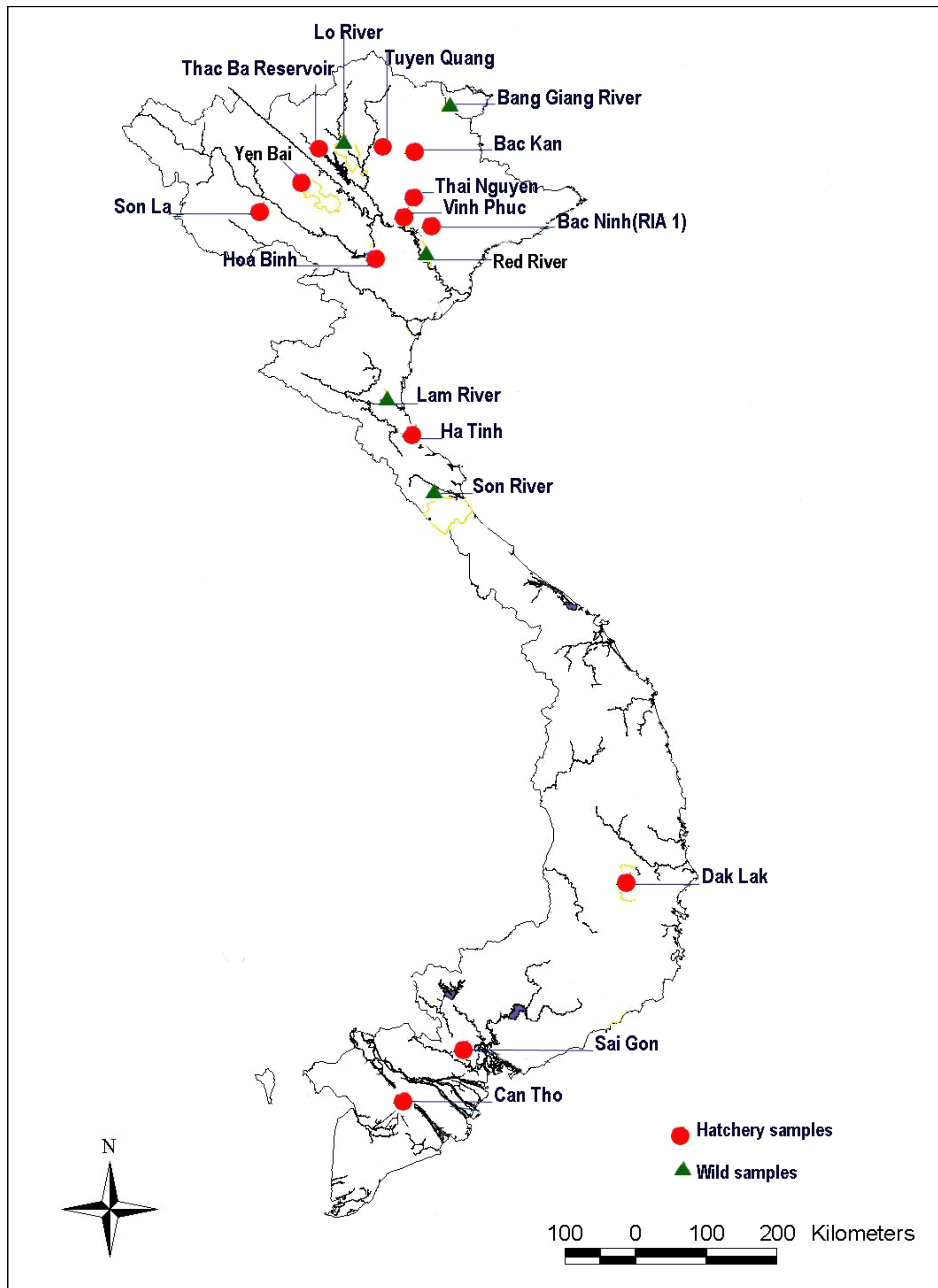

Figure. 4.1. Collection localities for *Cyprinus carpio* L. samples in Vietnam

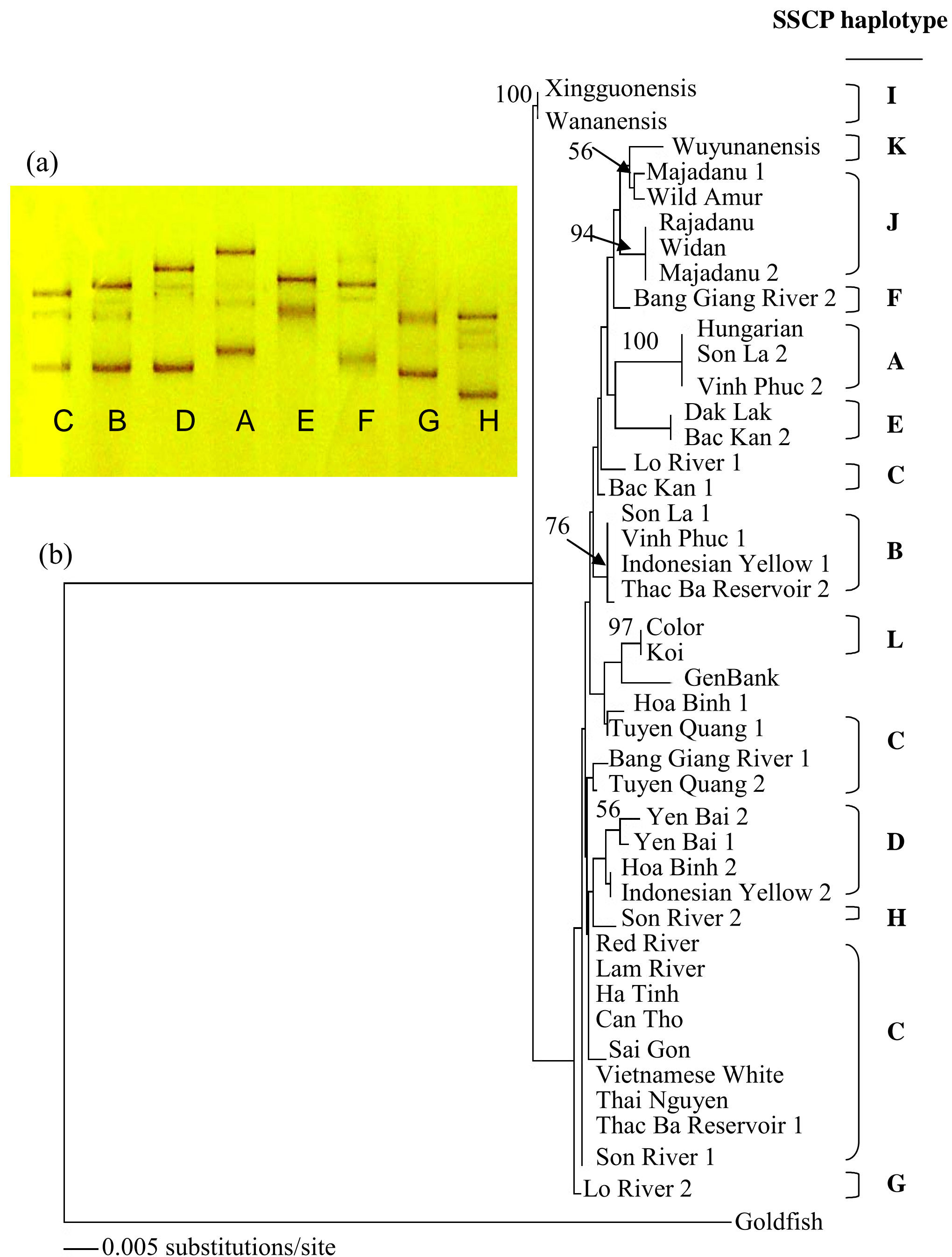

Figure. 4.2. Silver stained polyacrylamide gel showing the eight SSCP variants detected in common carp populations in Vietnam (a). Neighbour-joining tree reconstruction derived from *CR* sequences, using HKY+I+G model of evolution. Bootstrap values are based on 1,000 replicates. Bootstrap value is given for nodes with at least 50% or more support (b).

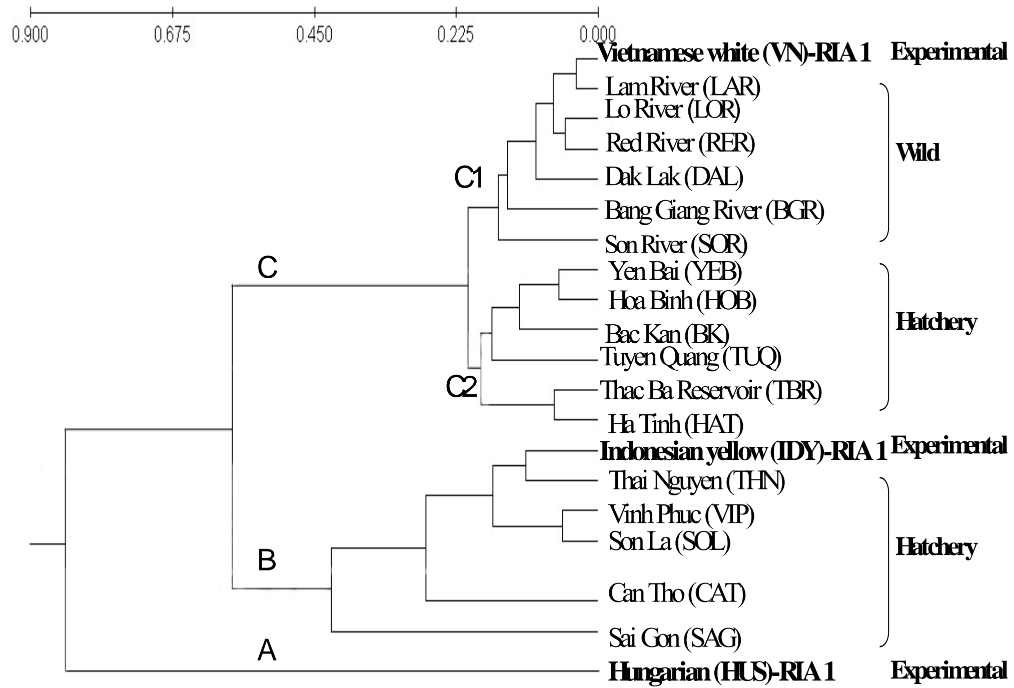

Figure. 4.3. Relationships among common carp from wild and hatchery population in Vietnam using the unbiased genetic distance of Rogers (1972) and UPGMA joining method. A, B, and C are SSCP haplotypes which predominate in each cluster.

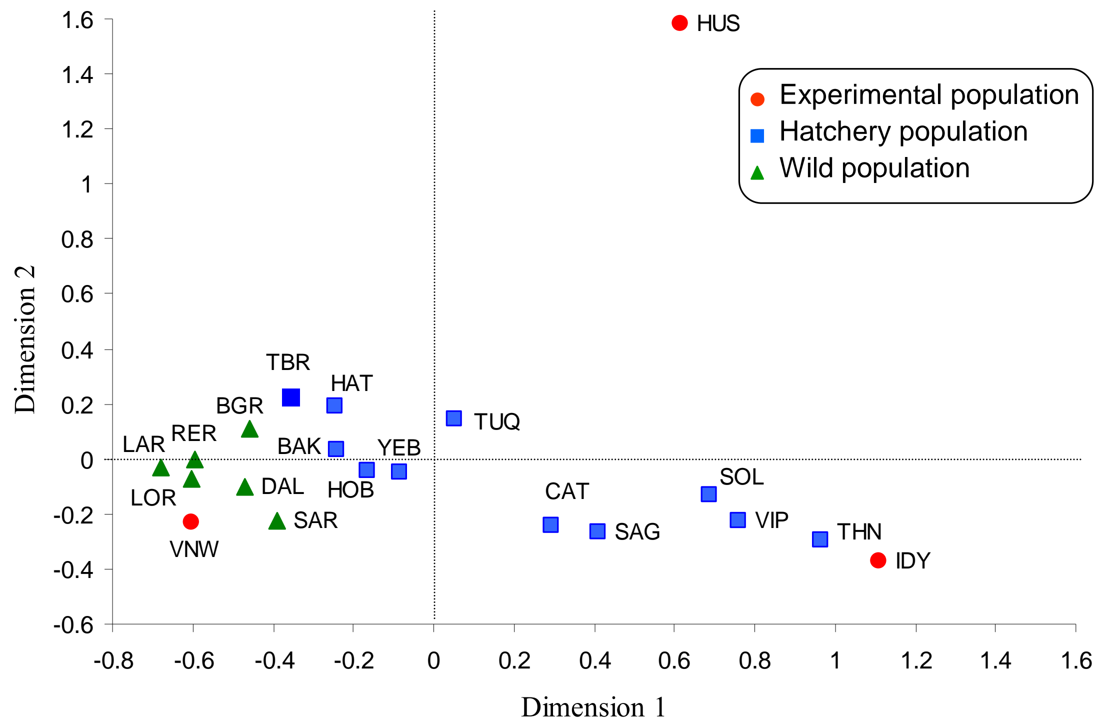

Figure. 4.4. MDS plot of pairwise *F*st value among hatchery and wild populations of common carp in Vietnam. Population codes are given in Table 4.1.

Chapter 5.

Genetic diversity of common carp (*Cyprinus carpio* L.) in Vietnam using four microsatellite loci*

5.1 Introduction

The understanding of genetic diversity is one of the most important steps in managing fisheries resources and aquaculture selective breeding programs (Beaumont, Hoare, 2003; Dunham, 2004; Ward, Grewe, 1995). Aquaculture practices may inadvertently decrease the genetic variability present in farmed stocks by breeding among related individuals or by the use of small numbers of founding broodstock. Selective breeding programs can also lead to decreased diversity when they utilize only a small number of "superior" families that may be related or use a mass selection approach with high selection intensities. Unless pedigree records are maintained, there is often a probability of selecting related individuals as parents for constructing the next generation and thereby increasing inbreeding (Norris *et al.*, 1999).

Conversely, breeding programs may deliberately introduce divergent stocks and utilize crossbreeding programs to increase diversity and productivity, via, for example, hybrid vigour (Hulata, 1995).

---

* A peer reviewed publication was derived from this chapter: Thai, T. B., C. P. Burridge and C.M. Austin (2007) Genetic diversity of common carp in Vietnam using four microsatellite loci. *Aquaculture* (in press)

However, the extent of stock mixing, the relative survival of the different stocks and the extent to which they are disseminated are important issues that frequently need to be addressed for effective management of aquaculture species (Dunham, 2004). There is also often a need to evaluate the status of wild stocks in aquaculture species as escapes of aquaculture stocks are common and these fish can have negative effects on resident indigenous forms (Cross, 2000).

Many of these issues apply to the common carp, *Cyprinus carpio* L., arguably the most important and oldest cultured fish species in the world, and providing a valuable source of protein especially in rural communities in many developing countries. The species has been widely translocated both within and outside its natural range leading to uncertainties concerning the genetic quality and origin of domesticated stocks and the status of wild populations (Kohlmann *et al.*, 2005; Lehoczky *et al.*, 2005).

In Vietnam, common carp has a natural distribution restricted to the north but is now distributed throughout the country as a result of translocations for aquaculture (Nguyen, Ngo, 2001). Enhancement of common carp production in Vietnam has focused substantially upon the development of genetically improved strains (Tran, Tran, 1995). For this purpose, Hungarian and Indonesian yellow carp strains were imported into Vietnam almost 30 years ago for crossbreeding and mass selection programs with local Vietnamese white carp. The genetic improvement strategy adopted for common carp in Vietnam involved the development of hybrid common carp by crossbreeding among three genetic lines (Vietnamese white, Hungarian scale

and Indonesian yellow common carp) coupled with mass selection (Nguyen *et al.*, 2005; Tran, Nguyen, 1992); (see 1.1.4). Aquaculture trials have demonstrated that hybrid common carp grew much faster than Vietnamese common carp. However, experimental lines of common carp are thought to have lost their purity, due to inadvertent mixing (Nguyen *et al.*, 2005). Recently, the improvement strains (three bloods) grow slowly in the farms. It may result of breeding depressions.

The ongoing development of common carp farming in Vietnam requires greater understanding of the genetic status of domesticated stocks. To this end an initial study of genetic variation within and among populations using mitochondrial markers was undertaken (Thai *et al.*, 2006). While this study provided useful insights into variation within and between wild and hatchery stocks and the dissemination of introduced stocks and genetic improved lines, mtDNA only represents a single locus and the maternal gene pool.

Microsatellites are highly variable nuclear genetic markers, which are inherited co-dominantly in a Mendelian fashion and, as such, offer an alternative perspective to mtDNA data. Microsatellites have been found suitable for a variety of applications in fisheries and aquaculture research, particularly where genetic differentiation within and between populations may be limited. Potential applications in aquaculture include monitoring changes in genetic variation as a consequence of different breeding strategies, the investigation of interactions between wild and cultured populations, parentage assignment and estimation of relatedness between potential breeding pairs (Cross, 2000; Cross *et al.*, 2005; Davis, Hetzel, 2000; Liu, Cordes, 2004; Norris *et al.*, 1999).

In recent years, a number of population genetic studies on common carp have been reported using a variety of molecular markers including, allozyme (Brody *et al.*, 1979; Kohlmann, Kersten, 1999; Murakaeva *et al.*, 2003) RAPDs (Bartfai *et al.*, 2003), AFLPs (David *et al.*, 2001; Wang *et al.*, 2000), mtDNA sequencing (Mabuchi *et al.*, 2005; Thai *et al.*, 2004) and microsatellites (Kohlmann *et al.*, 2005; Lehoczky *et al.*, 2005). However, no studies have examined genetic variation in common carp populations in Vietnam using nuclear data.

In this study, genetic diversity of wild and cultured of common carp in Vietnam are investigated using four microsatellite loci, and the same samples as a previous study that utilized mtDNA data (Chapter 4, Thai *et al*. 2006). The results of this study, together with the mtDNA data, provide important new insights into the management of Vietnamese common stocks and the selective breeding and dissemination programs for this species in this country.

5.2 Materials and Methods

5.2.1 Sample collection and DNA extraction

Nine hundred and sixty eight samples of common carp were collected from six rivers or reservoirs, 11 hatcheries and three experimental lines of common carp (Hungarian, Vietnamese white, Indonesian yellow common carp from the RIA 1, Bac Ninh, Vietnam). Fish were captured using lift and seine nets and tissue samples were obtained as fin clips and preserved in 90% ethanol. Total genomic DNA was isolated from the fin clips using the method of Crandall *et al*. (1999). Details of sample collection for common carp are given in Table 4.1 and Fig. 4.1.

5.2.2 Data collection

Four microsatellite loci *MFW1*, *MFW6*, *MFW7*, and *MFW9* (Crooijmans *et al.*, 1997) were PCR amplified with conditions comprising 1 X PCR buffer, 0.2mM dNTPs, and 0.2 unit *Taq* DNA polymerase (Invitrogen) with $MgCl_2$ concentration as detailed in Table 5.1. For all loci, dye label incorporation followed Schuelke (2000); the forward primer 5' was appended with an 18 bp M13 sequence (TGT AAA ACG ACG GCC AGT), and employed at 0.03 µM, while the reverse primer and a FAM- or HEX- labeled "M13"primer were employed at 0.5 µm. Thermal cycling conditions for each locus were: 3 min at 94° C, followed by 8 cycles of 94 ^{o}C for 30 s, annealing temperature (Table 5.1) for 30s, and 72 ^{o}C for 1 min, followed by 33 cycles as before but annealing at 53 ^{o}C, and then a final extension of 72 ^{o}C for 5 min. PCR products were separated on a 6% denaturing polyacrylamide gel using and ABI 373 (Applied Biosystems Inc) following manufacturer's instructions. Lengths of PCR products were determined relative to the GS400 size standard (ABI).

5.2.3 Genetic diversity analysis

5.2.4 Variation at microsatellite loci and tests from Hardy-Weinberg equilibrium

Allelic variation at the four microsatellite loci in the 20 populations was determined as number of alleles per locus (A) and heterozygosity (H). Heterozygosity and allelic frequencies for each population at each locus were calculated directly from microsatellite phenotypes using GENEPOP version 3.1 (Raymond, Rousset, 1995).

To test for departures from Hardy-Weinberg equilibrium (HWE) comparisons were made between observed heterozygosity (Ho), and expected heterozygosity (He) using exact tests as implemented by GENEPOP 3.1. This software employs the Markov chain method to estimate the probability of significant deviation from HWE using the following parameters: dememorization = 10,000, batches = 500 and iterations = 1,000. Levels of significance were adjusted for the number of simultaneous tests using the sequential Bonferroni procedure (Rice, 1989).

5.2.5 Genetic differentiation between populations

Genetic differences between populations was evaluated by calculating pairwise F_{st} values and testing their significance by bootstrapping analysis (1,000 replicates) using ARLEQUIN 2.000 (Schneider *et al.*, 2000). This program was also used to partition variation within and between populations using AMOVA procedure. Genetic differentiation between three groups (experimental, hatchery, wild) carp were also evaluated using F-STAT (Goudet, 2002). An assignment test was implemented by GeneClass v.1.0.02 software in order to determine the extent to which individuals could be correctly assigned to their population of origin (Piry *et al.*, 2004) as a measure of population differentiation and as a means of investigating population mixing.

To examine the genetic relationships among populations, a matrix of pairwise D_A distance (Nei *et al.*, 1983) was first calculated using the Microsatellite Analyser (MSA) program (Daniel, Christian, 2003). The distance matrix was then used to construct a UPGMA dendrogram using the software package PHYLIP 3.5c of Felsenstein (1993). Multidimensional Scaling (MDS) as implemented by SPSS

(Version 10), was also used to examine relationships based on this matrix as this procedure can better represent non-hierarchical patterns of relationships among populations (Shepard, 1974).

5.3 Results

5.3.1 Within population variation

All four loci were polymorphic and were variable in all populations (Table 5.2). A total of 72 different alleles ranging in size from 100 to 262 bp were found over the four loci. The number of alleles ranged from 10 at *MFW9* to 23 at *MFW1* and with from three to 15 alleles per population per locus.

Within populations, the lowest mean number of alleles per locus (4.25) was observed in the Indonesian yellow common carp experimental line (IDY), while the highest mean number of alleles per locus (11.00) was found in the wild Red River population (RER). Average observed heterozygosity ranged from 0.40 in the Indonesian yellow carp experimental line to 0.83 in the Red River population. Average allelic diversity was the lowest in the experimental lines (5.50-8.25), highest in the wild populations (8.75-10.00) and generally intermediate in the hatchery populations (6.50-9.50). Average observed heterozygosity showed a similar trend with the experimental lines having the lowest average observed heterozygosity (0.40-0.59), the wild populations the highest (0.77-0.83) and the hatchery populations again with generally intermediate value (0.51-0.81) (Table 5.2).

Of the 80 test of deviation from Hardy-Weinberg equilibrium (HWE) proportion, 37 were significant and all but five of the tests were associated with heterozygote deficiencies. A much greater proportion of significant HWE tests occurred within the experimental lines (8 of 12) and the hatchery stocks (27 of 44) compared with the wild population (7 of 24). Based on average *F*is values it can be seen that the pattern of heterozygotes deficiencies was most pronounced in the experimental lines (0.19-0.26). Heterozygote deficiencies were also apparent in the hatchery populations with the exception of Thac Ba Reservoir (TBR) but not to the same degree (*F*is = 0.05-0.22). The wild population showed minimal heterozygote deficits (-0.11-0.05).

The proportion of private alleles showed the converse pattern to the *F*st values. Across the four loci only three private alleles occurred within the three experimental lines and only 12 within the 11 hatchery populations, which compares with 19 private alleles in the six wild populations (Table 5.3).

5.3.2 Genetic differentiation and relationships among populations

Pairwise *F*st analyses indicates significant genetic heterogeneity among populations with the majority of pairwise comparisons yielding signification differences (Table 5.4). The three experimental strains were well differentiated from each other (*F*st = 0.16-0.34). The Hungarian strain was the most divergent (*F*st= 0.10-0.34) followed by the Indonesian strain (*F*st= 0.05-0.21). While the Vietnamese experimental strain was significantly different from most of the other samples *F*st value were generally lower (Table 5.4).

Allelic frequencies at the four loci for the three experimental lines are depicted in Fig. 5.1. The differences between the three experimental lines are largely a matter of degree and none of the loci provide a profile that is diagnostic for any of the three experimental lines. Nevertheless, some loci are more effective than other in distinguishing particular strains. For example, the Hungarian strain has almost exclusively small sized alleles at locus *MFW7* compared to the Indonesian and Vietnamese strain. The distinction between the three strains is generally a combination of allelic difference that accumulates across the loci. This is most marked in the Indonesian strain which has one or two alleles at high or moderately high frequencies at *MFW1*, *MFW6* and *MFW7* that are absent or at low frequencies in the other two experimental lines. The Vietnamese strain is distinguished substantially by a large number of private alleles occurring at low frequencies that are spread across all loci. The ability of allelic variation at these loci as a group to distinguish between three strains is demonstrated by assignment test using just these strains. This test resulted in only 15 (10%) of the 150 individuals being misclassified.

Levels of differentiation were limited among the wild population and similarly with that hatchery samples except for samples BAK and TBR. The UPGMA dendrogram emphasises the distinctiveness of the Hungarian sample which forms the most basal branch (A) (Fig. 5.2). The remaining samples form two distinct clusters (B) and (C). One (C) contains the Indonesian sample, which is the most distinct of all samples in this cluster and all but two of the hatchery samples. The other cluster (B) contains the Vietnamese experimental line, all the wild populations and two of the hatchery samples (BAK and TBR). The MDS analysis (Fig. 5.3) reflects these same relationships and emphasizes the distinctiveness of the three experimental lines and

the distinctiveness of the Hungarian strain. In addition this analysis indicates that the hatchery samples (excluding BAK and TBR), while generally closest to the Indonesian experimental line fall into an intermediate position between the Vietnamese and Hungarian samples. Some of these samples are almost equidistant between the experimental lines. For example HAT is almost exactly halfway between the Vietnamese and Indonesian experimental lines and Tuyen Quang (TUQ) is almost equidistant between the Hungarian and Indonesian samples (Fig. 5.3).

The results of the assignment test using all 20 populations confirms and extends the population genetic (*F*st) and phylogenetic genetic distance-based analyses (Table 5.5). Overall there is a relatively high proportion of misclassifications, reflecting the generally limited divergence among populations. However the pattern of misclassification differs among the three groups. The highest proportion of correct classification is among the experimental strains and the wild populations, although the Vietnamese white carp (VNW) experimental strains and wild population exchange a relatively higher proportion of misclassifications. The hatchery populations as a group suffer the highest proportion of misclassifications. With the exception of samples TBR and BAK, there is a consistent pattern of misclassification, involving the other hatchery stocks and the Indonesian and Hungarian experimental lines. There was also a small but consistent amount of misclassification between the hatchery samples and the Vietnamese experimental lines and wild samples. The samples TBR and BAK differ from the other hatchery samples as the extent of misclassification between the Indonesian or Hungarian lines and hatchery stocks on the one hand and the Vietnamese experimental line and wild stocks on the others, are about equal.

The AMOVA analysis is consistent with the previous analysis and highlights the differences between the partioning of genetic variation between the experimental lines and wild population (Table 5.6). For the experimental group the proportion of inter-population variation (23.80%) is much greater than for the wild populations (1.03%). The distribution of intra- and inter-population variation in the hatchery samples is intermediate between the experimental and wild populations but more similar to the latter group. From Table 5.6 it can be seen that the hatchery samples are also intermediate with respect to allelic richness and observed heterzygosity.

5.4 Discussion

5.4.1 Genetic diversity within common carp populations

Genetic diversity is important to both natural and cultured population because it provides the necessary spectrum of genotypes for adaptive response to changing conditions and heterozygous individuals usually are superior to less heterozygous individuals in many economically important characteristics like growth, fertility and disease resistance (Beardmore *et al.*, 1997). As a consequence, there has been increasing attention being paid to loss of genetic diversity in domesticated fish stocks including carp. Desvignes *et al*. (2001), Bartfai *et al*. (2003), and Kohlmann *et al*. (2005) have analysed microsatellite variation in common carp, however all three studies have limitations due to the sampling of restricted number of populations or the used of small sample sizes. Nevertheless the levels of variation detected are broadly similar to the results of this study.

Kohlman *et al*'s (2005) study is by far the most comprehensive of these, and also looked at variation at four microsatellite loci in common carp from 22 populations from mostly Western Europe found similar levels of intrapopulation genetic variation to this study. The average number of allele per locus per population range from 2.50-14.00 in Kohlmann *et al*'s (2005) study which compares with 4.25-11.00 in this study. Kohlmann *et al*. (2005) also found significant differences in allelic richness (Ar) between domestic populations (4.44) and wild caught populations (8.24) which is very similar to the values in this study for the experimental lines (5.83) and the wild populations (9.26). The loss of genetic diversity in hatchery populations is also found in other aquaculture species such as channel catfish (Mickett *et al.*, 2003; Simmons *et al.*, 2006), Atlantic salmon (Elliott, Reilly, 2003; Innes, Elliott, 2006), cutthroat trout (Allendorf, Phelps, 1980), and abalone (Evans *et al.*, 2004). Loss of variation in closed hatchery populations can occur during establishment (founder effect) and over subsequent generations though genetic drift arising from low effective broodstock number (Allendorf, Phelps, 1980). The large reduction in genetic variability in the experimental lines observed in this study by Thai *et al*. (2006) and cultured carp in Europe by Kohlmann *et al*. (2005) indicate the potential negative impact of genetic improvement programmes and captive breeding on domesticated common carp stocks in Vietnam and elsewhere. The large number of deviation from HWE proportions in experimental lines and hatchery stocks due to a deficiency in heterozygotes is consistent with a Wahlund effect which results from the mixing of genetically diverse stocks (Ridley, 2003).

This study found diversity levels in the majority of hatchery populations in Vietnam to be similar (A = 8.0) to the wild carp populations and those studied by Kohlmann *et*

al. (2005) is unusual. Such findings are usually attributed to the mixing of stocks during founding and subsequent propagation (Ferguson, 1995b; Thompson, 1985). In the case of hatchery stocks of common carp in Vietnam, this elevated diversity can be attributed to the successful dissemination of genetically improved carp lines that were generated by crossbreeding with imported stocks. These results and conclusions are very similar to those based on the mtDNA data set derived from the same populations and the two data sets produce remarkably similar results which will be discussed in more detail below.

5.4.2 Differentiation between populations

While significant differences were detected in allelic frequencies between many pairwise combinations of populations, levels of differentiation as measured by *F*st are generally less than observed by Kohlmann *et al*. (2005). Kohlmann's study however, included much wider geographic sampling of populations from Europe and central Asia. Importantly, the microsatellite variation was effective in separating the Vietnamese experimental line and wild populations from the introduced Hungarian and Indonesian lines. Thus, it is possible to determine the affinities of the hatchery stocks and the extent to which they represent mixed stocks. From the MDS and the assignment test, it is apparent that all but two of the hatchery stocks represent mixtures of all three experimental lines with the Indonesian line predominating, followed by the Vietnamese and the Hungarian line having the least influence. Two hatchery populations BAK and TBR stand out as distinct as they cluster closely with the Vietnamese samples. This is consistent with the known history of these populations as the BAK sample is from a small private hatchery in the highland that utilizes local broodstock and has not imported fish from central hatcheries. The population TBR is from fish caught in Thac Ba reservoir, which, although stocked

with fish from Government hatcheries, also maintains an indigenous carp population which clearly dominates this sample.

The finding that the majority of hatchery stocks had elevated diversity levels compared with inbred experimental lines and were genetically intermediate is consistent with stock mixing among the three experiment lines and the known genetic history of carp in Vietnam. Research by the Vietnamese Government into the genetic improvement of carp over the last 15 years has led to the development of three genetically improved lines through a combination of crossbreeding and mass selection. These lines are referred to as "three blood" carp as all three strains were used over two generations to produce carp comprising a genetic mixture of the strains in a 2: 1: 1 ratio. The three strains even though derived by hybridization , are referred to as the "Indonesian", "Hungarian" and "Vietnamese" improved strains on the basis of the female parent used to construct the second generation of the hybrid crosses (Tran, Nguyen, 1993).

These hybrid lines have been preferentially disseminated to Government provincial hatcheries (e.g. sample VIP, SOL, HAT, THN) which is consistent with our results and indicates that the Vietnamese Government's crossbreeding strategy has been effective in preventing the erosion of genetic variation which is clearly prevalent in many domesticated lines of carp (Kohlmann *et al.*, 2005). An important activity for the future is to investigate the genetic affinities and diversity of carp stocks used by small scale farms and from the market places to determine the extent to which three blood carp lines have been disseminated beyond provincial hatcheries. Another important issue to be addressed is the relative performance of indigenous carp stock

and the three blood lines in aquaculture ponds. While the three blood lines potentially represent a richer array of genotypes, they may not perform as well as locally adapted genotypes present in indigenous and possibly even more inbred stocks. An effective way to examine these questions would be to undertake communal rearing experiments using both three blood carp lines and local farmed carp in farmers' ponds.

5.4.3 Comparison of microsatellite and mitochondrial DNA data

The use of multiple data sets and information from different molecular markers is becoming more common in aquaculture research (Davis, Hetzel, 2000; Liu, Cordes, 2004). An example in common carp are Kohlmann *et al*'s studies which have employed allozymes, mtDNA and microsatellites (Kohlmann *et al.*, 2003; Kohlmann, Kersten, 1999; Kohlmann *et al.*, 2005) to study variation and taxonomic questions. Confidence in the conclusions and insights derived from the patterns of variation at the four microsatellite utilized in this study are greatly increased due to the high level of concordance with mtDNA data collected from the same populations (Thai *et al.*, 2006). Indeed, the level of congruence in the interpopulation relationships derived from each type of marker as depicted by MDS plots is remarkable (Fig. 5.3; 4.4). In both analyses the first axis separates the Vietnamese and Indonesian experimental lines with extreme negative and positive score respectively. Similarly both analyses highlight the distinctiveness and isolation of the Hungarian experimental line which has extremely high positive score on axis two compared with the other 19 samples. The close relationship between the Vietnamese experimental line and the wild populations and BAK and TBR hatchery populations and the Indonesian experimental line and the CAT, SAG, SOL, VIP, THN hatchery populations are also

supported by both data sets. A contrast in the two analyses is the placement of the HAT, HOB, TUQ, and YEB hatchery samples which are close to the Vietnamese wild samples in the mtDNA based analyses but are aligned with the main group of hatchery samples and experimental Indonesian line based on microsatellite data. These differences may reflect asymmetry in the sex ratio of the parents of the stocks disseminated to the hatcheries with more female than male Vietnamese carp used to generate the hybrid lines.

There are also other noteworthy similarities and differences between the data sets. While both kinds of data were effective at distinguishing the experimental strains the extent of differentiation was much greater for the mtDNA (*F*st = 86.30%) compared with the microsatellite data (*F*st = 23.80%). MtDNA data is expected to be more effective at detecting differences of this kind as its effective population size is 1/4 of nuclear DNA and therefore mtDNA haplotypes will drift towards fixation much more quickly than microsatellite alleles (Ward, Grewe, 1995). The greater differentiation of the experimental strains in mtDNA profiles, and because they are largely characterized by different haplotypes, means that the mtDNA is more effective at detecting population mixing and the origin of the stocks contributing to mixing. The analysis of microsatellite data using assignment tests is also effective in this regard and its power can be enhanced by examining additional loci, an option that is not open to mtDNA as all mtDNA markers are linked.

The mtDNA, due to the greater divergence among experimental lines revealed higher levels of diversity among hatchery populations compared to the wild populations. For the microsatellite data the converse was true, which illustrates the point that even

a crossbreeding strategy to increase diversity can fail if the degree of divergence between the crossbred strains is not substantial and the parental strains have diminished variation due to inbreeding. Noteworthy is the significantly reduced proportion of private alleles in the hatchery populations compared with wild populations. Thus, if the breeding goal is simply to maximize diversity of farmed carp, it would be more effective to obtain a reasonably large sample of Vietnamese wild carp for broodstock, rather than crossbreeding between introduced lines of common carp which may not generate novel genotypes, that are superior to those produced by local carp stocks under local conditions.

Table 5.1. Characteristics of *Cyprinus carpio* microsatellite loci tested

Locus	Primer	Size (bp)	Annealing temp. (T °)	$MgCl_2$
MFW1	F13: GTC CAG ATC GTC ATC AGG AG R: GTT TGA GGT GTA CAC TGA GTC AGG C	184 -230	62	1.5
MFW6	F13: ACC TGA TCA ATC CCT GGC TC R: GTT TGG GAC TTT TAA ATC ACG TTG	130 - 219	62	1.5
MFW7	F13: TAC TTT GCT CAG GAC GGA TGC R: GTT TAT CAC CTG CAC ATC GCC ACT C	192 -262	62	1.5
MFW9	F13: GATCTGCAAGCATATCTGTCG R: GTTTATCTGAACCTGCAGCTCCTC	92 - 144	60	1.5

Table 5.2. Genetic variability of four microsatellite loci in 20 populations for common carp in Vietnam. Population codes given in Table 4.1.

Locus	Parameter	HUS	IDY	VNW	VIP	THN	SOL	BAK	TUQ	YEB	HOB	HAT	CAT	SAG	TBR	BGR	LOR	RER	LAR	SOR	DAL	Average cross population
MFW1	*N*	50	50	50	50	50	50	50	50	50	50	50	36	34	50	50	50	50	50	50	50	48.50
	A	6	4	9	10	10	11	13	7	12	11	11	11	7	12	11	15	12	11	12	14	10.45
	He	0.62	0.57	0.82	0.83	0.85	0.88	0.84	0.80	0.81	0.81	0.83	0.79	0.79	0.85	0.88	0.87	0.90	0.87	0.80	0.86	0.81
	Ho	0.48	0.24	0.46	0.48	0.44	0.46	0.62	0.44	0.60	0.48	0.58	0.56	0.47	0.88	0.76	0.86	0.72	0.80	0.68	0.78	0.59
	P_{HW}	**	**	*	**	*	*	*	**	*	*	**	*	*	n.s.	n.s.	n.s.	*	n.s.	*	*	
	Fis	0.21	0.53	0.41	0.41	0.47	0.46	0.24	0.44	0.24	0.39	0.27	0.27	0.35	-0.04	0.12	0	0.18	0.06	0.02	0.08	0.26
MFW6	*N*	50	50	50	50	50	50	50	50	50	50	50	36	34	50	50	50	50	50	50	50	48.50
	A	4	7	10	9	8	8	10	6	11	8	9	5	7	11	14	11	11	10	11	10	9.00
	He	0.74	0.69	0.86	0.69	0.73	0.62	0.86	0.74	0.77	0.72	0.78	0.59	0.74	0.85	0.85	0.82	0.82	0.85	0.81	0.83	0.77
	Ho	0.60	0.68	0.56	0.78	0.80	0.62	0.68	0.64	0.84	0.64	0.72	0.64	0.85	0.68	0.84	0.83	0.86	0.68	0.70	0.74	0.72
	P_{HW}	*	n.s.	**	**	*	n.s.	*	*	*	**	n.s.	*	n.s.	*	n.s.	n.s.	n.s.	*	*	n.s.	
	Fis	0.17	0	0.33	-0.14	-0.11	-0.01	0.19	0.12	-0.09	0.96	0.06	-0.1	-0.18	0.18	0.1	-0.01	0.05	0.18	0.28	0.09	0.10
MFW7	*N*	50	50	50	50	50	50	50	50	50	50	50	36	34	50	50	50	50	50	50	50	48.50
	A	9	3	10	9	12	15	10	6	11	11	11	10	8	10	11	12	15	9	10	9	10.05
	He	0.70	0.27	0.84	0.68	0.59	0.72	0.82	0.51	0.78	0.49	0.64	0.50	0.34	0.72	0.84	0.86	0.86	0.79	0.79	0.84	0.68
	Ho	0.52	0.22	0.60	0.56	0.34	0.32	0.44	0.28	0.44	0.22	0.32	0.42	0.21	0.88	0.82	0.74	0.82	0.84	0.82	0.80	0.53
	P_{HW}	**	n.s.	*	*	*	*	*	*	*	*	*	n.s.	n.s.	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.	
	Fis	0.24	0.09	0.28	0.16	0.42	0.53	0.46	0.42	0.41	0.53	0.4	0.13	0.34	-0.22	0.01	0.12	0.06	-0.07	-0.04	0.03	0.22
MFW9	*N*	50	50	50	50	50	50	50	50	50	50	50	36	34	50	50	50	50	50	50	50	48.50
	A	3	3	4	4	6	4	4	4	4	4	5	5	4	5	4	5	6	5	5	4	4.40
	He	0.62	0.58	0.74	0.54	0.64	0.62	0.75	0.61	0.63	0.69	0.69	0.63	0.55	0.78	0.73	0.75	0.75	0.83	0.86	0.74	0.69
	Ho	0.42	0.46	0.72	0.58	0.60	0.66	0.74	0.76	0.68	0.74	0.94	0.58	0.53	0.78	0.76	0.78	0.88	0.92	0.98	0.80	0.72
	P_{HW}	*	n.s.	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*	n.s.	n.s.	n.s.	n.s.	
	Fis	0.3	0.13	0	-0.11	0.04	-0.78	-0.01	-0.25	-0.09	-0.07	-0.37	0.06	0.02	-0.01	-0.05	-0.56	-0.17	-0.61	-0.48	-0.09	-0.16
Mean (all loci)	*A*	5.50	4.25	8.25	8.00	9.00	9.50	9.25	5.75	9.50	8.50	9.00	7.75	6.50	9.50	10.00	10.75	11.00	8.75	9.50	9.25	
	He	0.67	0.53	0.81	0.68	0.70	0.71	0.82	0.67	0.75	0.68	0.73	0.63	0.60	0.80	0.82	0.82	0.83	0.83	0.81	0.82	
	Ho	0.51	0.40	0.59	0.60	0.55	0.52	0.62	0.53	0.64	0.52	0.64	0.55	0.51	0.81	0.80	0.80	0.82	0.81	0.80	0.78	
	P_{HW}	**	*	**	*	*	**	*	*	*	*	*	*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	*	
	Fis	0.23	0.19	0.26	0.08	0.21	0.05	0.22	0.18	0.12	0.45	0.09	0.09	0.13	-0.02	0.05	-0.11	0.03	-0.11	-0.06	0.03	

(N = sample size; A= total number of alleles; He = expected heterozygosity; Ho = observed heterozygosity; P_{HW} = Hardy-Weinberg probability test: * P <0.05; ** P <0.01; *F*is = fixation indices; n.s. = non-significant)

Table 5.3. Number of private alleles at four microsatellite loci in 20 common carp populations and group. The population codes are given in Table 4.1

Population	Locus				Total
	MFW1	*MFW6*	*MFW7*	*MFW9*	
HUS	1				1
IDY		2			2
VN					0
VIP		1			1
TN			1	2	3
SOL			1	2	3
BAK	2				2
TUQ					0
YEB	1		2		3
HOB					0
HAT					0
CAT					0
SAG					0
TBR					0
BGR		1			1
LOR	1	2		2	5
RER					0
LAR	1			2	3
SOR	1		3		4
DAL	3				3
Experimental group	1	2	-	-	3
Hatchery group	2	1	4	4	12
Wild group	6	3	3	4	16

Table 5.4. Pairwise *F*st values between 20 common carp populations in Vietnam based on four microsatellite loci. Population codes given Table 4.1.

	HUS	IDY	VNW	VIP	THN	SOL	BAK	TUQ	YEB	HOB	HAT	CAT	SAG	TBR	BAG	LOR	RER	LAR	SOR	DAL
HUS																				
IDY	0.34*																			
VNW	0.16*	0.21*																		
VIP	0.20*	0.10*	0.13*																	
THN	0.17*	0.06*	0.10*	0.01																
SOL	0.18*	0.11*	0.10*	0.00	0.01															
BAK	0.16*	0.17*	0.00	0.10*	0.07*	0.07*														
TUQ	0.20*	0.07*	0.13*	0.01	0.00	0.01	0.10*													
YEB	0.19*	0.08*	0.08*	0.02*	0.01	0.02	0.05*	0.02												
HOB	0.22*	0.07*	0.11*	0.03*	0.01	0.01	0.08*	0.01	0.03*											
HAT	0.21*	0.05*	0.09*	0.03*	0.00	0.02	0.06*	0.01	0.01	0.01										
CAT	0.26*	0.06*	0.13*	0.02	0.01	0.01	0.09*	0.01	0.02	0.00	0.01									
SAG	0.29*	0.05*	0.15*	0.03*	0.02	0.03*	0.12*	0.01	0.03*	0.02	0.01	0.01								
TBR	0.18*	0.15*	0.02*	0.08*	0.06*	0.06*	0.01	0.08*	0.05*	0.07*	0.05*	0.08*	0.10*							
BAG	0.14*	0.20*	0.01	0.11*	0.08*	0.08*	0.01	0.11*	0.07*	0.10*	0.08*	0.12*	0.15*	0.01						
LOR	0.13*	0.18*	0.02*	0.09*	0.06*	0.07*	0.01	0.10*	0.06*	0.09*	0.06*	0.10*	0.12*	0.01	0.00					
RER	0.10*	0.19*	0.02*	0.09*	0.06*	0.07*	0.02	0.10*	0.06*	0.09*	0.07*	0.11*	0.13*	0.01	0.00	0.00				
LAR	0.18*	0.21*	0.03*	0.14*	0.10*	0.11*	0.03*	0.14*	0.08*	0.13*	0.09*	0.13*	0.16*	0.02	0.01	0.01	0.02*			
SOR	0.18*	0.16*	0.04*	0.10*	0.07*	0.08*	0.03*	0.10*	0.06*	0.09*	0.06*	0.09*	0.11*	0.02*	0.02*	0.02*	0.02*	0.01		
DAL	0.15*	0.19*	0.00	011*	0.08*	0.08*	0.00	0.11*	0.06*	0.10*	0.07*	0.11*	0.14*	0.01	0.00	0.00	0.01	0.01	0.02*	

* $P < 0.05$ following sequential Bonferroni correction. See Table 1 for samples code details.

Table 5.5. Results of assignment test (self-classification) of common carp individuals based on four microsatellite loci. Population code are given in Table 4.1.

	n	HUS	IDY	VNW	VIP	THN	SOL	BAK	TUQ	YEB	HOB	HAT	CAT	SAG	TBR	BAG	LOR	RER	LAR	SOR	DAL
HUS	50	45		2				1	1								1				
IDY	50		36										3	11							
VNW	50		2	20				5	1			1			1	3	2		4	4	7
VIP	50	6	18		8	1		2				1	2	12							
THN	50	4	16	1		10			1			2	2	10					1	3	
SOL	50	8	15	2			2	2			2	2	5	7				1	1		3
BAK	50	4	4	10	3			14	2			2	1	1			2		3	3	1
TUQ	50	12	11						10	1	2		4	9						1	
YEB	50	3	10	5				1		18				9			1		1	2	
HOB	50	3	11	5			1	2			20		2	5							1
HAT	50	10	5	3								11	3	10			1		3	2	2
CAT	36		11	3									12	8						1	1
SAG	35		10	4			1						2	16			1				1
TBR	50	8	3	5		1			1		1	1	6		12		7	1	1	2	1
BAG	50	1	1												4	30	10	1		3	
LOR	50	1	2	2												1	42			1	1
RER	50	2	2	3				1							1	2	4	34		1	
LAR	50	1	2	4															39	2	2
SOR	50			4														3	4	35	4
DAL	47			12					1				3	1	4		3		1		22

Table 5.6. Genetic diversity of experimental, hatchery and wild common carp groups in Vietnam based on variation at four microsatellite loci (Intra = intrapopulation, Inter = interpopulation, values are %).

Group	Intra	Inter	*P* -value	*F* st	A-richness	Ho
Experiment	76.18	23.8	<0.001	0.24	5.83	0.50
Hatchery	96.12	3.88	<0.001	0.03	8.00	0.59
Wild	98.97	1.03	<0.001	0.01	9.26	0.80
All	92.04	7.96	<0.001	0.08	-	-
P -value					<0.001	<0.001

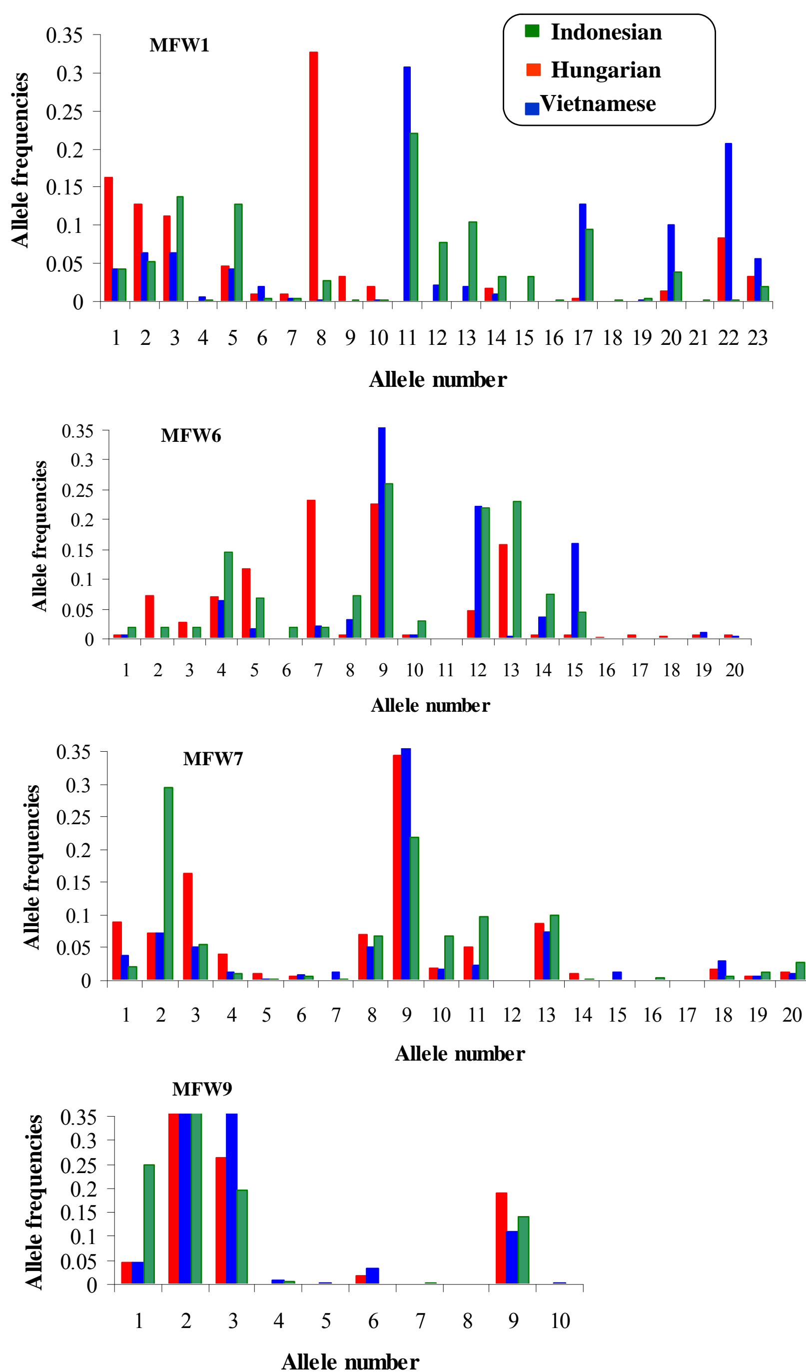

Figure. 5.1. Allele number and distribution of allele frequencies for four loci between Hungarian, Indonesian and Vietnamese common carp populations in experimental group.

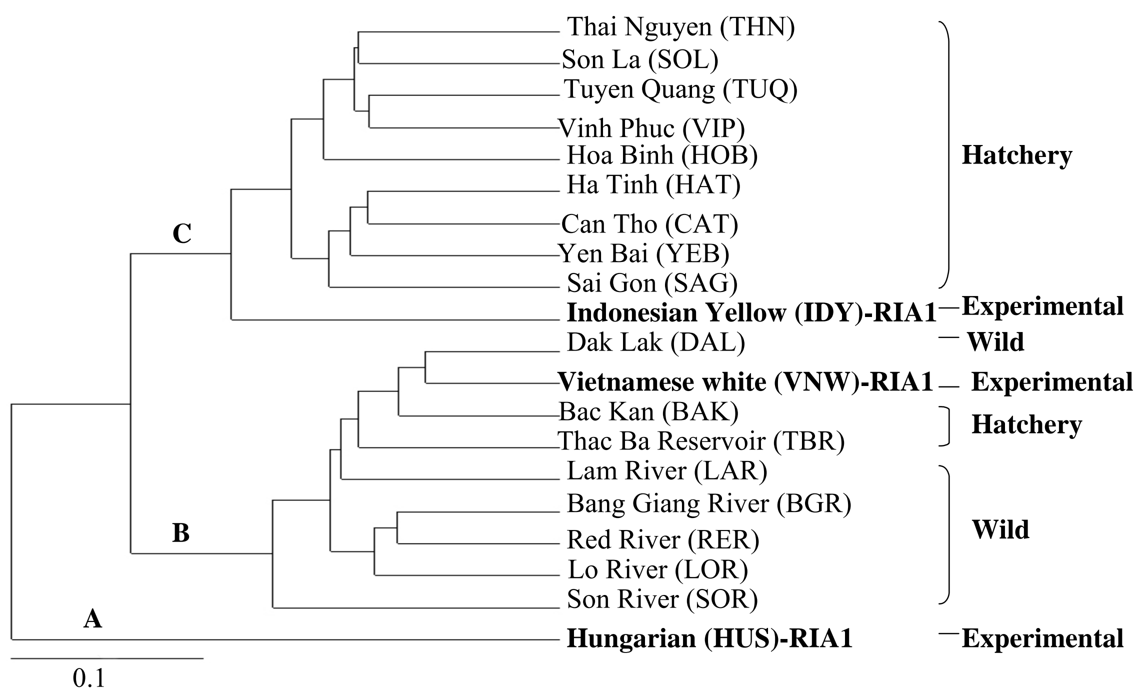

Figure. 5.2. UPGMA dendrogram of common carp populations in Vietnam based on matrixes of genetic distance (Nei *et al.*, 1983).

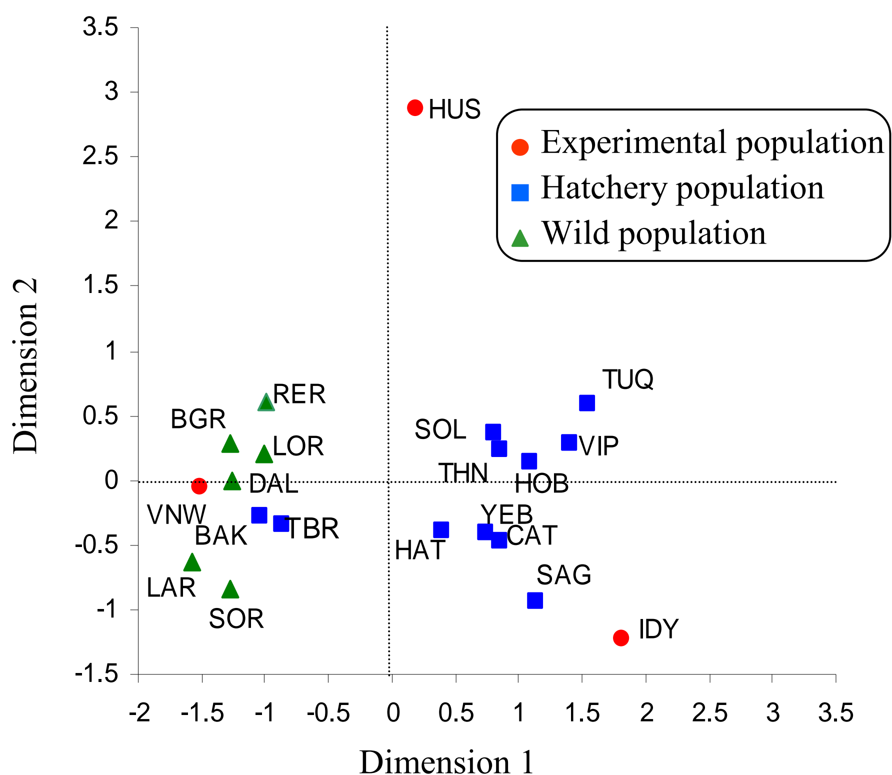

Figure. 5.3. MDS plot of genetic distance (Nei *et al.*, 1983) among experimental, hatchery and wild populations of common carp in Vietnam. Population codes are given in Table 4.1.

Chapter 6.

Phylogenetic evaluation of subfamily classification of the Cyprinidae, focusing on Vietnamese species

6.1 Introduction

The Cyprinidae is the largest freshwater fish family in the world with over 200 genera and 2,000 species (Liu, Chen, 2003). While the family has a relatively diverse fauna in Africa, Europe and North America, over 1,200 species are recorded from Asia with the centre of diversity being China and South East Asia (Liu, Chen, 2003). A large number of well known fish species belong to the Cyprinidae including the barbell, the barbs, the common carp, goldfish, chubbs and roach. The family also contains many species important to aquaculture and inland fish production with an annual world production over 17 million tones (FAO, 2003). The Cyprinidae is, thus, perhaps the most important taxonomic group of fish consumed by humans.

As with many Asian countries, Vietnam has an abundant cyprinid fauna with over 220 recognised species. Members of the family play an important role in aquaculture in Vietnam (Nguyen, Ngo, 2001). There are 13 indigenous and five introduced species that contribute to about 75% of inland fish production in the country. Cyprinids are mainly cultured in polyculture systems, the main species being silver carp (*Hypophthalmichthys molitrix* Cuvier & Valencienes 1844), grass carp (*Ctenopharyngodon idella* Steindachner 1866), bighead carp (*Aristicchthys nobilis* Richardson 1844), rohu (*Labeo rohita* Hamilton 1822), mrigala (*Cirrhinus cirrhosus* Hamilton 1822) and local fish species such as common carp (*Cyprinus carpio*

Linnaeus 1758) (Nguyen *et al.*, 2005). A number of Vietnamese cyprinids have restricted distributions and are threatened due to over-fishing, interbreeding between indigenous and introduced exotic species or translocated native species (Nguyen, Ngo, 2001), environmental degradation and anthropogenic changes such as construction of reservoirs and hydroelectric dams.

Taxonomically, the Cyprinidae have been divided into a greater or lesser number of subfamilies (Chen *et al.*, 1984; Nguyen, Ngo, 2001; Rainboth, 1996). For example, Chen *et al*. (1984), based on a cladistic analysis divided the cyprinids into 10 subfamilies (Labeoninae + Cyprininae + Barbinae + Tincinae + Acheilognathinae + Gobioninae + Xenocyprininae + Cultrinae + Leuciscinae + Danioninae or Rasborinae). In contrast, Rainboth (1996) divided the Cyprinidae into just four subfamilies (Alburinae + Danioninae + Leuciscinae + Cyprininae).

The taxonomic confusion and uncertainties within the cyprinids are evident by considering just the taxonomic treatment of Vietnamese cyprinids. Mai (1978) recognized 9 subfamilies (Cyprininae + Barbinae + Acheilognathinae + Gobioninae + Gobiobotinae + Xenocyprinae + Cultrinae + Leuciscinae + Hypophthalmichthyinae). In contrast, Truong and Tran (1993) and Mai *et al*. (1992) placed Vietnamese cyprinids just into four groups (Cyprininae + Abraminae + Rasborinae + Garrinae). Recently, Nguyen and Ngo (2001) divided cyprinids in Vietnam into 11 subfamilies (Labeoninae + Cyprininae + Barbinae + Acheilognathinae + Gobioninae + Gobiobotinae + Xenocyprinae + Cultrinae + Leuciscinae + Danioninae + Hypophthalmichthyinae). Such contrasting opinions on cyprinid classification hinder evolutionary, biogeographic and even comparative

studies, and are clearly undesirable for such a widespread and important group of fishes.

Molecular phylogenetic studies are increasingly being used to investigate cyprinid classification and evolution at a variety of taxonomic levels including the validity of various families and their inter-relationships using nucleotide sequences from the mtDNA *Cyt b* (Briolay *et al.*, 1998; Cunha *et al.*, 2002; Durand *et al.*, 2002; Fuchs *et al.*, 2000; Gilles *et al.*, 1998; Zardoya, Doadrio, 1998; Zardoya *et al.*, 1999) and *CR* (Gilles *et al.*, 2001; Liu, Chen, 2003).While these studies bring important new insight into the evolutionary history of the family and its taxonomic classification, most studies have focused on European, Eurasian, North America and East Asian cypinids with the sampling of species from South East Asia including Vietnam having been neglected so far.

In the present study, sequences of the mtDNA *16S*, *CR* and *Cyt b* fragments were used to evaluate taxonomic and phylogenetic relationships within the Cyprinidae. Using sequences obtained from previous studies and from a set of species obtained from Vietnam, subfamily groupings are critically examined and the relationships suggested by Chen *et al*. (1984) and Cavender and Coburn (1992) and Gilles *et al*. (2001) are evaluated using maximum likelihood based hypothesis testing procedure (Shimodaira, Hasegawa, 1999).

6.2 Materials and Methods

6.2.1 Sample collection

Vietnamese cyprinid species were identified using the taxonomies of Cavender and Coburn (1992) and Nguyen and Ngo (2001). Tissue samples of cyprinids were obtained from fish kept in the National Brood Stock Center of RIA1, Hai Duong, Vietnam. The fish were originally obtained by Fish Gene Conservation Programs in 2004 and 2005, or collected from lakes, reservoirs and rivers in Vietnam using seine net and baited traps. Tissue samples were preserved in 90% ethanol and voucher specimens were preserved in 70% ethanol and deposited in the Fish Museum of RIA1. Sampled species, GenBank accession numbers and collection localities are given in Table 6.1.

6.2.2 DNA extraction and PCR amplification

Total DNA was extracted from fin-clip tissue, following the Crandall *et al*. (1999) method. One or two individuals were first analysed by direct sequencing from each species. The *Cyt b* gene was polymerase chain reaction amplified using the primers H15891 (5'GTT TGA TCC CGT TTC GTG TA 3') and L 15267 (5' AAT GAC TTG AAG AAC CAC CGT 3') (Briolay *et al.*, 1998). The *CR* was amplified by using the primers Carp-Pro (5' AAC TCT CAC CCC TGG CTA CCA AAG 3'), and Carp-Phe (5' CTA GGA CTC ATC TTA GCA TCT TCA GTG 3') (Thai *et al.*, 2004). The *16S* region was amplified using the primers 16Sar (5' GCC TGT TTA ACA AAA ACA T 3') and 16Sbr (5' CCG GTCTGA ACT CAG ATC ATG T 3') (Simon *et al.*, 1991). PCR was carried out in 50 µl reaction volumes (1 X reaction buffer, 2 mM dNTP, 1.5 mM $MgCl_2$, 0.5 µM of each primer, 0.5 units *Taq*

polymerase, and approximately 200 ng DNA template). Thermal cycling comprised 95 ^{o}C for 3 min, followed by 34 cycles of 95 ^{o}C for 30s, annealing at 55 ^{o}C (*CR* and *Cyt b*) and 58 ^{o}C (*16S*) for 30s, and an extension temperature of 72 ^{o}C for 1 min. This was then followed by a final extension of 72 ^{o}C for 3 min. PCR products were purified using the Qiagen (Hiden Germany) QIA quick PCR purification kit, following ABI PRISM BigDye Terminator (Foster city, CA, USA) protocols. For each individual, sequencing reactions were performed using both primers.

6.2.3 Data analysis

According to Gilles *et al*. (2001), and Liu *et al*. (2002), both morphology and molecular genetic data supports a monophyletic Cyprinidea. Following Liu and Chen (2003), sequences of *Crossostoma lacustre* from the Balitoridae (GenBank access number: M91245) were used as the outgroup. Two data sets were assembled for the analysis of cyprinid relationships. For the first data set, the 29 *CR* sequences generated in this study were combined with 27 *CR* sequences of the same length available for cyprinid species from GenBank. The second data set consisted of *CR*, *16S* and *Cyt b* sequence obtained in this study from 23 species and were combined with 9 additional cyprinid species for which sequence for these same mtDNA regions and length are available from GenBank.

Sequences were aligned using CLUSTAL X (Thompson *et al.*, 1997). To test for phylogenetic signal, a *g*1 statistic was calculated using 100,000 random trees as described by Hillis and Huesenbeck (1992). Four tree building methods of the were used to reconstruct phylogenetic relationships: maximum-likelihood (ML), neighbour-joining (NJ) and maximum parsimony (MP) were implemented using

PAUP* 4.0b10 (Swofford, 2000); and Bayesian methods were carried out using MrBayes 3.0 (Huelsenbeck, Ronquist, 2001). The appropriate model of evolution for ML, NJ and Bayesian analyses was obtained via testing alternative models of evolution using Modeltest (Posada, Crandall, 1998). Heuristic searches used for ML analyses consisted of 100 replicates of random sequence additions, while non-parametric bootstrapping consisted of 100 replications with 10 random sequence additions. MP analyses were performed with gaps treated as missing data, heuristic searches as per maximum-likelihood analyses, but with 1,000 non-parametric bootstrap replicates. The NJ tree was constructed with distances calculated under the same model of evolution as the ML analysis, with bootstrapping performed using 1,000 replicates. Bayesian analyses were performed using the same general model identified by Modeltest. Analyses were initiated with random starting trees and run for 1.0×10^6 generations, sampling the four Markov chains every 100 generations resulting in 10,000 trees. The likelihood scores of the sampled trees were plotted against generation time to ensure that stationarity was reached, trees generated prior to stationarity being reached were discarded as “burn-in” (1,500 trees in this case). Bayesian posterior probabilities of each bipartition, representing the percentage of times each node was recovered were calculated from a 50% majority rule consensus of the remaining trees.

6.2.4 Phylogenetic hypothesis testing

To test taxonomic and phylogenetic hypotheses proposed by other authors, comparisons were made between trees derived from these hypotheses and the optimal trees recovered by our analysis, using the SH test (Shimodaira, Hasegawa, 1999). Due to incomplete or limited taxonomic sampling Gobiobotinae, Acheilognathinae, Phoxininae, Alburninae and Schizothoracinae were excluded from

hypothesis testing. Three hypotheses for taxonomic and phylogenetic relationships of Vietnamese cyprinid species, as proposed by Chen *et al*'s. (1984), Cavender and Coburn's (1992) and Gilles *et al*'s. (2001) studies, were tested (Table 6.2).

6.3 Results

6.3.1 Sequence variation

All sequences obtained in this study have been submitted to GenBank (accession numbers DQ464904-464992; DQ864654-864655). A summary of the characteristics of each mitochondrial region is presented in Table 6.3. It can be seen from Table 6.3 that all three fragments show significant phylogenetic signal (based on $g1$ values). The *CR* sequences show the most variation and the *16S* region the least.

The combined *16S*/*Cytb*/*CR* data set consisted of 1,786 aligned nucleotide positions. Of these, 861 were variable and 636 parsimony informative. The partition homogeneity test did not reject phylogenetic congruence between these mtDNA fragments ($P < 0.05$), allowing their combination for phylogenetic analyses. Tree length frequency distributions were significantly skewed for all taxa ($g1 = -1.250$; $P < 0.05$), suggesting the presence of phylogenetic signal. The model selection for the NJ and ML analysis was TRN + I + G which accommodates differing transition/ tranversion mutation rates. Percentage sequence divergence among taxa ranged from 0.0% (*Tor tambroides* and *Tor stracheyi*) to 31.6% (*Danio rerio* and *Hampala macrolepidota*).

6.3.2 Phylogenetic analysis

Following the classification of Cavender and Coburn (1992), the relationship among 51 cyprinid species representing 41 genera and 12 subfamilies were evaluated using the *CR* sequences. In general, shallower relationships were resolvable to a much greater extent than at the deeper levels, which is consistent with the known rapid rate of evolution of this mtDNA region (Fig. 6.1; 6.2; 6.3; 6.4). The NJ, ML and Bayesian methods of analysis generated almost identical relationships but with varying levels of nodal support. Many of the relationships were unresolved using MP but those that were, mostly similar to the other analyses. The tree indicated many inconsistencies with the current classification. At the family level, species of Cultrinae and Xenocyprinae did not form the two anticipated monophyletic groups, although together these species form a well supported monophyletic lineage. Sister to this lineage is a well supported group containing representatives of the Leuciscinae, Gobininae, Achelognathinae, Tinciae and Gobiobtinae, entirely consistent with the current classification, although only the Leuciscinae is represented by more than two species.

The remaining species form a sister group to the two previously discussed lineages that failed to clarify deeper level relationships. Further, the analyses do not support a monophyletic Barbininae with species of this subfamily distributed across four divergent lineages with varying levels of support. The Labeoninea is also non-monophyletic on the basis of the placement of *Hampala macrolepidota* (Barbininae) as sister to *Lobocheilos melamotaenia*. Otherwise there is some support for the Labeoninae as a natural group. In contrast the Cyprininae receives significant support as a monophyletic group based on five species representing three genera.

There are number of genera in the data set that are represented by two or more species and it is apparent that the morphologically based classifications of cyprinids fails in many cases at this level as well. While the genera *Tor* (three species), *Barbus* (two species), *Cyclocheilichthys* (two species) and *Gobio* (two species) are monophyletic, the genera *Culter* (three species), *Leuciscus* (three species), *Cirrhinus* (two species), *Cyprinus* (two species) and *Carassiodes* (two species) are all non-monophyletic.

An examination of divergence levels between species within genera, and species belonging to different genera, also indicates inconsistencies in the morphologically based cyprinid taxonomy. Divergence levels between monophyletic congeneric species supported by the phylogenetic analyses, range from 0.90 to 16.4% which overlaps broadly with divergence levels for monophyletic species pair placed in different genera which ranges from 0.52% to 50.03 %. There was also no support for *Cyprinus melanes* as a distinct taxon as this sample clustered within the *C. capio* sample with very low divergence levels (1.02-1.91%). Other species pairs that show very high levels of similarity are *Tor Tambroides* and *Tor stracheyi* (0.00%) and *Cyclocheilichthys repasson* and *Cyclocheilichthys apogon* (1.25%), and are therefore of questionable status as valid species.

The data set consisting of the concatenated *16S*, *CR*, and *Cyt b* sequences successfully clarified deeper level relationships, despite more limited taxon sampling. The different methods of analysis recovered similar results with the exception that the maximum likelihood analysis placed *Carassioides cantonensis*, rather than

Carassiodes phongnhaensis as a sister to *Carassius auratus* and the parsimony analysis placed *Danio rerio* in an alternative position (Fig. 6.5; 6.6; 6.7; 6.8). The cyprinids were divided into the same two major clades by each analysis, with exception that *Danio rerio*, was placed in a more basal clade either as sister to a clade containing (Tincinae, Gobionia, Leuciscinae, Xenocyprinae and Cultrinae) (Fig. 6.5; 6.6; 6.8) or as sister to all other taxa (Fig. 6.6). Similarly to the *CR* analysis, all the European subfamilies (excluding Danioninae) form a monophyletic group which is sister to a clade containing representatives of the Xenocyprinea and Cultrinae. Also consistent with the *CR* only analysis, Xenocyprinae is non-monophyletic and the monophyly of the Cultrinae is only weakly supported.

The other major lineage which is well supported by the concatenated data contains the Cyprininae, the Labeoninae, and the Barbininae. While the Cyprininae is supported as monophyletic the Barbininae and Labeoninae are non-monophyletic which is also consistent with the *CR* analysis. All members of this lineage are from Asia with the exception of the sample of *C. carpio* from Hungary. The *16S*, *Cyt b*, and *CR* analysis also fails to support several generic level grouping including *Leuciscus* (Leuciscinae), *Cirrhinus* (Labaoninae) and *Carassiodes* (Cyprininae). Lastly, the analysis also fails to provide support for the recognition of *Cyprinus melanes* as a distinct species from common carp.

The testing of specific taxonomic hypotheses using the Shimodaira and Hasegawa (1999) procedure rejects several of them. The hypothesis of Chen *et al*. (1984) was rejected as significantly inferior to the optimal tree, as was also the monophyly of the Xenocyprinae and the Cutrinae. Based on the data set utilized in this study, none of

the alternative taxonomic hypotheses could be statistically rejected, even though subfamilies such as Barbininae are non-monophyletic based on the reconstructed trees.

6.4 Discussion

Comparison of the results of this study with the literature on Cyprinidae systematics is complicated because of the diversity of the family and the many and varied data sets with respects to kinds of data (molecular and morphological) and taxa sampled and methods of analysis that have been used (Cavender, Coburn, 1992; Chen *et al.*, 1984; Gilles *et al.*, 2001; Gosline, 1978; Howes, 1991; Nelson, 1994). Nevertheless some key points of agreement emerge between this and other studies.

At the deepest taxonomic level the results from the combined *16S*/ *Cyt b*/ *CR* data supports a fundamental division between the Cyprinine (Cyprininae + Barbinae + Labeoninae + Schizothoracinae) and the Leuciscine (Leuciscinae + Acheilognathinae + Gobioninae + Gobiobotinae + Tincinae), with the Rasborinae or Danioninae joining the Leuciscine lineage at the most basal position in all analyses other than parsimony. This association of the Danioninae with the Leuciscine lineage, rather than as the sister group to (Leuciscine + Cyprinine) is consistent with the morphological based analysis of Cavender and Cobum (1992) and the molecular study of Liu and Chen (2003). However, with respect to the placement Danioninae, our results are contrary to both the morphologically based study of Chen *et al*. (1984) and Gilles *et al*'s (2001) molecular study. The weight of evidence would seem to favour the Danioninae as the sister lineage to the Leuciscinae as it is supported by two molecular studies using different tree building methods and does not require the

independent evolution of a complex morphological trait associated with the pleural rib in two separate lineages (Gilles *et al.*, 2001). However, it should be noted that the parsimony-based analysis in this study supported Gilles *et al*' s. (2001) study, which used only parsimony-based analyses suggesting their conclusion may be a result of inherent limitations to parsimony methods known as "long-branch" attraction. Further, while Chen's hypothesis could be rejected at very low level of significance, Gilles *et al*'s (2001) hypothesis could not. Thus, this hypothesis requires further testing through greater taxon and gene sampling before it can be categorically refuted.

The close relationship of the Tincins to the Leuciscin taxa is supported by three previous molecular studies (Gilles *et al.*, 2001; Liu, Chen, 2003; Zardoya, Doadrio, 1998) and by the Cavender and Coburn' s (1992) morphological study. This contradicts the morphological data of Chen *et al*. (1984). However, the precise relationships of the Tincins is uncertain as it is variously placed by different analyses as basal to the other Leuciscine taxa, as part of a polyphyletic node, or associated with a clade containing representative of the Gobioninae, Leuciscinae and Acheilognathinae.

Some strongly supported relationships such as between Cultrin and Xenocyprinin species is entirely consistent with both morphological (Cavender, Coburn, 1992) and molecular studies (Liu, Chen, 2003), although support for the monophyly of each subfamily is inconsistent based on this study and that of Liu and Chen (Liu, Chen, 2003). Another strongly supported relationships which is consistent with Liu and

Chen but is contrary to the morphologically based analyses of Cavender and Coburn (1992), is that between species of Leuciscinae and Gobioninae.

The other major lineage contains the Barbins and Labeonins, which together make up the Cyprinidae, a group that has been recovered by the major morphological analyses and all major molecular analyses (Durand *et al.*, 2002; Gilles *et al.*, 2001; Liu, Chen, 2003; Zardoya, Doadrio, 1999). While this study suggested the Cyprinine may be monophyletic, it is only based on limited taxon sampling. In contrast, the data suggest that Labeoninae and Barbinae are non-phylogenetic. Other studies sampling different species have concluded that the Barbinae is polyphyletic including the genus *Barbus* its self. In addition, Durand *et al*. (2002) commented that morphological characters are “sometimes irrelevant” in phylogenetic inference in the cyprinids. This comment is also seen to be at least partially true for several other genera that were found to be non-monophyletic in this study and the study of Gilles *et al*. (2001) and Zardoya and Doadria (1999) including *Culter*, *Leuciscus*, *Labeo*, *Cirrhinus*, *Carassioides*, to which can be added *Rutilus* and *Chodrostonma*, based on Zardoya and Doadrio (1999).

The data from this study indicates that deficiencies in morphological information extend to the lowest taxonomic level. Thus re-examination of species boundaries in several genera, including *Tor*, *Cyclocheilichthys* and *Cyprinus* are required. Based on morphological data, six species were identified in *Cyprinus*: *C. carpio* Linnaeus, 1758; *C. melanes* Yen, 1978; *C. quidatensis* Tu et al, 1999; *C. multitaeniata* Pellegrin and Chevey, 1936; *C. hyperdosalis* Hao, 1991, and *C. exopghthalmus* Yen, 1978 (Nguyen, Ngo, 2001). In this study, *C. carpio* and *C. melanes* is not

differentiated at the molecular level. In fact, from the phylogenetic analysis, it is very clear that the levels of divergence within the genus *Cyprinus* are very modest compared with other genera and this analysis does not support the division of the genus into more than one species. An interesting observation is that *Carassiodes phong- nhaensis* may be more closely related to *Cyprinus* than it is to *Carassioides cantonensis*. This suggests that *Carassioides phongnhaensis* may be a more appropriate outgroup for phylogenetic studies of *Cyprinus* and that the species could be placed in the genus *Cyprinus* rather than *Carassiodes*.

In summary, the result of this study confirms and contradicts elements of both morphological and molecular studies of the Cyprinidae at various taxonomic levels. Nevertheless, it is clear that there are two principal lineages within the Cyprinidae: Cyprinie and Leuciscine and that further molecular studies are required to define well supported monophyletic groups within each of these lineages that can be associated with existing named groups and morphological information. Such studies will need substantial taxon sampling to ensure the generality of the results and that stable taxonomic classification can be constructed.

Table 6.1. Species, sampling localities, GenBank access numbers subfamily designation as proposed by previous studies.(* samples from this study)

Species	Code	Locality	*Cyt b*	*16S*	*CR*	Chen *et al.* (1984)	Cavender & Coburn (1992)	Gilles *et al.* (2001)	Nguyen & Ngo (2001)
Ancherythroculter daovantieni	ANC*	Vietnam	DQ464975	DQ464929	DQ464940	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Aristicchthys nobilis	BHC*	Vietnam	DQ464976	DQ464908	DQ464949	Xenocyprinae	Xenocyprinae	Xenocyprinae	Hypophthalmichthyinae
Barbonymus gonionotus	BOB*	Vietnam			DQ464945	Barbinae	Barbinae	Barbinae	Barbinae
Babus fluviatilis	BAF	Europe			AJ388415	Barbinae	Barbinae	Cyprininae	Cyprininae
Barbus meridionali	BAM	Europe			AJ388417	Barbinae	Barbinae	Cyprininae	Cyprininae
Carassiodes phongnhaensis	CHA	Vietnam			DQ464946	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Carassioides cantonensis	NHU	Vietnam	DQ464980	DQ464930	DQ464962	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Carassius auratus	CRU*	Vietnam	DQ464978	DQ464926	DQ464961	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Chondrostonma nanus	CHO	Europe	AY026402	AJ247047	AJ388396	Leuciscinae	Leuciscinae	Leuciscinae	Leuciscinae
Cirrhinus cirrhosus	MRI*	Vietnam	DQ464981	DQ464904	DQ464952	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Cirrhinus molitorella	MUD*	Vietnam	DQ464968	DQ464921	DQ464964	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Crossotoma lacustre (outgroup)	CRO	Taiwan?	M91245	M91245	M91245	Balitorinae	Balitorinae	Balitorinae	Balitorinae
Ctenopharyngodon idellus	GRC*	Vietnam	DQ464983	DQ464928	DQ464953	Leuciscinae	Xenocyprinae	Leuciscinae	Leuciscinae
Culter alburns	CUT	China			AY095331	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Culter mongolieus	CUM	China			AY095329	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Culter oxycephaloides	CUO	China			AY095328	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Cultrichthy erythropterus	THI*	Vietnam	DQ464977	DQ464934	DQ464954	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Cyclocheilichthys repasson	CLO*	Vietnam			DQ464938	Barbinae	Barbinae	Barbinae	Barbinae
Cyclocheilichthys apogon	CYL*	Vietnam	DQ464989	DQ464918	DQ464955	Barbinae	Barbinae	Barbinae	Barbinae
Cyprinus carpio	XIN*	China	AY347282	DQ864655	DQ532110	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Cyprinus carpio	BBC1	China			AY347303	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Cyprinus carpio	LBW	Japan	AB158803	AP009047	AB158808	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Cyprinus carpio	HUS*	Hungary	DQ532114	DQ864654	AY597981	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Cyprinus carpio	CYC*	Vietnam	DQ464969	DQ464909	DQ464944	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Cyprinus melanes	CYM*	Vietnam	DQ464970	DQ464910	DQ464943	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Dangila lineatus	DAN*	Vietnam	DQ464991	DQ464907	DQ464939	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Danio rerio	DAR	Europe	NC002333	NC002333	NC002333	Danioninae	Rasborinae	Rasborinae	Rasborinae
Discogobio tetrabarbatus	DIS	China			AY095326	Labeoninae	Labeoninae	Labeoninae	Labeoninae

Table 6.1. continued..

Species	Code	Locality	*Cyt b*	*16S*	*CR*	Chen *et al.* (1984)	Cavender & Coburn (1992)	Gilles *et al.* (2001)	Nguyen & Ngo (2001)
Distoechodon tumirostros	DTT	China			AY014165	Xenocyprinae	Xenocyprinae	-	Xenocyprinae
Gobio gobio 1	GOB	Europe			AJ388393	Gobioninae	Gobioninae	Gobioninae	Gobioninae
Gobio gobio 2	GOG	Europe	AJ388431	AJ247056	AJ388392	Gobioninae	Gobioninae	Gobioninae	Gobioninae
Gobiobotia filifer	GOF	China			AY095341	Gobiobotinae	Gobiobotinae	Gobiobotinae	Gobiobotinae
Hampala macrolepidota	HAN*	Vietnam	DQ464974	DQ464916	DQ464947	Barbinae	Barbinae	Barbinae	Barbinae
Hemicuter leucisculus	MXA*	Vietnam	DQ464973	DQ464923	DQ464957	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Hypophthalmichthys molitrix	SIL*	Vietnam	DQ464966	DQ464936	DQ464958	Xenocyprinae	Xenocyprinae	Xenocyprinae	Hypophthalmichthyinae
Label bicolor	LBI	Europe			AJ388414	Cyprininae	Cyprininae	Cyprininae	Cyprininae
Labeo rohita	ROH*	Vietnam	DQ464965	DQ464935	DQ464950	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Leuciscus cabeda	LEU	Europe			AJ388406	Leuciscinae	Leuciscinae	Leuciscinae	Leuciscinae
Leuciscus cephalus	LEC	Europe	AJ252805	AJ247054	AJ388407	Leuciscinae	Leuciscinae	Leuciscinae	Leuciscinae
Leuciscus sofia	LES	Europe			AJ388398	Leuciscinae	Leuciscinae	Leuciscinae	Leuciscinae
Lobocheilos melanotaenia	LOB*	Vietnam	DQ464990	DQ464917	DQ464948	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Mylopharyngodon piceus	BLC*	Vietnam	DQ464971	DQ464905	DQ464937	Leuciscinae	Xenocyprinae	Leuciscinae	Leuciscinae
Paracheilognathus imberbis	PAR	China			AY017147	Acheilognathinae	Acheilognathinae	Acheilognathinae	Acheilognathinae
Puntius brevis	PUB	Vietnam	DQ464967	DQ464912	DQ464942	Barbinae	Barbinae	Barbinae	Barbinae
Rasbora trilineata	RAT	Europe			AJ388423	Danioninae	Rasborinae	Rasborinae	Rasborinae
Rhodeus amarus	RHA	Europe			AJ388412	Acheilognathinae	Acheilognathinae	Acheilognathinae	Acheilognathinae
Rutilus rubilio	RUT	Europe			AJ388400	Leuciscinae	Leuciscinae	Leuciscinae	Leuciscinae
Schizothorax chongi	SCH	China			AY095325	Schizothoracinae	Schizothoracinae	Schizothoracinae	Schizothoracinae
Semilabeo obscurus	AVU*	Vietnam	DQ464988	DQ464913	DQ464963	Labeoninae	Labeoninae	Labeoninae	Labeoninae
Sinibrama macrops	SIM	China			AY095332	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Spinibarbus denticulatus	BON*	Vietnam	DQ464984	DQ464906	DQ464956	Barbinae	Barbinae	Barbinae	Barbinae
Tinca tinca	TIN	Europe	Y10451	AJ247053	AJ388411	Tincinae	Tincinae	Tincinae	Tincinae
Tor duronensis	TOD*	Vietnam	DQ464986	DQ464925	DQ464959	Barbinae	Barbinae	Barbinae	Barbinae
Tor stracheyi	TOS*	Vietnam	DQ464987	DQ464915	DQ464951	Barbinae	Barbinae	Barbinae	Barbinae
Tor tambroides	TOT*	Vietnam	DQ464985	DQ464914	DQ464960	Barbinae	Barbinae	Barbinae	Barbinae
Toxabramis houdemeri	TOX*	Vietnam	DQ464972	DQ464924	DQ464941	Cultrinae	Cultrinae	Cultrinae	Cultrinae
Xenocypris hupeiensis	XEH	China			AY014164	Xenocyprinae	Xenocyprinae	Xenocyprinae	Xenocyprinae

Table 6.2. Major hypotheses for phylogenetic relationships of cyprinid species.

Hypothesese	Topology
Chen *et al* ., (1984)	(((Barbininae, Cyprininae), Labeoninae), Tincinae), (Danioninae, (Gobioninae, ((Xenocyprinae, Cultrinae), Leuciscinae)));
Cavender and Coburn's (1992)	(((((((Barbininae, Cyprininae), Labeoninae), Tincinae), (Xenocyprinae, Cultrinae)), Leuciscinae), Gobioninae), Danioninae);
Gilles *et al* ., (2001)	(((((((Barbininae, (Cyprininae, Labeoninae), Tincinae), (Xenocyprinae, Cultrinae)), Gobioninae), Leuciscinae), Danioninae);

Table 6.3. Summary of results of phylogenetic analysis in three mitochondrial DNA gene regions of cyprinid species.

	16S	*Cyt b*	*CR*
# sites	446	582	758
% variable sites	13.229	7.045	17.282
% Parasimony infomative	10.538	3.952	46.438
Ti/Tv	3.13	2.11	1.365
Sequence divergence (%)	0.5-17.23	0.5-26	0.6-38
g 1	-0.428	-0.481	-0.548
Model of evolution	HKY + I + G	GTR + I	HKY + I + G

Table 6.4. Tests of alternate phylogenetic hypotheses using combine *16S*, *CR*, *Cyt b* regions, Shimodaira-Hasegawa (SH) test. (* significant difference between optimal and alternate topologies, *P* <0.05.

Tree	lnL	Diff-lnL	*P*
Optimal	16786.822	(best)	
Chen *et al* . (1984)	16887.860	97.316	0.000*
Cavender and Coburn (1992)	16793.900	7.077	0.400
Gilles *et al* . (2001)	16792.410	7.062	0.562
Cyprininae	16798.804	21.012	0.530
Barbininae	16737.129	47.337	0.120
Labeoninae	16792.151	2.360	0.710
Xecyprinae	16891.804	72.012	0.001*
Cultrinae	16873.140	80.976	0.035*
Leuciscinae	16705.702	15.910	0.560

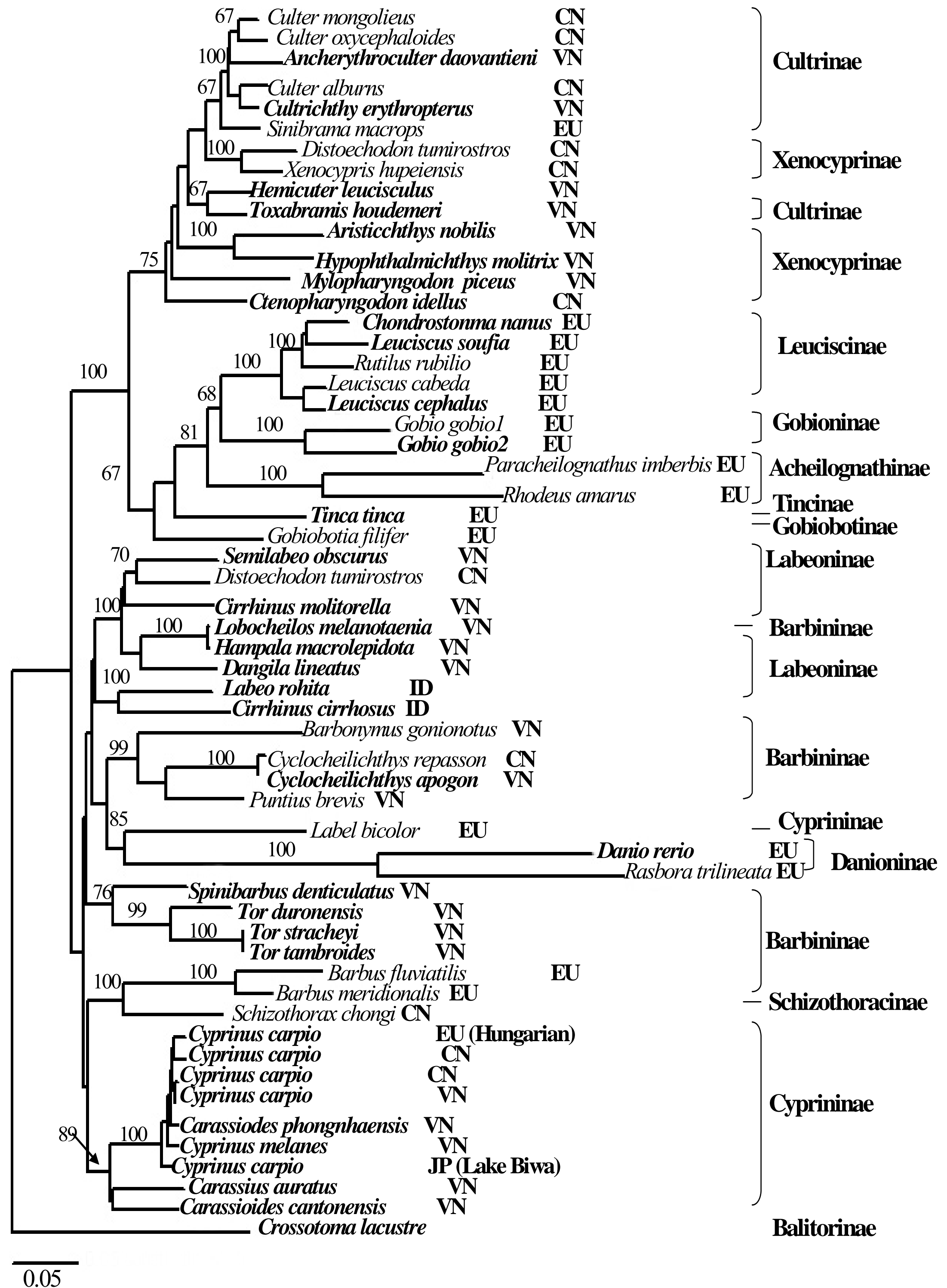

Figure. 6.1. Phylogenetic tree resulting from neighbour-joining analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support . The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combined three region data sets.

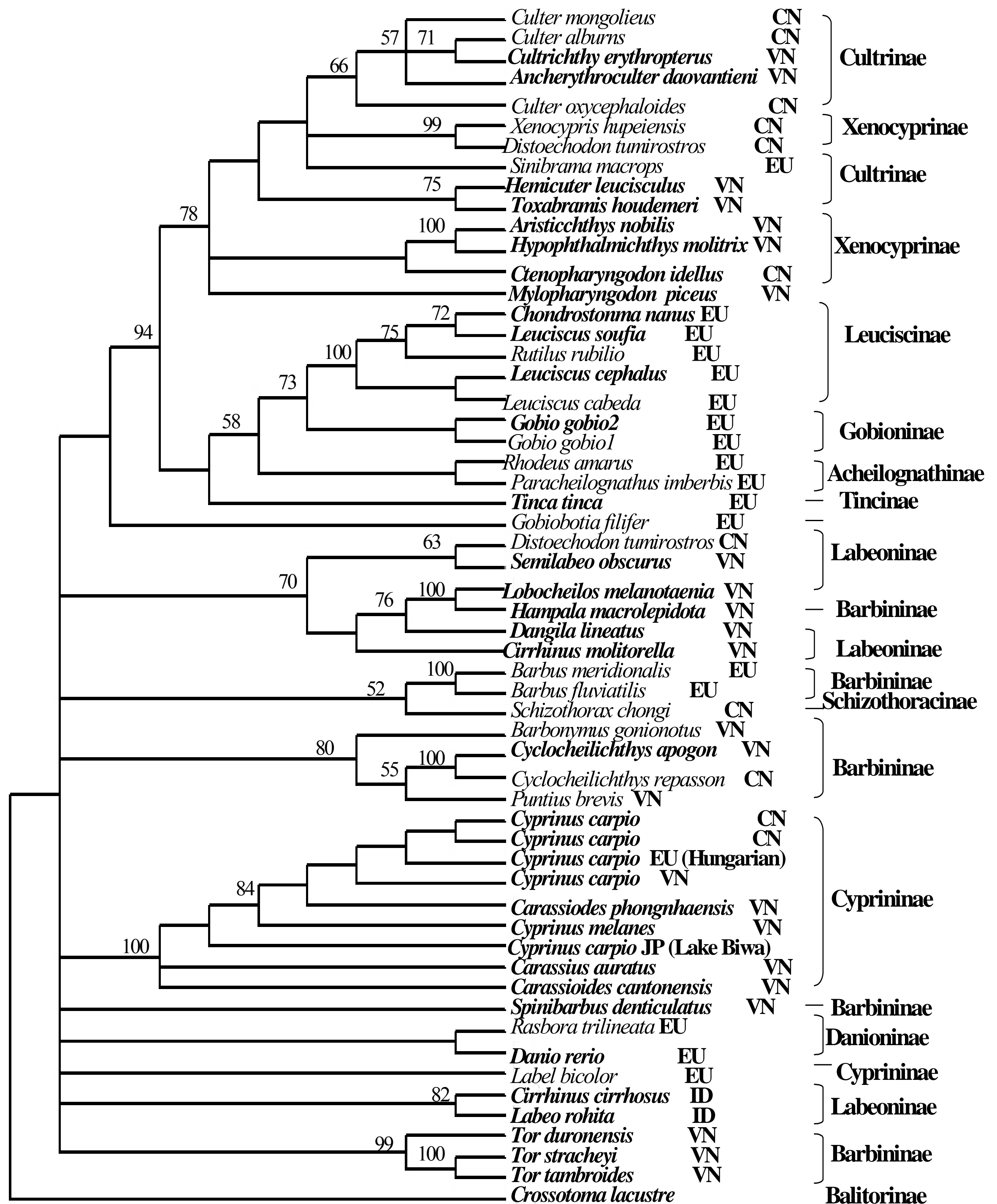

Figure. 6.2. Phylogenetic tree resulting from maximum parsimony analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support. The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combined three region data sets.

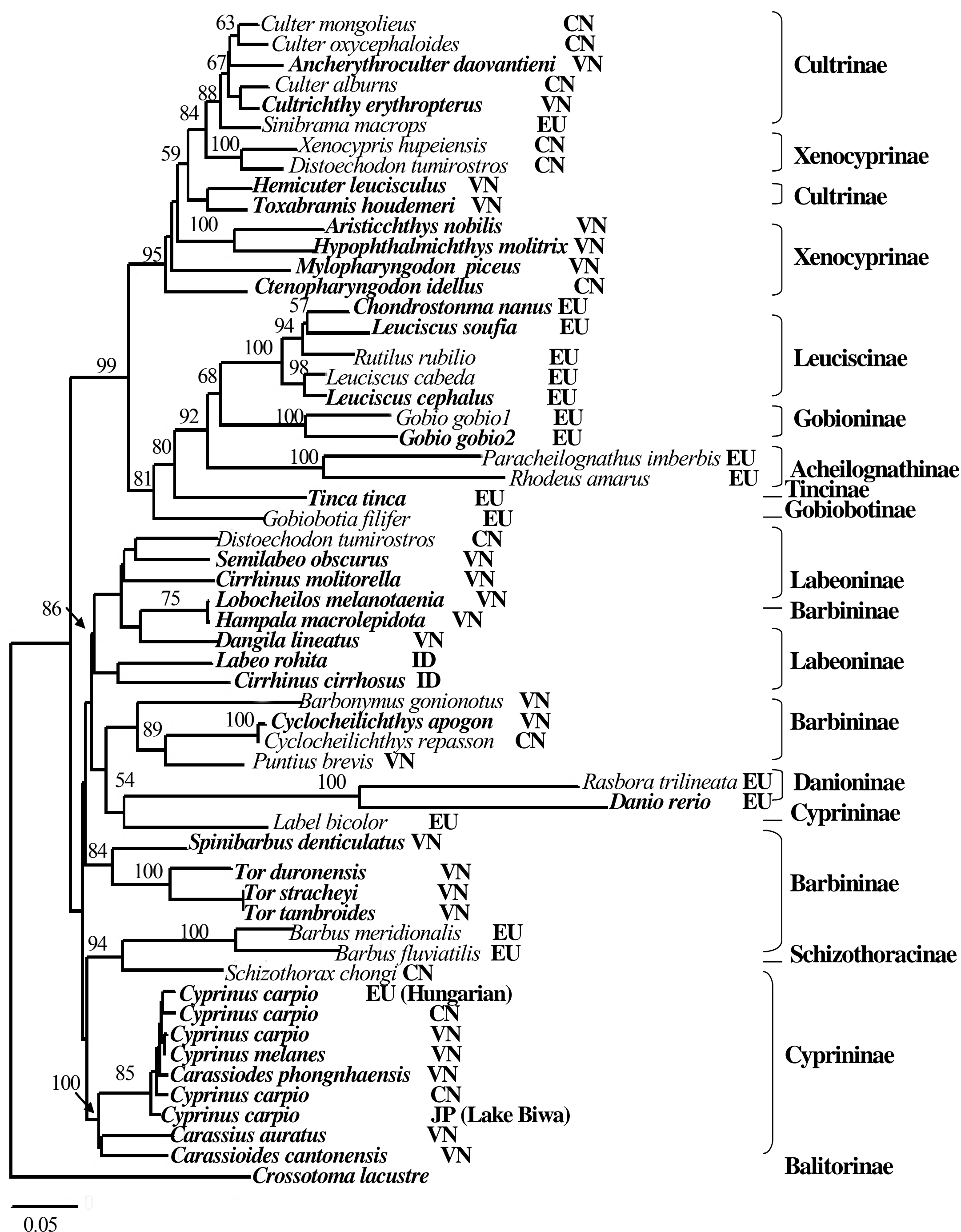

Figure. 6.3. Phylogenetic tree resulting from maximum likelihood analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent bootstrap support. The subfamily groups are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates sample used in combine three region data sets

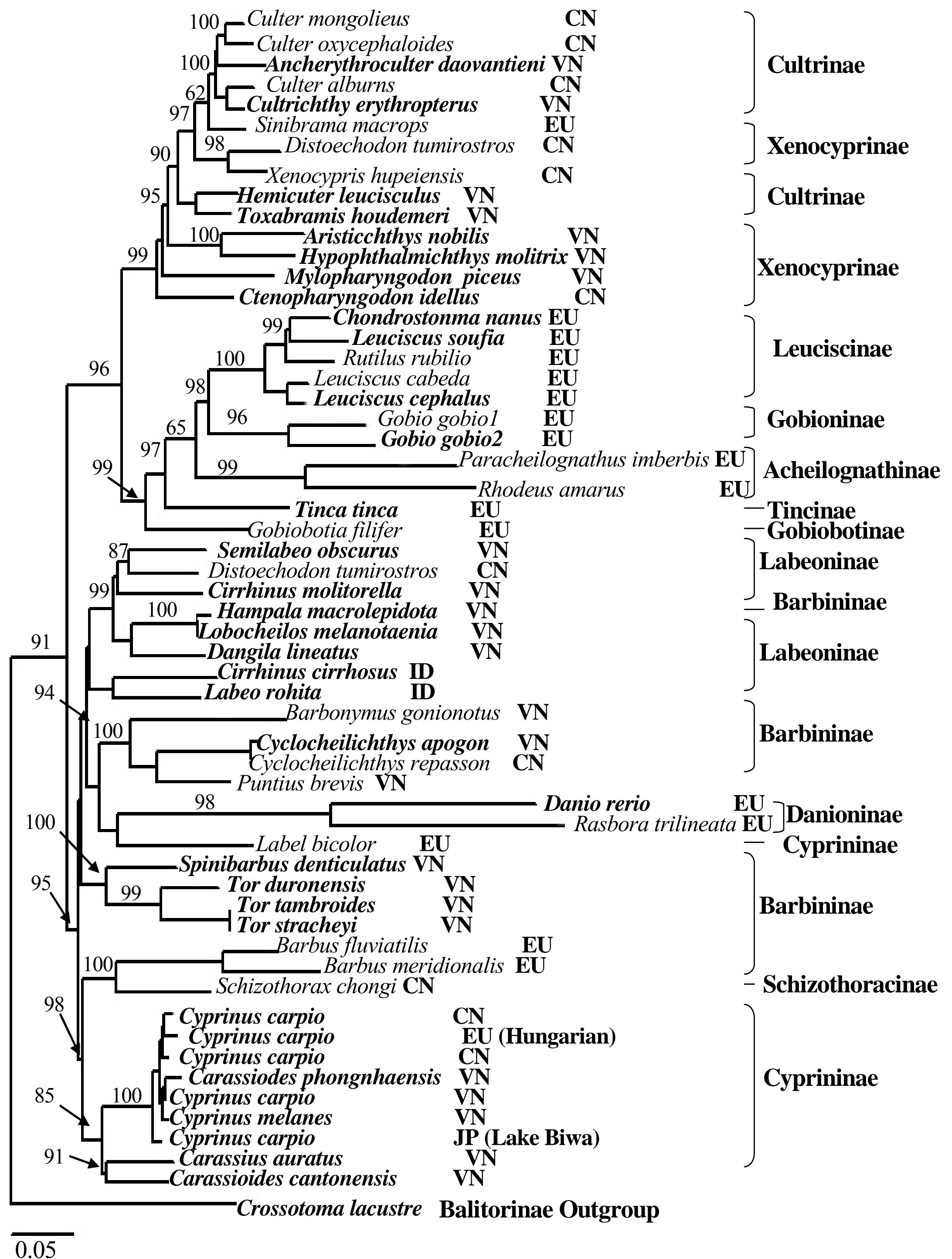

Figure. 6.4. Phylogenetic tree resulting from Bayesian analysis of mitochondrial DNA *CR* of 51 cyprinid species. Numbers on each branch represent Bayesian posterior probabilities. The subfamily group are based on Cavender and Coburn (1992).VN: Vietnam, CN: China, ID: India, JP: Japan, EU: Europe. Bold indicates samples used in combine three region data sets.

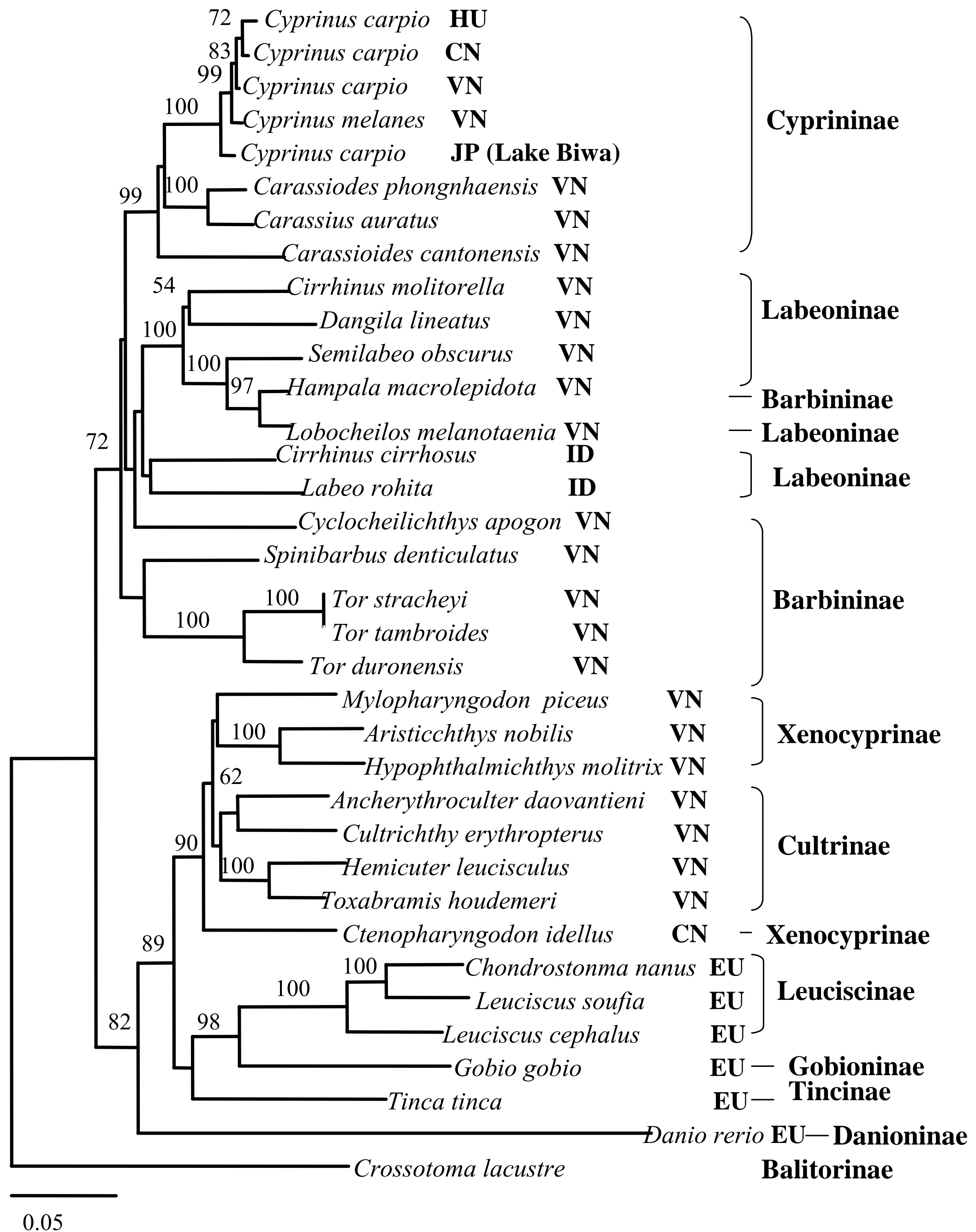

Figure. 6.5. Neighbour-joining tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe.

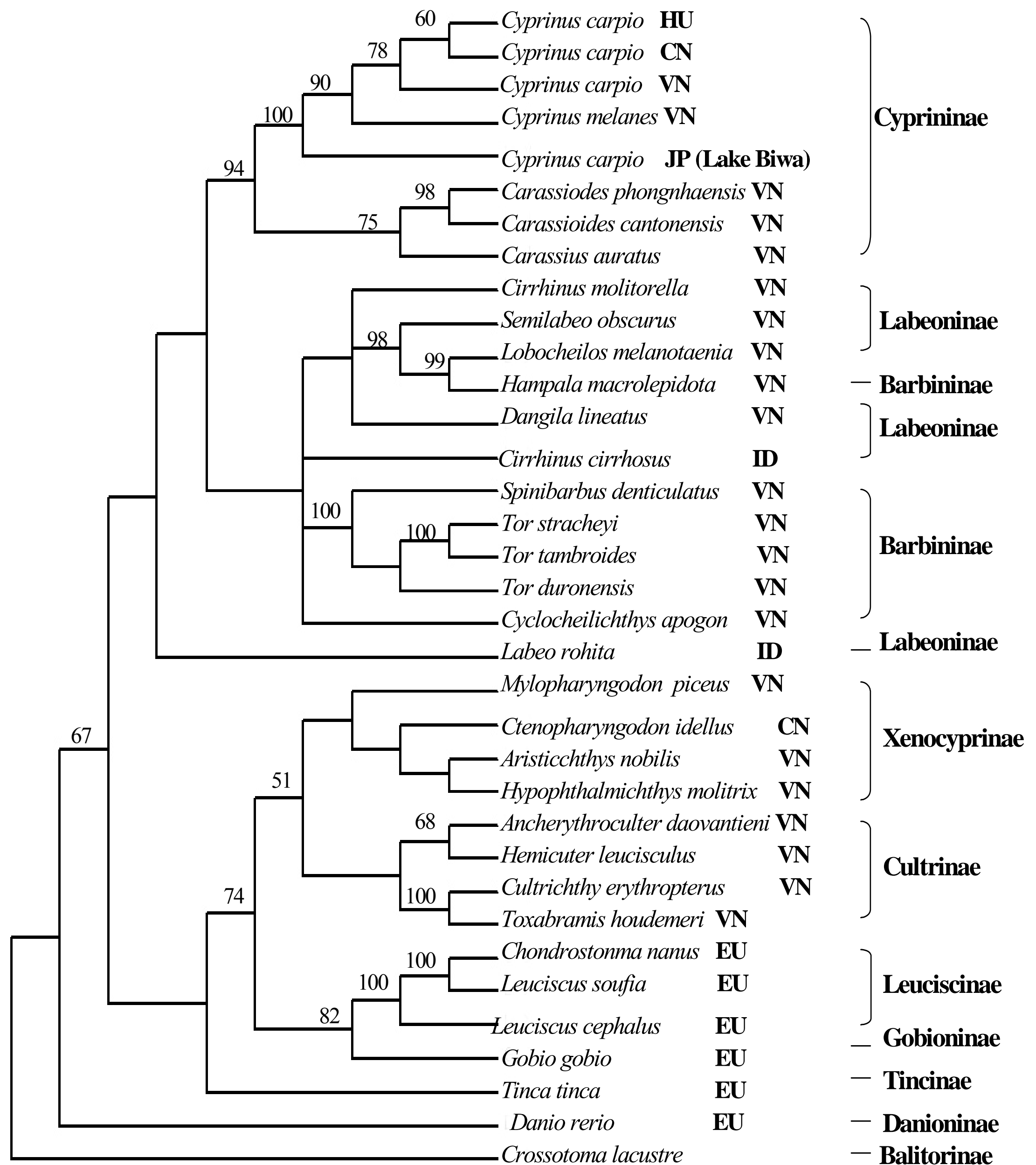

Figure. 6.6. Maximum parsimony tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe.

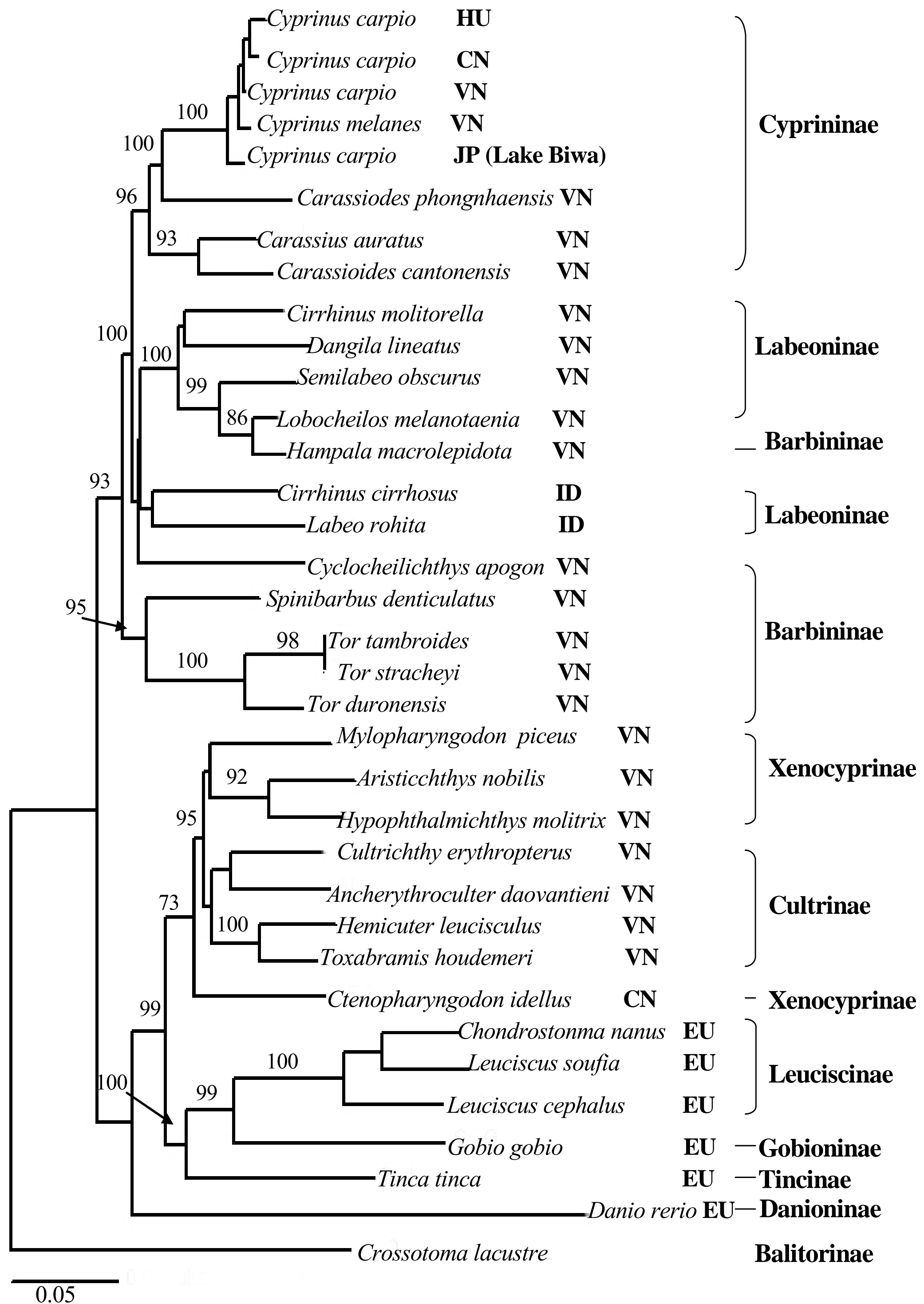

Figure. 6.7. Maximum likelihood tree from combined *16S*, *Cyt b* and *CR* mtDNA data. Number on branches indicate bootstrap value (- = <50 bootstrap support). The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe.

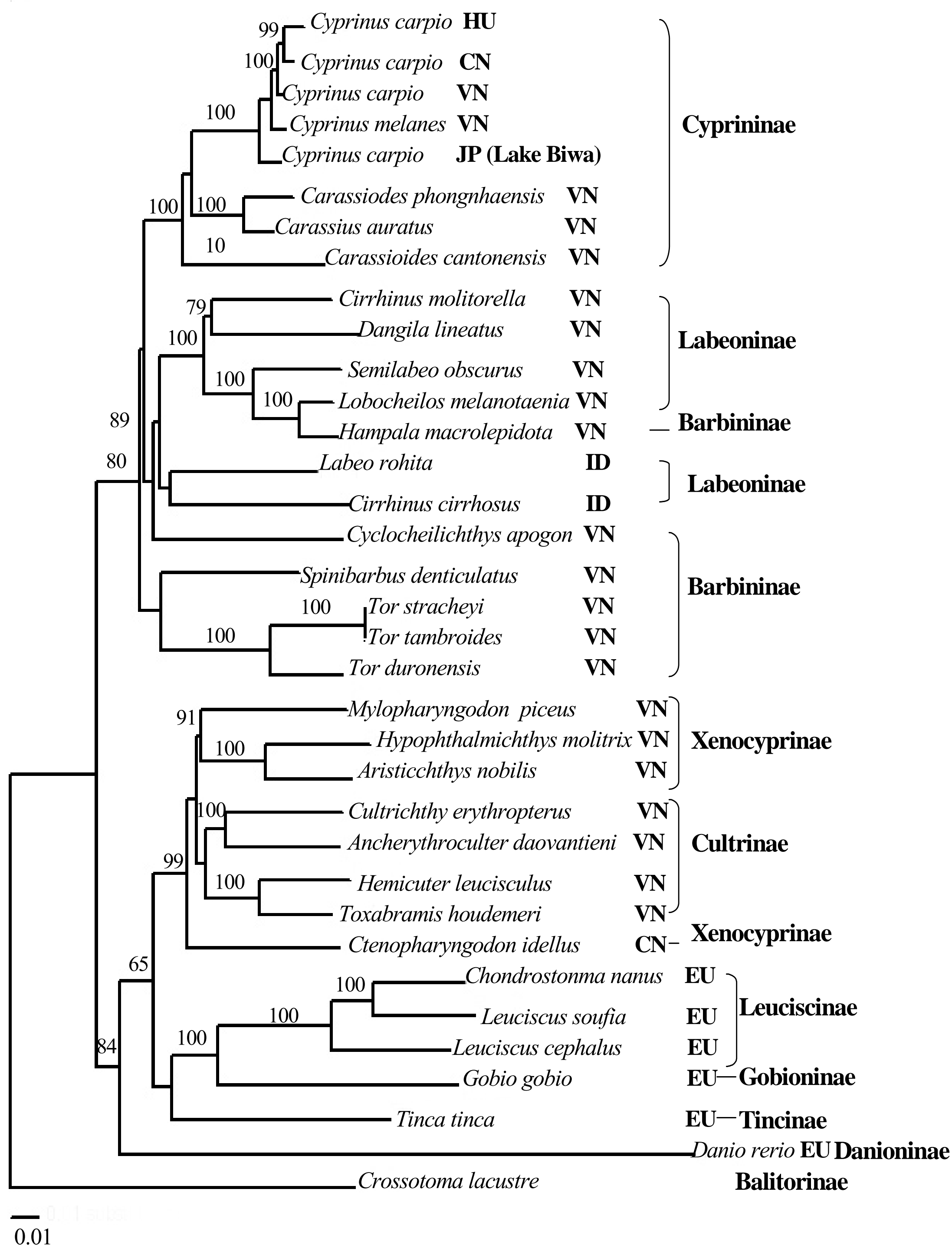

Figure. 6.8. Bayesian tree from combined *16S*, *Cyt b* and *CR* mtDNA data. The subfamily groups are based on Cavender and Coburn (1992). VN: Vietnam, CN: China, ID: India, HU: Hungary, JP: Japan, EU: Europe.

Chapter 7.

General Discussion

The wide distribution, long history of domestication, taxonomic uncertainties and evolutionary disputation makes common carp an interesting, challenging and at times controversial species for study (Zhou *et al.*, 2003). The advancement of knowledge of the population genetics and the provision of information relevant to the genetic improvement and effective mangement of domesticated and wild stocks of common carp is also of the utmost importance given the species significance for food security and household incomes in many developing countries.

The series of related studies presented in this thesis, involving the acquistion of different kinds of molecular genetic data, contributes significant new knowledge about common carp. Indeed this thesis arguably represents, the single most comprehensive study of common carp molecular genetics so far undertaken, contributing to the understanding of genetic impacts of domestication, taxonomy and evolution of the species. This study is certainly the most indepth molecular genetic investigation of any Vietnamese fish species so far completed.

Following from the primary aims of this study, the following section summarises and discusses the principal findings and conclusions, together with recommendations for future research.

7.1 Global genealogy, taxonomy and evolution of common carp

Over 50 common carp strains or populations, obtained on a worldwide basis and from comprehensive collections from Vietnam, have been used to reconstruct genealogical relationships using mtDNA sequences. Depending on the question being addressed data from three different mtDNA fragments have been utilised: *ATPase6/ATPase8*, *CR*, *cytochrome b*. Sequence data with up to 2,130 bp was analysed using various combinations of samples as detailed in Chapters 2, 3 and 4. A remarkably consistent picture emerged from the different nucleotide sequences and genealogical analyses. Firstly, compared to other freshwater fish species, genetic divergence among common carp stocks and strains is very low, especially given its wide geographic distribution. Secondly, the European and Asian strains and stocks do not segregated into discrete monophyletic units as anticipated based on current taxonomic opinion (Balon, 1995; Kohlmann *et al.*, 2003). In fact, the European strains show remarkably little within and between strain variation, consistent with translocated stocks.

These results have important implications for the taxonomy of *Cyprinus carpio*, which is highly confused (Balon, 2004) with some authors considering *Cyprinus* to be a monotypic genus (Lever, 1996), while others consider it to be represented by multiple species (Baruš *et al.*, 2002; Kottelat, 1997; Nguyen, Ngo, 2001; Zhou, Chu, 1986). Fundamentally, the very low levels of genetic divergence within the species, and the failure to resolve any lineages consistent with recent taxonomic opinions, suggest that *Cyprinus* is best considered a monotypic genus, with *C. carpio* being a

morphologically plastic species, with much of its extensive distribution likely the result of translocation (see also Mabucchi *et al*. 2006).

These studies are therefore also consistent with an Asian origin for common carp with at least one translocation event to Western Europe, because the European common carp group is nested in the Asian clades and wild Amur carp have a haplotype identical to European carp. Such a finding makes no biogeographical sense unless European carp were founded by translocation. Additional evidence supporting an Asian origin for common carp is the much higher haplotype diversity in Asia, the occurrence of its closest relatives *Carassius auratus* and *Carassiodes spp* in south east Asia (Li, 1999) and the fact European common carp is the only member of the subfamily, Cyprininae to occur outside of Asia (Chapter 6).

7.2 Genetic variation and domestication of Vietnamese common carp

Common carp is one of the most important cultured fish species in Vietnam (Nguyen, Ngo, 2001). The aquaculture of common carp is based on indigenous and introduced strains and thus it is important that the management of stocks and genetic improvement programes utilize the best information to ensure benefits accrue to small scale farmers who depend on this species for their livelihood and food security.

While common carp have been studied using a wide variety of molecular markers in many parts of the world, detailed knowledge of the population genetics of common carp is generally lacking, due to limited population sampling. One of the few

exceptions is Kohlmann *et al*'s (2005) study; however this study mostly focused on European and Central Asian populations and the sampling of a number of populations was sub-optimal. Thus almost no information on the genetic diversity of Vietnamese common carp was available prior to this study.

This thesis presents the first comprehensive population genetic study on common carp in Vietnam, using markers that have been previously used for common carp (e.g. sequences of mtDNA *CR* and microsatellites) and other markers, including direct DNA sequencing of the mtDNA ATPase genes and the SSCP procedure that have not been applied to this species before (Chapters 2 and 4).

In general, the mtDNA and nuclear DNA are consistent and showed Vietnamese wild common carp populations to be genetically homogeneous, but distinct from other strains from both Europe and Asia. In addition, these studies provided evidence of mixing between indigenous and introduced common carp hatchery stocks consistent with the Vietnamese Government's genetic improvement and dissemination program. An important finding was the asymmetry in the frequency of the haplotypes and alleles characterising the two introduced strains. While the haplotypes and alleles characterising the Indonesian yellow carp were quite common, those typical of the Hungarian common carp were surprisingly rare. This suggests that Hungarian pure stock or crossbred stocks have been disseminated to a much lesser extent than realised or these stocks have a significantly reduced survival compared to the Vietnamese and Indonesian derived stocks.

The examination of variation in both mitochondrial and nuclear markers documented substantially reduced genetic diversity in experimental lines (Hungarian, Indonesian yellow, and Vietnamese white) maintained for a number of generations in the RIA 1 research ponds. Reduced genetic variation in aquaculture stocks such as these have been observed in other species and can be explained by founder effects, inbreeding and genetic drift .Therefore, there is a clear need to improve breeding and hatchery programmes for preserving the genetic variation of important experimental and cultured common carp stocks in Vietnam.

An impotant direction for future research for both common carp and other aquaculture species is to examine the relationship between levels of genetic variation measured by molecular genetic markers and aquaculture performance. Do, for example, stocks hypothesised to have been subject to inbreeding and low N_e show reduced growth, survival and response to selection compared to those stocks that appear to have been formed from crossing breeding? Such tests should be carried out under communal stocking condtions in appropriate experimental situations that mimic commercial condtions or better still in on-farm environments.

Lastly, the population genetic analyses of Vietnamese wild carp identified distinct haplotypes and genetically divergent populations. This indicates there is a significant genetic resource provided by wild common carp in Vietnam and steps need to be taken to ensure that this resource is preserved for the future.

7.3 Utility of SSCPs and Microsatellites

The advances in molecular genetics have led to the development of a number of new techniques over the last two decades. Data generated from one technique may be better than others for addressing particular problems, but no technique is best under all circumstances (Hillis *et al.*, 1996a). In this study, the utility of data generated from sequencing, SSCPs and microsatellite techniques can be evaluated within the context of addressing population genetic questions related to common carp in Vietnam.

The SSCP method is novel approach to the investigation of the population genetics of common carp. In this study, the SSCP technique was sufficiently powerful to allow effective comparisons among Vietnamese samples, even though the extent of nucleotide variability in the *CR* fragment was relatively limited. The SSCP phenotypic diversity revealed significant and contrasting patterns of variation within and between experimental lines, hatchery and wild populations. In addition, the SSCPs technique provides a cheap and rapid method for distinguishing Vietnamese from non-Vietnamese carp (e.g Hungarian, Indonesian, Chinese, and Japanese).

The use of multiple data sets and information from different molecular markers has becoming more common in aquaculture research (Cross *et al.*, 2000; Liu, Cordes, 2004). For this reason, the microsatellite technique was used to evaluate the population genetics of Vietnamese common carp to provide a comparison with the results of the SSCP analyses based on mtDNA. As with many studies on fish, the microsatellites provided very useful data even though it was based on variation at only four loci. While there were some differences between the two data sets the

microsatellite data showed an extraordinary level of congruence with the SSCP - mtDNA data; thus effectively distinguishing the experimental strains and giving very strong evidence for the mixing of stocks in a number of hatcheries, and the asymmetrical dissemination/survival of the introduced strains and inbreeding within experimental lines.

7.4 Utility of Micochondrial DNA gene regions

Recently, direct sequencing techniques and data from a greater number of gene regions are being used to optimize phylogenetic signal and construct more robust organismal phylogeny (Creer *et al.*, 2003; Pamilo, Nei, 1998). The examination of multiple gene regions not only ensures greater confidence in results, but also allowed the utility of the individual gene regions to be assessed. In this study the mtDNA, *ATPase6*/*ATPase8*, *Cyt b*, and *CR* gene regions have been used to reconstruct relationships between common carp strains or populations, while *16S*, *Cyt b* and *CR* have been used to examine taxonomic and evolutionary relationship among cyprinid species, including common carp.

The widely utilized *16S* gene was very effective for resolving generic and subfamily level relationships, however this gene did not effectively resolve the shallower relationships such as amongst common carp strains and between closely related species. The reasons for the lack of resolution are uncertain as it has been effective in other taxonomic groups (eg Austin *et al*. 2003), however it may reflect the slower evolution of the *16S* gene in fishes and, that common carp has a particularly shallow evolutionary history (see below).

While the mitochondrial protein coding *ATPase6/ATPase8* and *Cyt b* showed intermediate levels of variation, the *CR* was by far the most variable region examined in this study. Nevertheless levels, of *CR* variation were found to be extremely low in common carp. The low level of *CR* divergence could be due to a slower rate of evolution for this mtDNA fragment in carp; however, it most likely reflects a shallow evolutionary history for common carp because all other mtDNA gene regions also exhibited reduced variation among samples as did also the microsatellite loci.

7.5 Classification and evolution of the Cyprinidae

Molecular data also proved useful for studies at deeper evolutionary levels and indicate that the taxonomic problems encountered within common carp also extend to the higher level classification and establishment of phylogenetic relationships within the cyprinids (Chapter 6). Sequences of mtDNA *CR*, *Cyt b*, and *16S* gene fragments from little known Vietnamese cyprinids and sequences of cyprinids from GenBank were combined with those from common carp. This analysis clearly demonstrated that *Cyprinus* shows very low levels of intra-generic divergence compared to other cyprinid genera thus providing further support for the recognition of a single *Cyprinus* species.

With respect to deeper relationships, the phylogenetic analyses supported the traditional division of Cyprinidae into two major lineages: the cyprinis and leuciscins, and helped resolve controversy over the evolutionary affinities of the Danioninae (Rasborinae). However, many recognised subfamilies, however, were not recoverable as monophyletic groups and levels of genetic divergence within genera varied widely, indicating that the difficulties in the interpretation of relationships

based on morphology encountered in common carp extends to the higher taxonomic levels in the cyprinids. Further molecular phylogenetic studies are required to define well supported monophyletic groups within the major cyprinid lineage that can be associated with existing named groups and morphological information. Such studies will need substantial taxon sampling to ensure the generality of the results and sustainable taxonomic conclusion and would benefit from the use of nucleotide data from additional gene regions, especially those from the nuclear genome.

7.6 Further study

While this study has significantly enhanced the understanding of the population genetics, domestication, phylogenetic relationships and taxonomy of common carp, especially in relation to Vietnamese stocks, there are still many areas that require further research. The lack of comprehensive molecular genetic knowledge, predetermined that a number of aspects of this study would be essentially exploratory. However as a result of the data collected in this study there is now a framework on which subsequent molecular and morphological studies can be based.

In general, proper genetic, phylogenetic, biogeographical and ecological studies cannot be conducted without a stable species level taxonomy. It is thus extremely important that further molecular taxonomic studies should be conducted and coupled with more extensive morphological studies similar to those carried out by Farias *et al*. (2000) and Lopez-Fernandez *et al*. (2005a) in cichlid fish studies. There are still a number of named species of *Cyprinus*, especially from China (Zhou, Chu, 1986) and a lesser extent, Vietnam (Kottelat, 2001; Nguyen, Ngo, 2001), that need to be

examined to determine their taxonomic status. The data collected in this study will provide a sound basis and a source of comparative information for such studies.

While we now have a more substantial framework for assessing the relationships of common carp strains, there are still knowledge gaps in relation to this species. Common carp most likely reaches its greatest diversity in China and it appears that these stocks have not so far been adequately sampled and described in terms of both their genealogical relationships and population genetics. There are also other countries such as Indonesia that also have significant aquaculture production of common carp and are yet to have their common carp stocks genetically characterised to any significant extent. New population genetic studies of common carp can now take advantage of the SSCP technique used in this study and can make use of the increasing number of microsatellite loci that have been developed and the increasing data that are being accumulated on their variability (Chapter 5, Kohlmann *et al*. 2005). A logical extension of these microsatellite studies would be to construct a genomic map for common carp and investigate the potential for the identification of QTL and MAS as means of enhancing aquaculture productivity (Davis, Hetzel, 2000; Dunham, 2004; Lande, Thompson, 1990; Reid *et al.*, 2005).

Considerable scope exists for extending the generic and subfamily level phylogenetic studies to the whole family in order to provide a stable framework for establishing a sustainable taxonomic and evolutionary framework for the Cyprinidae. While mtDNA sequences from different gene regions will undoubtedly be useful for such a task, the mtDNA molecule is but a single a locus and can represent just one gene tree. For this reason, mtDNA sequences need to be combined with sequences from nuclear

genes such as recombination-activating genes, RAG1(San Mauro *et al.*, 2004) and RAG2 (Lopez-Fernandez *et al.*, 2005b), to ensure robust phylogenies are generated.

In addition, more recently developed molecular techniques such as AFLPs and SNPs should be used to investigate Vietnamese common carp populations in further studies. These techniques will provide more information to help complete the picture of the genetics of common carp in Vietnam.

APPENDICES

Appendix 1.

Thai, T. B., C. P. Burridge., T. A. Pham., C. M. Austin. (2004). Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.). Animal Genetic 36, 23-28.

ASBTRACT

Direct sequencing of mitochondrial DNA *CR* region (745 bp) and *MTATPase6/MTATPase8* (857 bp) regions was used to investigate genetic variation within common carp and develop a global genealogy of common carp (carp) strains. The *CR* region was more variable than the *MTATPase6/MTATPase8* region, but given the wide distribution of carp the overall levels of sequence divergence were low. Levels of haplotype diversity varied widely among countries with Chinese, Indonesian, Vietnamese carp showing the greatest diversity and Japanese Koi and European carp without detectable nucleotide variation. A genealogical analysis supports a close relationship between Vietnamese, Koi and Chinese Color carp strains and to a lesser extent, European carp. Chinese and Indonesian carp strains were the most divergent, and their relationships do not support the evolution of independent Asian and European lineages and current taxonomic treatments.

Appendix 2.

B.T. Thai, T. A. Pham, U. D. Thai, C.M. Austin (2006) Progress towards a global genealogy of common carp (*Cyprinus carpio* L.) strains using mitochondrial nucleotide sequences data. NAGA 29 (3&4), 55-61.

ASBTRACT

As part of a study of genetic variation in Vietnamese strains of common carp (*Cyprinus carpio* L.) using direct DNA sequencing of mitochondrial *CR* and *ATPase6/ATPase8* gene regions, samples from a number of other countries: China, Japan, Indonesia, India, Hungary, Czech Republic, Israel and Australia were analysed for comparative purposes. Results show that the levels of sequence divergence in common carp is low on a global scale with the Asian carp having the highest diversity, whereas Koi and European carp were invariant. A genealogical analysis supports a close relationship among Vietnamese, Koi and Chinese Color and to a lesser extent, European carp. Koi carp appear to have originated from a strain of Chinese red carp. There is considerable scope to extend this research through the analysis of additional samples of carp from around the world and especially China in order to generate a comprehensive global genealogy of common carp strains.

Appendix 3.

B.T. Thai., T. A. Pham., C.M. Austin (2006). Genetic diversity of common carp in Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA control region. Aquaculture 258, 228-240.

ASBTRACT

The aquaculture of common carp (*Cyprinus carpio* L.) in Vietnam is based upon both indigenous and introduced stocks which have not been subject to study using modern molecular genetic methods to any significant extent. Twenty strains or populations of common carp in Vietnam represented by 968 fish were screened for variation in mitochondrial *CR* region fragments using a combination of direct DNA sequencing and Single Strand Confirmation Polymorphism (SSCP) analysis. Common carp samples from China, Japan, Indonesia, and Hungary were analysed for comparison. Sequencing revealed that Vietnamese common carp have high haplotype, but low nucleotide diversity and represent a mixture of indigenous and introduced strains. The SSCP procedure resolved eight haplotypes that distinguished Indonesian, Hungarian and Vietnamese strains and which varied significantly among hatchery and wild Vietnamese common carp populations. Intra-population diversity was least in experimental common carp lines (Hungarian scale, Indonesian yellow and Vietnamese white), maintained at the Research Institute for Aquaculture No. 1 and greatest within hatchery stocks. Relationships among populations based on SSCP haplotype frequencies showed that the RIA1 Vietnamese white common carp strain is closely related to wild populations, which in turn are closely related to six of the eleven hatchery stocks. The other five hatchery stocks had higher levels of interpopulation variation and mostly showed a closer relationship to Indonesian

yellow carp strains. The Hungarian carp strain was highly divergent from all other populations suggesting that this strain has not contributed significantly to the establishment of cultured stocks in Vietnam, based on this maternally inherited marker. The SSCP procedure shows considerable potential for rapid genotyping and genetic characterization of common carp and therefore, for investigation of diversity in wild stocks and broodstock management.

Appendix 4

B. T. Thai, C. P. Burridge, C. M. Austin (2006). Insights into the management and dissemination of domesticated strains of common carp (*Cyprinus carpio* L.) in Vietnam revealed by variation at four microsatellite loci. Aquaculture submitted.

ASBTRACT

Four highly variable microsatellite loci were used to investigate genetic diversity and population structure of common carp in Vietnam. A total of 968 fish were genotyped representing three groups comprising: three experimental lines from the Research Institute for Aquaculture No 1 (Bac Ninh); 11 hatcheries; and six wild populations from rivers and reservoirs giving 72 alleles over all loci. The mean number of alleles, per locus per population ranges from 4.25 to 11.00 and the mean observed heterozygosity at the four loci ranges from 0.40 to 0.83. An analysis of the distribution of genetic variation indicated within population variation is very high (90.6%), while among populations within groups and among groups is low (5.0% and 4.5% respectively). Highly significant deviations from Hardy-Weinberg, mostly due to deficits of heterozygotes, were found in both experimental and hatchery groups suggesting either inbreeding or recent stock mixing. Wild common carp populations exhibited more genetic diversity than cultured populations in term of allele richness and observed heterozygosity. Results from assignment tests for the 20 populations of carp indicated that the experimental common carp lines can be largely distinguish from one another and that mixing between indigenous and introduce carp is occurring in the hatchery and possibly also in wild populations. Multidimensional Scaling (MDS) and UPGMA analyses show that the experimental Vietnamese white carp line is closely related to wild common carp populations; the hatchery stocks are

mostly closely related to the experimental Indonesian yellow carp line but with evidence of some mixing; and the domesticated Hungarian population is highly divergent and not closely related to any other carp population.

BIBLIOGRAPHY

Agresti JJ, Seki S, Cnaani A*, et al.* (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 FW, Phelps S (1980) Loss genetic variation in hatchery stock of cutthroat trout. *Transactions of the American Fisheries Society* **109**, 537-543.

Amano M (1968) *Colourful 'Live Jewels". General Survey of Fancy Carp* Kajima Shoten Publishing Co, Tokyo.

Anjum R (1995) *Biochemical and chromosomal genetic characteristics of several breeding populations of common carp, Cyprinus carpio L.* PhD, Olsztyn University of Agriculture and Technology.

Apte S, Gardner JPA (2002) Population genetic subdivision in the New Zealand greenshell mussel (*Perna canaliculus*) inferred from single-strand conformation polymorphism analysis of mitochondrial DNA. *Molecular Ecology* **11**, 1617-1628.

Apte S, Star B, Gardner APJ (2003) A comparison of genetic diversity between cultured and wild populations, and a test for genetic introgression in the New Zealand greenshell mussel *Perna canaliculus* (Gmelin 1791). *Aquaculture* **219**, 193-220.

Aurelle D, Berrebi P (2001) Genetic structure of brown trout (*Salmo trutta* L.) populations from south-western France: data from mitochondrial control region variability. *Molecular Ecology* **10**, 1551-1561.

Austin CM, Nguyen TTT, Meewan MM, Jerry DR (2003) The taxonomy and phylogeny of the '*Cherax destructor*' complex (Decapoda : Parastacidae)

examined using mitochondrial 16S sequences. *Australian Journal of Zoology* **51**, 99-110.

Avise JC (1974) Systematic value of electrophoretic data. *Systematic Biology* **23**, 465-481.

Avise JC (1983) Protein variation and phylogenetic construction. In: *Protein Variation: Adaptive and Taxonomic Significance* (eds. Oxford G, Rollinson D). Systematic Association, British Museum Natural History, London.

Avise JC (1994) *Molecular Markers, Natural History, and Evolution* Chapman and Hall, New York.

Avise JC (2000) *Phylogeography - The History and Formation of Species* Harvard University Press, United States of America.

Bakos J, Gorda S (2001) Genetic resources of common carp at the Fish Culture Research Institute Szarvas, Hungary. *FAO Fisheries Technical Paper, Rome*, 106.

Balon EK (1995) Origin and domestication of wild carp, *Cyprinus carpio*: from Roman gourmets to the swimming flowers. *Aquaculture* **129**, 3-48.

Balon EK (2004) About the oldest domesticates among fishes. *Fish Biology* **65**, 1-27.

Baranski M, Rourke M, Loughnan S, Robison N (2006) Isolation and characterization of 125 microsatellite DNA marker in blacklip abalone *Haliotis rubara*. *Molecular Ecology Note* **6**, 740-746.

Bardakci F, Skibinski DOF (1994) Application of the rapd technique in tilapia fish - species and subspecies identification. *Heredity* **73**, 117-123.

Bartfai R, Egedi S, Yue GH*, et al.* (2003) Genetic analysis of two common carp broodstocks by RAPD and microsatellite markers. *Aquaculture* **219**, 157-167.

Baruš V, Penaz M, Kohlmann K (2002) *Cyprinus carpio* (Linnaeus, 1758). In: *Freshwater Fishes of Europe* (eds. Banarescu PM, Paepke HJ), pp. 85-179. Academic Press, London.

Baum DA, Shaw KL (1995) *Genealogical perspectives on the species problem*, Missouri.

Baverstock PR, Moritz C (1996) Sampling. In: *Molecular Systematics* (eds. Hillis DM, Moritz C, Mable BK), pp. 17-28. Sinaer Associates, Sunderland.

Beardmore AL, Mair CG, Lewis CG (1997) Biodiversity in aquatic systems in relation to aquaculture. *Aquaculture Research* **28**, 829-839.

Beaumont RA, Hoare K (2003) *Biotechnology and Genetics in Fisheries and Aquaculture* Blackwell Publishing, Oxford.

Ben-Dom N, Cherfas NB, Gomelsky B, Hulata G (2000) Genetic stability of Israeli common carp stocks inferred from electrophoretic analysis of transferring, phosphoglucomutase and glucose-6-phosphate isomerase. *Isr. J. Aquac. Bamidgeh* **52**, 30-35.

Bernal-Ramirez JH, Adcock GJ, Hauser L, Carvalho GR, Smith PJ (2003) Temporal stability of genetic structure in the New Zealand snapper, *Pagrus auratus*, and relationships to coast currents. *Marine Biology* **142**, 567-574.

Blackwelder RE (1967) *Taxonomy: A text and Reference Book* Wiley, New York.

Blaxter ML (2004) The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society of London* **359**, 669-679.

Bouza C, Sanchez L, Martinez P (1997) Gene diversity analysis in natural populations and cultured stocks of turbot (*Scophthalmus maximus* L). *Animal Genetics* **28**, 28-36.

Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular phylogeny of Cyprinidae inferred from cytochrome b DNA sequences. *Molecular Phylogenetics and Evolution* **9**, 100-108.

Brody T, Kirsht D, Parag G*, et al.* (1979) Biochemical genetic comparison of the Chinese and European races of the common carp. *Animal Blood Group Biochemical. Genetic* **10**, 141-149.

Buckley TR, Arensburger P, Simon C, Chambers GK (2002) Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. *Systematic Biology* **51**, 4-18.

Buth DG (1984) The application of electrophoretic data in systematic studies. *Annual Review of Ecology and Systematics* **15**, 501-522.

Cataudella S, Sola L, Corti M*, et al.* (1987) Cytogenetic, genic and morphometric charaterization of groups of common carp, *Cyprinus carpio*. In: *Proc. World Symp. On Selection, Hybridization, and Genetic Enginerring in Aquaculture* (ed. Tiews K), pp. 113-129. Schriften der Bundesanstalt fur Fischerei Hamburg, Bordeaux.

Cavender T, Coburn MM (1992) Phylogenetic relationships of Norht American Cyprinidea. In: *In Systermatics, Historical Ecology and North American freshwater fishes* (ed. Mayden LR), pp. 293-327. Stanford University Press, Stanford, California.

Chang Y, Huang HD, Lo CF (1994) The complete nucleotide sequence and gene organisation of carp (*Cyprinus carpio*) mitochondrial genome. *Molecular Evolution* **38**, 138-155.

Chen CA, Ablan MC, McManus JW*, et al.* (2004) Population structure and genetic variability of six bar wrasse (*Thallasoma hardwicki*) in Northern South China

Sea revealed by mitochondrial control region sequences. *Mar Biotechnol* **6**, 312-326.

Chen I, Hsu C-H, Hui C-F*, et al.* (1998) Sequence length and variation in the mitochodrial control region of two freshwater fishes belong to Rhinogobius (Teleostei: Gobioidei). *J. Fish. Biology* **53**, 179-191.

Chen X, Yue P, Lin R (1984) Major group within the family Cyprinidae and their phylogenetic relationships. *Acta Zootaxon. Sin* **9**, 424-440.

Chiba TN, Kurihama, Kenichi Y (1966) *Culture of common carp* Green Book House Publisher, Tokyo.

Claridge MA, Dawah HA, Wilson MR (1997) Species: The Units of Diversity. Chapman and Hall, London.

Clifford LS, McGinnity P, Ferguson A (1998) Genetic Changes in Atlantic salmon (*Salmo salar*) populations of Northwest Irish rivers resulting from escapes of adult salmon. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 358-363.

Cooper EL (1987) *Carp in North America* Maryland: Bethesda.

Cracraft J (2000) Species concepts in theoretical and applied biology: A systematic debate with consequences. In: *Species Concepts and Phylogenetic Theory* (eds. Wheeler QD, Meier R). Columbia University Press, New York.

Crandall KA, Fetzner JW, Lawler SH, Kinnersley M, Austin CM (1999) Phylogenetic relationships among the Australian and New Zealand genera of freshwater crayfishes (Decapoda : Parastacidae). *Australian Journal of Zoology* **47**, 199-214.

Creer S, Malhotra A, Thorpe RS (2003) Assessing the phylogenetic utility of four mitochondrial genes and a nuclear intron in the Asian pit viper genus,

Trimeresurus: Separate, simultaneous, and conditional data combination analyses. *Molecular Biology and Evolution* **20**, 1240-1251.

Cronin MA (2003) Research on deer taxonomy and its relevance to management. *Ecoscience* **10**, 432-442.

Crooijmans RPM, Bierbooms V, Komen J, Vanderpoel JJ, Groenen M (1997) Microsatellite markers in common carp (*Cyprinus carpio* L). *Animal Genetics* **28**, 129-134.

Cross T, Dillance E, Galvin P (2000) Which molecular markers should be chosen for different specific applications in fisheries and aquaculture? National University of Ireland.

Cross T, Dillance E, Galvin P (2003) Which molecular markers should be chosen for different specific applications in fisheries and aquaculture? National University of Ireland.

Cross TF (2000) Genetic implications of translocation and stocking of fish species, with particular reference to Western Australia. *Aquaculture Research* **31**, 83-94.

Cross TF, Coughlan J, Burnell G*, et al.* (2005) Utility of microsatellite loci for detecting reduction of variation in reared aquaculture strains compared with wild progenitors and also as genetic "tags" in breeding programmes: evidence from abalone, halibut and salmon. *Aquaculture* **247**, 9-10.

Csizmadia C, Jeney Z, Szerencses I, Gorda S (1995) Transferrin polymorphism of some races in a live gene bank of common carp. *Aquaculture* **129**, 193-198.

Cunha C, Mesquita N, Dowling TE, Gilles A, Coelho MM (2002) Phylogenetic relationships of Eurasian and American cyprinids using cytochrome b sequences. *Journal of Fish Biology* **61**, 929-944.

Daniel D, Christian S (2003) Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes*, 167-169.

David L, Rajasekaran P, Fang J, Hillel J, Lavi U (2001) Polymorphism in ornamental and common carp strains (*Cyprinus carpio* L.) as revealed by AFLP analysis and new set of microsatellite markers. *Molecular genetic Genomics* **266**, 353-362.

Davis GP, Hetzel DJS (2000) Integrating molecular genetic technology with traditional approaches for genetic improvement in aquaculture species. *Aquaculture Research* **31**, 3-10.

Davis KM, Dixon PI, Harri JH (1999) Allozyme and mitochondrial DNA analysis of carp, *Cyprinus carpio* L., from south-eastern Australia. *Marine Freshwater Research* **50**, 253-260.

De Francisco AK, Galetti PM (2005) Genetic distance between broodstocks of the marine shrimp *Litopenaeus vannamei* (Decapoda, Penaeidae) by mtDNA analyses. *Genetics and Molecular Biology* **28**, 258-261.

Desvignes JF, Laroche J, Durand JD, Bouvet Y (2001) Genetic variability in reared stocks of common carp (*Cyprinus carpio* L.) based on allozymes and microsatellites. *Aquaculture* **194**, 291-301.

Diniz FM, Maclean N, Ogawa M, Cintra IH, Bentzen P (2005) The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. *Mar Biotechnol* **7**, 462-473.

Dunham AR (2004) *Aquaculture and Fisheries Biotechnology Genetic Approaches* CABI Publishing, Oxfordshire.

Duran S, Palacin C, Becerro MA, Turon X, Giribet G (2004) Genetic diversity and population structure of the commercially harvested sea urchin *Paracentrotus lividus* (Echinodermata, Echinoidea). *Molecular Ecology* **13**, 3317-3328.

Durand JD, Tine M, Panfili J, Thiaw OT, Lae R (2005) Impact of glaciations and geographic distance on the genetic structure of a tropical estuarine fish, *Ethmalosa fimbriata* (Clupeidae, S. Bowdich, 1825). *Molecular Phylogenetics and Evolution* **36**, 277-287.

Durand JD, Tsigenopoulos CS, Unlu E, Berrebi P (2002) Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA - Evolutionary significance of this region. *Molecular Phylogenetics and Evolution* **22**, 91-100.

Dzialuk A, Chybicki I, Burczyk J (2005) PCR multiplexing of nuclear microsatellite Loci in *Quercus* Species. *Plant Molecular Biology Reporter* **23**, 121-128.

Edwards P, Hiep DD, Anh PM, Mair G (2000) Traditional culture of indigenous common carp in rice fields in northern Vietnam. *World Aquaculture* **31**, 34-40.

Ellegren H (1992) Polymerase Chain Reaction (PCR) analysis of microsatellites a new approach to studies of genetic relationships in birds. *Auk* **109**, 886-895.

Elliott NG, Reilly A (2003) Likelihood of bottleneck event in the history of the Australian population of Atlantic salmon (*Salmo salar* L.). *Aquaculture* **215**, 31-44.

Emery AM, Wilson IJ, Craig S, Boyle PR, Noble LR (2001) Assignment of paternity groups without access to parental genotypes: multiple mating and developmental plasticity in squid. *Molecular Ecology* **10**, 1265-1278.

Evans B, Bartlett J, Sweijd N, Cook P, Elliott NG (2004) Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South Africa (*Haliotis midae*). *Aquaculture* **233**, 109-127.

FAO (1998) Database on introductions of aquatic species. Food and Agriculture Organization of the United Nations, Rome, Italy. http://www.fao.org/waicent/faoinfo/fishery/statist/fisoft/dias/index.htm.

FAO (2001) FAO/SIDP - species identification and data programme http://www.fao.org/figis/servlet/Fred?ds=species&fid=2957.

FAO (2003) Yearbook of fisheries statistic summary table. Available at: htt://ftp.fao.org/FI/STAT/summary/default.htm.

FAO (2004) Cultured Aquatic Species Information Programme *Cyprinus carpio* http://www.fao.org/figis/servlet/static?dom=culturespecies&xml=Cyprinus_carpio.xm.

Farias IP, Orti G, Meyer A (2000) Total evidence: Molecules, morphology, and the phylogenetics of cichlid fishes. *Journal of Experimental Zoology* **288**, 76-92.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783-791.

Felsenstein J (1993) PHYLIP version 3.50. Seattle, WA, USA. P. O. Box 357360.

Ferguson MM (1995a) Molecular genetics in fisheries: current and future perspectives. In: *Molecular Genetics in Fisheries* (eds. Carvalho GR, Pitcher TJ), pp. 110-115. Chapman & Hall, Great Britain.

Ferguson MM (1995b) The role of molecular genetic markers in the management of cultured fishes. In: *Molecular Genetics in Fisheries* (eds. Carvalho GR, Pitcher TJ), pp. 81-104. Chapman & Hall, Great Britain.

Ferguson MM, Danzmann RG (1998) Role of genetic markers in fisheries and aquaculture: useful tools for stamp collecting? *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 1553-1563.

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies* **326**, S22-S29.

Froufe E, Magyary I, Lehoczky I, Weiss S (2002) MtDNA sequence data supports an Asian ancestry and single introduction of common carp into Danube Basin. *Fish Biology* **61**, 301-304.

Fuchs BH, Schlee P, Blusch J*, et al.* (2000) Short Communication Phylogenetic studies in cyprinid species from central Europe by cytochrome b sequences of mitochondrial DNA. *J Appl Ichthyol* **16**, 79-82.

Garoia F, Guarniero I, Piccinetti C, Tinti F (2004) First microsatellite loci of red mullet (*Mullus barbatus*) and their application to genetic structure analysis of Adriatic shared stock. *Mar Biotechno* **6**, 446-452.

Gharbi K, Gautier A, Danzmann RG*, et al.* (2006) A linkage map for brown trout (*Salmo trutta*): Chromosome homeologies and comparative genome organization with other salmonid fish. *Genetics* **172**, 2405-2419.

Gilles A, Lecointre G, Faure E, Chappaz R, Brun G (1998) Mitochondrial phylogeny of the European cyprinids: Implications for their systematics, reticulate evolution, and colonization time. *Molecular Phylogenetics and Evolution* **10**, 132-143.

Gilles A, Lecointre G, Miquelis A*, et al.* (2001) Partial combination applied to phylogeny of European cyprinids using the mitochondrial control region. *Molecular Phylogenetics and Evolution* **19**, 22-33.

Giuffra E, Guyomard R, Forneris G (1996) Phylogenetic relationships and introgression patterns between incipient parapatric species of Italian brown trout (*Salmo trutta* L complex). *Molecular Ecology* **5**, 207-220.

Goldman N, Anderson JP, Rodrigo AG (2000) Likelihood-based test of topologies in phylogenetics. *Systematic Biology* **49**, 652-670.

Gosline WA (1978) Unbranched dosal-fin rays and subfamily classification of the fish family Cyprinidae. *Occas. Pap. Mus. Zoll. Univ. Mich* **684**, 1-21.

Goudet J (2002) FSTAT, A Program to estimate and test gene diversities and fixation indices (version 2.9.3). Institute of Ecology, University of Lausanne, Switzerland. http://www.unil.ch/izea/softwares/fstat.html.

Graur D, Li WH (2000) *Fundamentals of Molecular Evolution*, Second Edition edn. Sinauer Associates, Inc., Sunderland.

Gross R, Kohlmann K, Kersten P (2002) PCR-RFLP analysis of the mitochondrial ND-3/4 and ND-5/6 gene polymorphisms in the European and East Asian subspecies of common carp (*Cyprinus carpio* L.). *Aquaculture* **204**, 507-516.

Gupta VM, Dey M, M., Penman JD (2005) Importance of carp genetic resource. In: *Carp genetic resource for aquaculture in Asia* (eds. Penman JD, Gupta VM, Dey MM). WorlFish Center, Penang.

Hailer F, Gaustschi B, Helander B (2005) Development and multiplex PCR amplification of novel microsatellite markers in the White-tailed Sea Eagle, *Haliaeetus albicilla* (Aves: Falconiformes, Accipitridae). *Molecular Ecology* **5**, 938-940.

Hansen MM, Ruzzante DE, Nielsen EE, Mensberg KD (2000) Microsatellite and mitochondrial DNA polymorphism reveals life-history dependent

interbreeding between hatchery and wild brown trout (*Salmo trutta* L.). *Molecular Ecology* **9**, 853-601.

Hardjamulia A, Sukadi F, Subagyo., Gustiano R (2001) Aquaculture genetic research in Indonesia. In: *Fish genetic research in member countries and institution of the International Network on Genetic in Aquaculture* (eds., B.O.Acosta Ga), p. 179. ICLARM.

Hastein T, Hill BJ, Berthe F, Lightner DV (2001) Traceability of aquatic animals. *Revue Scientifique Et Technique De L Office International Des Epizooties* **20**, 564-583.

Hayashi K (1992) A method for detection of mutations. *GATA* **9**, 73-79.

Hayashi K, Yandell DW (1993) How sensitive is PCR-SSCP? *Human mutation* **2**, 338-346.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 313-321.

Hillis DM, Huelsenbeck JP (1992) Signal, noise and reliability in molecular phylogenetic analysis. *Heredity* **83**, 189-195.

Hillis DM, Mable BK, Moritz C (1996a) Applications of molecular systematics: The state of the field and a look to the future. In: *Molecular Systematics* (eds. Hillis DM, Moritz C, Mable BK), pp. 515-543. Sinaeur Associates, Inc., Canada.

Hillis DM, Moritz C, Mable BK (1996b) Molecular Systematics, p. 655. Sunderland, Sinauer Associates.

Hillis DM, Moritz C, Mable BK (1996c) *Molecular systematics. Sunderland, Sinauer Associates, 1-13 pp.*

Hindar K, Ryman N, Utter F (1991) Genetic-effects of cultured fish on natural fish populations. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 945-957.

Hoarau G, Piquet AMT, van der Veer HW*, et al.* (2004) Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: a comparison of resolving power between microsatellites and mitochondrial DNA data. *Journal of Sea Research* **51**, 183-190.

Hollebecq GM, Haffray P (1999) Genetic improvement. In: *Carp Biology and Culture* (ed. Billard R), pp. 101-102. Praxis Publishing, Chichester, UK.

Howes G (1991) Systermatics and biogeography: an overview. In: *Cyprinids fishes: systematic biology and exploitation* (eds. Win I, Nelson J), pp. 1-54. Chapman and Hall, New York.

Huelsenbeck JP, Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics* **28**, 437-466.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F (2002) Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology* **51**, 673-688.

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinfomatics* **17**, 754-755.

Hulata G (1995) A review of genetic improvement of the common carp (*Cyprinus carpio* L.) and other cyprinids by crossbreeding, hybridization and selection. *Aquaculture* **129**, 143-155.

Hurt AC, Murphy NP, Patil JG, Austin CM (2001) Molecular genetic evidence for a new species of bream of genus *Acantholpagrus* Peter (Perciformes: Sparidae). *Asian Fisheries Science* **14**, 425-433.

Innes BH, Elliott NG (2006) Genetic diversity in a Tasmanian hatchery population of Atlantic salmon (*Salmo salar* L.) compared with its Canadian progenitor population. *Aquaculture Research* **37**, 563-569.

Ivankovic A, Kavar T, Caput P*, et al.* (2002) Genetic diversity of three donkey populations in the Croatian coastal region. *Animal Genetics* **33**, 169-177.

Jeong SD, Uminou T, Kuroda K*, et al.* (2003) Genetic divergence and population structure of black sea bream (*Acanthopagrus schlegeli*) inferred from microsatellite analysis. *Fisheries Science* **69**, 896-902.

Jerry DR, Baverstock PR (1998) Consequences of a catadromous life-strategy for levels of mitochondrial DNA differentiation among populations of the Australian bass, Macquaria novemaculeata. *Molecular Ecology* **7**, 1003-1013.

Johnson MS (2000) Measuring and interpreting genetic structure to minimise the genetic risk of translocation. *Aquaculture Research* **31**, 133-143.

Julie T, Louis B (2003) Reticulate evolution and phenotypic diversity in North American ciscoes, Coregonus ssp.(Teleostei: Salmonidae): implications for the conservation of an evolutionary legacy. *Conservation Genetics* **4**, 67-81.

Karsi A, Patterson A, Feng JN, Liu ZJ (2002) Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression. *Gene* **291**, 177-186.

Kavar T, Habe F, Brem G, Dovc P (1999) Mitochondrial D-loop sequence variation among the 16 maternal lines of the Lipizzan horse breed. *Animal Genetics* **30**, 423-430.

Kierstein G, Vallinoto M, Silva A*, et al.* (2004) Analysis of mitochondrial D-loop region casts new light on domestic water buffalo (*Bubalus bubalis*) phylogeny. *Molecular Phylogenetics and Evolution* **30**, 308-324.

Kim KI, Lee JH, Li K*, et al.* (2002) Phylogenetic relationships of Asian and European pig breeds determined by mitochondrial DNA D-loop sequence polymorphism. *Animal Genetics* **33**, 19-25.

Kirpitchnikov VS (1967) Homologous hereditary variation and evolution of wild common carp *Cyprinus carpio* L. *Genetika* **8**, 65-72 (in Russian).

Kirpitchnikov VS (1999) Genetic and Breeding of common carp (eds. R. Billar R, Reperant J, Rio JP, Ward R), p. 97. INR, Paris.

Kishino H, Hasegawa M (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data the braching order in the Hominoidea. *Molecular Evolution* **29**, 550-570.

Kocher TD, Lee WJ, Sobolewska H, Penman D, McAndrew B (1998) A genetic linkage map of a cichlid fish, the tilapia (*Oreochromis niloticus*). *Genetics* **148**, 1225-1232.

Kohlmann K, Gross R, Murakaeva A, Kersten P (2003) Genetic variability and structure of common carp (*Cyprinus carpio*) populations throughout the distribution range inferred from allozyme, microsatellite and mitochondrial DNA markers. *Aquatic Living Resources* **16**, 421-431.

Kohlmann K, Kersten P (1999) Genetic variability of German and foreign common carp (Cyprinus carpio L.) populations. *Aquaculture* **173**, 435-445.

Kohlmann K, Kersten P, Flajshans M (2005) Microsatellite-based genetic variability and differentiation of domesticated, wild and feral common carp (*Cyprinus carpio* L.) populations. *Aquaculture* **247**, 253-266.

Koljonen ML (1989) Electrophoretically detectable genetic variation in natural and hatchery stocks of Atlantic salmon in Finland. *Hereditas* **110**, 23-35.

Komen J (1990) *Clones of common carp, Cyprinus carpio.New perspectives on fish research. PhD thesis*, Agricultural University.

Kottelat M (1997) European freshwater fishes. *Biologia (Bratislava)* **52**, 1-217.

Kottelat M (2001) Freshwater Fish of Northern Vietnam, p. 140. Environment and Social Development Unit, East Asia and Pacific Region, World Bank, Washington.

Kuznetsova MV, Kholodova MV, Danilkin AA (2005) Molecular phylogeny of deer (Cervidae: Artiodactyla). *Russian Journal of Genetics* **41**, 742-749.

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743-756.

Lee WJ (2003) Detection of QTL for salinity tolerance of tilapia. *PAG XI Abstracts (http://wwwintl-pagorg)*.

Lee WJ, Conroy J, Howell WH, Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *Journal of Molecular Evolution* **41**, 54-66.

Lehoczky I, Jeney Z, Magyary I, Hancz C, Kohlmann K (2005) Preliminary data on genetic variability and purity of common carp (*Cyprinus carpio* L.) strains kept at the live gene bank at Research Institute for Fisheries, Aquaculture and Irrigation (HAKI) Szarvas, Hungary. *Aquaculture* **247**, 45-49.

Lever C (1996) *Naturalised Fishes of the World* CA: Academic Press, San Diego.

Li L, Guo XM (2004) AFLP-based genetic linkage maps of the Pacific oyster *Crassostrea gigas* Thunberg. *Marine Biotechnology* **6**, 26-36.

Li L, Xiang JH, Liu X*, et al.* (2005) Construction of AFLP-based genetic linkage map for Zhikong scallop, Chlamys farreri Jones et Preston and mapping of sex-linked markers. *Aquaculture* **245**, 63-73.

Li S-F (1999) Freshwater Fish Genetic Resources and Conservation Approaches. In: *Genetic in Sustainable Fisheries Management* (ed. Saleem M). Fishing News Books, Koto Kinabalu.

Li S-F, Wanqi C, Shuming Z, Chenghui W (2001) Genetic research program and panned activities in Shanghai Fisheries University (SFU), China. Report submitted to the Sixth Steering Committee Meeting of International Network on genetics in Aquaculture (INGA), 8-10 May 2001, Hanoi, Vietnam.

Li W-H (1997) *Molecular Evolution* Sinauer Associates Inc., Sunderland.

Linhart O, Gela D, Rodina M, Slechtova V, Slechta V (2002) Top-crossing with paternal inheritance testing of common carp (*Cyprinus carpio* L.) progeny under two altitude conditions. *Aquaculture* **204**, 481-491.

Linhart O, Kudo S, Billard R, Slechta V, Mikodina E (1995) Morphology composition and fertilization of carp eggs: A review. *Aquaculture* **129**, 75-93.

Liu H, Tzeng CS, Teng HY (2002) Sequences variation in the mtDNA control region and their implications for the phylogeny of the Cypriniformes. *Canadian Journal of Zoology* **80**, 569-581.

Liu HZ (2002) Mitochondrial DNA sequence variations and systematics of the genus Distoechodon (Teleostei : Cyprinidae). *Journal of Applied Ichthyology* **18**, 181-184.

Liu HZ, Chen YJ (2003) Phylogeny of the East Asian cyprinids inferred from sequences of the mitochondrial DNA control region. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **81**, 1938-1946.

Liu ZJ, Cordes FJ (2004) DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **238**, 1-37.

Liu ZJ, Karsi A, Li P, Cao DF, Dunham R (2003) An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. *Genetics* **165**, 687-694.

Long JR, Qiu XP, Zeng FT, Tang LM, Zhang YP (2003) Origin of rabbit (*Oryctolagus cuniculus*) in China: evidence from mitochondrial DNA control region sequence analysis. *Animal Genetics* **34**, 82-87.

Lopez-Fernandez H, Honeycutt RL, Stiassny MLJ, Winemiller KO (2005a) Morphology, molecules, and character congruence in the phylogeny of South American geophagine cichlids (Perciformes, Labroidei). *Zoologica Scripta* **34**, 627-651.

Lopez-Fernandez H, Honeycutt RL, Winemiller KO (2005b) Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes : Labroidei). *Molecular Phylogenetics and Evolution* **34**, 227-244.

Lovette I, J., Bermingham E, Seutin G, Ricklest R, E. (1998) Evolutionary differentation in three endemic West Indian warblers. *Auk* **115**, 890-903.

Mabuchi K, Miya M, Senou H, Suzuki T, Nishida M (2006) Complete mitochondrial DNA sequence of the Lake Biwa wild strain of common carp (*Cyprinus carpio* L.): further evidence for an ancient origin. *Aquaculture* **257**, 68-77.

Mabuchi K, Senou H, Suzuki T, Nishida M (2005) Discovery of an ancient lineage of *Cyprinus carpio* from Lake Biwa, central Japan, based on mtDNA sequence data, with reference to possible multiple origins of koi. *Journal of Fish Biology* **66**, 1516-1528.

Macaranas JM, Sato J, Fujio Y (1986) Genetic characterization of culture populations of Japanese common carp. *Tohoku J. Agric. Res* **37**, 21-29.

Mai DY (1978) *Freshwater Fish in Vietnam* Science and Technology Publisher, Hanoi.

Mai DY, Nguyen VT, Le HY, Nguyen VT (1992) *Taxonomy of Freshwater Fish in Southern Vietnam* Science and Technology Publisher, Hanoi.

Manjunath BJ, Pramod KR, Ajoy MK*, et al.* (2004) Phylogeography and origin of Indian domestic goats. *Molecular Biology and Evolution* **21**, 454-462.

Marklund S, Chaudhary R, Marklund L, Sandberg K, Anderson L (1995) Extensive mtDNA diversity in horses revealed by PCR-SSCP analysis. *Animal Genetics* **26**, 193-196.

Maxson LR, Maxson RD (1990) Protein II: Immunological Techniques. In: *Molecular Systematics* (eds. Hillis DM, Moritz C), pp. 127-155. Sinauer, Sunderland, MA.

Mayden RL (1997) Species. In: *The Units of Biodiversity* (eds. Claridge MF, Dawah HA, Wilson MR), p. 381–424. Chapman & Hall, London.

Mayr E (1963) *Animal Species and Evolution* The Belknap Press, Cambridge, MA.

Mayr E (2000) The biological species concept. In: *Species concepts and phylogenetic theory* (eds. Wheeler QD, Meier R), pp. 17-19. Columbia University, New York.

Mayr E (2002) Classification and other ordering systems. *Jornal of Zoological Systematic and Evolutionary Research* **40**, 169-194.

McConnell SKJ, Beynon C, Leamon J, Skibinski DOF (2000) Microsatellite marker based genetic linkage maps of *Oreochromis aureus* and *O.niloticus* (Cichlidae): extensive linkage group segment homologies revealed. *Animal Genetics* **31**, 214-218.

Mickett K, Morton C, Feng J*, et al.* (2003) Assessing genetic diversity of domestic populations of channel catfish (*Ictalurus punctatus*) in Alabama using AFLP markers. *Aquaculture* **228**, 91-105.

Miller MP (1997) Tools For Population Genetic Analysis (TFPGA) 1.3: A windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by author.

Mirol MP, Garcia PP, Dulout NF (2002a) Mitochondrial variability in the D-loop of four equine breeds shown by PCR-SSCP analysis. *Genetics and Molecular Biology* **25**, 25-28.

Mirol PM, Garcia PP, Vega-Pla JL, Dulout FN (2002b) Phylogenetic relationships of Argentinean Creole horses and other South American and Spanish breeds inferred from mitochondrial DNA sequences. *Animal Genetics* **33**, 356-363.

Miya M, Nishida M (2000) Use of mitogenomic information in teleostean molecular phylogenetics: a tree-base exploration under the maximum-parsimony optimality criterion. *Molecular Phylogenetics and Evolution* **17**, 437-455.

Mjolnerod IB, Refseth UH, Karlsen E*, et al.* (1997) Genetic differences between two wild and one farmed population of Atlantic salmon (*Salmo salar*) revealed by three classes of genetic markers. *Hereditas* **127**, 239-248.

Mullis K, Faloona F, Scharf S*, et al.* (1986) Specific enzymatic amplication of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Sym. quant. Biol* **51**, 263-273.

Murakaeva A, Kohlmann K, Kersten P, Kamilov B, Khabibullin D (2003) Genetic characterization of wild and domesticated common carp (*Cyprinus carpio* L.) populations from Uzbekistan. *Aquaculture* **218**, 153-166.

Murakami M, Yamashita Y, Fujitani H (1998) The complete sequence of mitochondrial genome from a gynogenetic triploid "ginbuna" (*Carassius auratus* langsdorfi). *Zoological Science* **15**, 335-337.

Neff BD (2001) Genetic paternity analysis and breeding success in bluegill sunfish (*Lepomis macrochiros*). *Journal of Heredity* **92**, 111-119.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of National Academy of Sciences USA* **70**, 3321-3323.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics* Oxford University Press, New York.

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170.

Nelson JS (1994) *Fishes of the world* John Wiley and Sons, New York.

Nguyen C, D., Tran MT, Pham A, T (2005) Carp genetic resources of Vietnam. In: *Carp genetic resources for aquaculture in Asia* (eds. David j, P., Modadugu V, G., Madan MD), pp. 72-81. World Fish Center, Penang.

Nguyen TTT, Austin CM, Meewan MM, Schultz MB, Jerry DR (2004) Phylogeography of the freshwater crayfish *Cherax destructor* Clark (Parastacidae) in inland Australia: historical fragmentation and recent range expansion. *Biological Journal of the Linnean Society* **83**, 539-550.

Nguyen VH, Ngo SV (2001) *Vietnamese freshwater fish* Agriculture Publishing House, Hanoi (in Vietnamese).

Nikolyukin NI (1956) *Fishes of Amur-river basin* Akademija Nauk SSSR, Moscow (Russian).

Nixon KC, Wheeler QD (1990) An amplification of the phylogenetic species concept. *Cladistics* **6**, 211-223.

Norris AT, Bradley DG, Cunningham EP (1999) Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture* **180**, 247-264.

Paaver T, Gross R (1991) Genetic variability of *Cyprinus carpio* L. stock reared in Estonia. *Soviet Genetic (Genetika)* **26**, 1269-1278.

Paaver T, Tammert M (1993) Genetic variability of bream and wild carp of lake Yaskaham. *Limnologica* **23**, 219-283.

Page MDR, Holmes CE (1998) *Molecular Evolution: A Phylogenetic Approach* Blackwell Science Ltd., Great Britain.

Page RDM, Charleston MA (1997) From gene to organismal phylogeny: reconciled trees and the gene tree species tree problem. *Molecular Phylogenetics and Evolution* **7**, 231-240.

Pamilo P, Nei M (1998) Relationship between gene tree and species trees. *Molecular Biology and Evolution* **5**, 568-583.

Parker PG, Snow AA, Schug MD, Booton GC, Fuerst PA (1998) What molecules can tell us about populations: choosing and using a molecular marker. *Ecology* **79**, 361-382.

Peakall R, Smouse P (2001) Genetic Analysis in Excel (Gen ALEx) version 5. Australian National University.

Peng ZG, He SP, Zhang YG (2004) Phylogenetic relationships of glyptosternoid fishes (Siluriformes : Sisoridae) inferred from mitochondrial cytochrome b gene sequences. *Molecular Phylogenetics and Evolution* **31**, 979-987.

Perdices A, Cunha C, Coelho MM (2004) Phylogenetic structure of *Zacco platypus* (Teleostei, Cyprinidae) populations on the upper and middle Chang-Jiang (= Yangtze) drainage inferred from cytochrome b sequences. *Molecular Phylogenetics and Evolution* **31**, 192-203.

Piry S, Alapetite A, Cornuet MJ*, et al.* (2004) GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536-539.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Primmer CR, Aho T, Piironen J*, et al.* (1999) Microsatellite analysis of hatchery stocks and natural populations of Arctic charr, *Salvelinas alpinus*, from the Nordic region: implications for conservation. *Hereditas* **130**, 277-289.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology and Evolution* **8**, 285-288.

Rainboth WJ (1996) *Fish of the Cambodian Mekong* Food and Agriculture Organization of the United Nation, Rome.

Rambaut A, Grassly NC (1997) Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic frees. *Computer Applications in the Biosciences* **13**, 235-238.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetic software for exact tests and ecumenicism. *Heredity* **86**, 248-249.

Reed DL, Carpenter KE, deGravelle MJ (2002) Molecular systematics of the Jacks (Perciformes: Carangidae) based on mitochondrial Cytochrome b sequences using parsimony, likelihood, and Baysian approaches. *Molecular Phylogenetics and Evolution* **23**, 513-524.

Reeder TW (2003) A phylogeny of the Australian *Sphenomorphus* group (Scincidae: Squamata) and the phylogenetic placement of the crocodile skinks (*Tribolonotus*): Bayesian approaches to assessing congruence and obtaining confidence in maximum likelihood inferred relationships. *Molecular Phylogenetics and Evolution* **27**, 384-397.

Rehbein H (2005) Identification of the fish species of raw or cold-smoked salmon and salmon caviar by single-strand conformation polymorphism (SSCP) analysis. *European Food Research and Technology* **220**, 625-632.

Rehbein H, Mackie IM, Pryde S*, et al.* (1999) Fish species identification in canned tuna by PCR-SSCP: validation by a collaborative study and investigation of intra-species variability of the DNA-patterns. *Food Chemistry* **64**, 263-268.

Reid DP, Szanto A, Glebe B, Danzmann RG, Ferguson MM (2005) QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*): comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). *Heredity* **94**, 166-172.

Rice WR (1989) Analysing tables of statistical test. *Evolution* **43**, 223-225.

Richardson BJ, Baverstock PR, Adams M (1986) *Allozyme Electrophoresis: A Handbook for Animal Systematics and Population Studies* Academic Press, Australia.

Ridley M (2003) *Wahlund Effect* Available:http://www.blackwellpublishing.com/ridley/a-z/Wahlund_effect.asp.

Rodriguez MF, LaPatra S, Williams S, Famula T, May B (2004) Genetic markers associated with resistance to infectious hematopoietic necrosis in rainbow and steelhead trout (*Oncorhynchus mykiss*) backcrosses. *Aquaculture* **241**, 93-115.

Rogers JS (1972) Measures of genetic similarity and genetic distance. In: *Studies in Genetics VII*. University of Texas Publication No. 7213.

Rozas J, Sanchez-Delbarrio CJ, Messeguer X, Rozas R (2003) DnaSP, DNA Polymorphism analyses by coalescent and other methods. *Bioinfomatics* **19**, 2496-2497.

Rutten MJM, Komen H, Deerenberg RM, Siwek M, Bovenhuis H (2004) Genetic characterization of four strains of Nile tilapia (*Oreochromis niloticus* L.) using microsatellite markers. *Animal Genetics* **35**, 93-97.

Saiki RK, Gelfand DH, Stoffel S*, et al.* (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Sakamoto T, Danzmann RG, Gharbi K*, et al.* (2000) A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* **155**, 1331-1345.

Sala O, Chapin FS, Armesto JJ*, et al.* (2000) Biodiversity scenarios for the year 2100. *Science* **287**, 1770-1774.

San Mauro D, Gower DJ, Oommen OV, Wilkinson M, Zardoya R (2004) Phylogeny of caecilian amphibians (Gymnophiona) based on complete mitochondrial genomes and nuclear RAG1. *Molecular Phylogenetics and Evolution* **33**, 413-427.

Sang TK, Chang HY, Chen CT, Hui CF (1994) Population structure of the Japanese Eel, *Anguilla japonica*. *Molecular Biology and Evolution* **11**, 250-260.

Sanger F, Nicken S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitor. *Proceedings of National Academy of Sciences USA* **74**, 5463-5467.

Sarich VM, Wilson AC (1967) Immunological time scale for hominid evolution. *Science* **158**, 1200-1203.

Sato M, Kawamata K, Zaslavskaya N*, et al.* (2005) Development of microsatellite markers for Japanese Scallop (*Mizuhopecten yessoensis*) and their application to a population genetic study. *Mar Biotechnol* **7**, 713-728.

Schander C, Willassen E (2005) What can biological barcoding do for marine biology? *Marine Biology Research* **1**, 79-83.

Schneider S, Kueffer JM, Roessli D, Excoffier L (2000) ARLEQUIN, Version 2.000 : a Software for population genetic data analysis. University of Geneva, Genetics and Biometry Laboratory. http://anthro.unige.ch/arlequin. Genetics and Bio-metry Laboratory, University of Geneva, Geneva, Switzerland.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* **18**, 233-234.

Schuh RT (2000) *Biological Systematic* Cornell University Press, New York.

Shannon CE, Weaver W (1949) *The mathematical theory of communication* University of Illinois Press, Urbana.

Shaw KL (1998) Species and diversity of natural groups. In: *Enless Forms: Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 44-56. Oxford University, Oxford.

Shaw KL (2001) The genealogical view of speciation - Commentary. *Journal of Evolutionary Biology* **14**, 880-882.

Shepard RN (1974) Representation of structure in similarity data: Problems and prospects. *Psycometrika* **39**, 373-421.

Shimodaira H, Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16**, 1114-1116.

Simmons M, Mickett K, Kucuktas H*, et al.* (2006) Comparison of domestic and wild channel catfish (*Ictalurus punctatus*) populations provides no evidence for genetic impact. *Aquaculture* **252**, 133-146.

Simon C, Franke A, Martin A (1991) The polymerase chain reaction: DNA extraction and amplification. In: *Molecular techniques in Taxonomy* (eds. Hewitt GM, Johnson A, Young JPW), pp. 329-355. Springer-Verlag, Berlin.

Sites JW, Crandall KA (1997) Testing species boundaries in biodiversity studies. *Conservation Biology* **11**, 1289-1297.

Sivasundar A, Bermingham E, Orti G (2001) Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American river. *Molecular Ecology* **10**, 407-417.

Slectova V, Slechta V, Flajshans M, Gela D (2002) Protein variability in common carp (*Cyprinus carpio*) breeds in the Czech Republic. *Aquaculture* **204**, 241-242.

Sokal R, Crovello T (1970) The biological species concepts: A critical evaluation. *American Naturalist* **104**, 127-153.

Sotka EE, Hempelmann JA, Biermann CH (2005) Genetic evidence of postglacial population expansion in Puget Sound rockfish (*Sebastes emphaeus*). *Mar Biotechnol* **7**, 223-230.

Steel M, Penny D (2000) Parsimony, likelihood, and the role of models in molecular phylogenetics. *Molecular Biology and Evolution* **17**, 839-850.

Stein JD, Phillips RB, Devlin RH (2001) Identification of sex chromosomes in chinook salmon (*Oncoorhynchus tshawytscha*). *Cytogenetics & Cell Genetic* **92**, 108-110.

Steinke D, Vences M, Salzburger W, Meyer A (2005) TaxI: a software tool for DNA barcoding using distance methods. *Philosophical Transactions of the Royal Society B-Biological Sciences* **360**, 1975-1980.

Sumantadinata K, Taniguchi N (1990) Comparison of electrophoretic allele frequencies and genetic variability of common carp stocks from Indonesia and Japan. *Aquaculture* **88**, 263-271.

Sun XW, Liang LQ (2004) A genetic linkage map of common carp (*Cyprinus carpio* L.) And mapping of a locus associated with cold tolerance. *Aquaculture* **238**, 165-172.

Sunnucks P, Wilson ACC, Beheregaray LB*, et al.* (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology* **9**, 1699-1710.

Swofford DL (2000) PAUP* phylogenetic analysis using parsimony and other methods. v.4.0b4a. Sinauer Associates, Sunderland, Masachusetts.

Tanck MWT, Baars HCA, Kohlmann K, Van Der Poel JJ, Komen J (2000) Genetic characterization of wild Dutch common carp (*Cyprinus carpio* L.). *Aquaculture Research* **31**, 779-783.

Tassanakajon A, Tiptawonnukul A, Supungul P*, et al.* (1998) Isolation and characterization of microsatellite markers in the black tiger prawn *Penaeus monodon*. *Molecular Marine Biology and Biotechnology* **7**, 55-61.

Templeton A (2002) Out of Africa again and again. *Nature* **416**, 45 - 51.

Templeton AR (1983) Convergent evolution and non-parametric inferences from restriction fragment and DNA sequence data. In: *Statistical Analysis of DNA Sequence Data* (ed. B W), pp. 151-179. Marcel Dekker, New York.

Templeton AR (2001) Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology* **10**, 779-791.

Thai BT, Burridge CP, Pham TA, Austin CM (2004) Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.). *Animal Genetics* **36**, 23-28.

Thai BT, Pham AT, Austin CM (2006) Genetic diversity of common carp in Vietnam using direct sequencing and SSCP analysis of the mitochondrial DNA control region. *Aquaculture* **258**, 228-240.

Thomas BS, Robert KW (1996) *Molecular Genetic Approaches in Conservation* Oxford University Press, New York.167-168.

Thompson JD (1985) Genetic identification of trout strains. *Aquaculture* **46**, 341-351.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgens DG (1997) The clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids research* **25**, 4876-4882.

Tommasini L, Batley J, Arnold GM*, et al.* (2003) The development of multiplex simple sequence repeat (SSR) markers to complement distinctness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. *Theoretical and Applied Genetics* **106**, 1091-1101.

Tran DT (1983) *A contribution to morphological mutation of common carp (Cyprinus carpio L.) in Vietnam* PhD, Pedagogical University (in Vietnam).

Tran MT, Nguyen CT (1992) Selection of common carp (*Cyprinus carpio*, L) in Vietnam. Selection of research works (1988 - 1992). Agriculture Publisher House, Hanoi (*In Vietnamese*).

Tran MT, Nguyen CT (1993) Selection of common carp (*Cyprinus carpio* L.) in Vietnam. *Aquaculture* **111**, 301-302.

Tran TM, Tran DT (1995) Genetic resources of common carp in Vietnam. *Aquaculture* **129**, 216-216.

Truong TK, Tran TH (1993) *Taxonomy of Freshwater Fish in Cuu Long Delta* Can Tho University, Can Tho.

Turgeon J, Bernatchez L (2001) Mitochondrial DNA phylogeography of lake cisco (*Coregonus artedi*): evidence supporting extensive secondary contacts between two glacial races. *Molecular Ecology* **10**, 987-1001.

Valles-Jimenez R, Cruz P, Perez-Enriquez R (2004) Population genetic structure of Pacific white shrimp (*Litopenaeus vannamei*) from Mexico to Panama: microsatellite DNA variation. *Mar Biotechnol* **6**, 475-484.

Vandeputte M (2003) Selective breeding of quantitative traits in the common carp (*Cyprinus carpio*): review. *Aquaculture Living Resources* **16**, 399-407.

Verspoor E (1998) Molecular markers and genetic management of farmed fish. In: *Biology of Farmed Fish* (eds. Black DK, Pickering DA), pp. 355-377. Sheffield Academic Press, England.

Vicario S, Caccone A, Gauthier J (2003) Xantusiid "night" lizards: a puzzling phylogenetic problem revisited using likelihood-based Bayesian methods on mtDNA sequences. *Molecular Phylogenetics and Evolution* **26**, 243-261.

Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503-1507.

Vooren CM (1972) Ecological aspect of the introduction of fish species into natural habitats in Europe, with special reference to the Netherlands. *Fish Biology* **4**, 565-583.

Waldbieser GC, Wolters WR (1999) Application of polymorphic microsatellite loci in a channel catfish *Ictalurus punctatus* breeding program. *Journal of the World Aquaculture Society* **30**, 256-262.

Wang CH, Li SF (2004) Phylogenetic relationships of ornamental (koi) carp, Oujiang color carp and Long-fin carp revealed by mitochondrial DNA COII gene sequences and RAPD analysis. *Aquaculture* **231**, 83-91.

Wang Z, Jayasankar P, Khoo KS*, et al.* (2000) AFLP fingerprinting reveals genetic variability in common carp stock from Indonesia. *Asian Fisheries Science* **13**, 139-147.

Waples RS (1991) Genetic interactions between hatchery and wild Salmonids - Lessons from the Pacific-Northwest. *Canadian Journal of Fisheries and Aquatic Sciences* **48**, 124-133.

Ward RD, Grewe PM (1995) Appraisal of molecular genetic techniques in fisheries. In: *Molecular Genetics in Fisheries* (eds. Carvalho GR, Pitcher TJ), pp. 29-54. Chapman & Hall, Great Britain.

Watanabe S, Aoyama J, Nishida M, Tsukamoto K (2005) A molecular genetic evaluation of the taxonomy of eels of the genus Anguilla (Pisces: Anguilliformes). *Bulletin of Marine Science* **76**, 675-689.

Waters JM, Burridge CP (1999) Extreme intraspecific mitochondrial DNA sequence divergence in Galaxias maculatus (Osteichthys : Galaxiidae). One of the world's most widespread freshwater fish. *Molecular Phylogenetics and Evolution* **11**, 1-12.

Weder JKP, Rehbein H, Kaiser KP (2001) On the specificity of tuna-directed primers in PCR-SSCP analysis of fish and meat. *European Food Research and Technology* **213**, 139-144.

Weir BS (1990) *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*, 2nd edn. Sinauer Associates, Inc., Sunderland, USA.

Wheeler QD, Meier R (2000) *Species Concept and Phylogenetic Theory* Columbia University Press, New York.

Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Systematic Biology* **51**, 69-91.

William CJ, Katherine F, Michael WB (1998) *Single -Strand Confirmation Polymorphism (SSCP) Analysis* Chapman & Hall.

Wilson K, Li YT, Whan V*, et al.* (2002) Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. *Aquaculture* **204**, 297-309.

Winans GA, Paquin MM, Van Doornik DM*, et al.* (2004) Genetic stock identification of steelhead in the Columbia River Basin: An evaluation of different molecular markers. *North American Journal of Fisheries Management* **24**, 672-685.

Wohlfarth G (1984) *Common carp, in Evolution of Domesticated Animal* Longman, London.

Wright S (1969) *Evolution and genetics of populations. Volume 2: The theory of gene frequencies* University of Chicago Press, Chicago.

Wu CI (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**, 851-865.

Wu HW (1977) Cyprinid fishes of China, Vol. 2, pp. 231-598. Shanghai: Technical Press.

Xu Z, Dhar AK, Wyrzykowski J, Alcivar-Warren A (1999) Identification of abundant and informative microsatellites from shrimp (*Penaeus monodon*) genome. *Animal Genetics* **30**, 150-156.

Xu ZK, Primavera JH, de la Pena LD*, et al.* (2001) Genetic diversity of wild and cultured Black Tiger Shrimp (*Penaeus monodon*) in the Philippines using microsatellites. *Aquaculture* **199**, 13-40.

Yeh FC, Yang RC, Boyle T (1999) POPGENE, Version 1.32, the user friendly software for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, Canada.

Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH (1998) A detailed linkage map of rainbow trout produced using doubled haploids. *Genetics* **148**, 839-850.

Yu Z, Guo X (2004) Genetic analysis of selected strains of eastern oyster (*Crassostrea virginica* Gmelin) using AFLP and microsatellite markers. *Mar Biotechnol* **6**, 575-586.

Zar JH (1984) *Biostatistical Analysis* Prentice-Hall, Inc., New Jersey.

Zardoya R, Doadrio I (1998) Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. *Proceedings of the Royal Society of London Series B-Biological Sciences* **265**, 1365-1372.

Zardoya R, Doadrio I (1999) Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. *Journal of Molecular Evolution* **49**, 227-237.

Zardoya R, Economidis PS, Doadrio I (1999) Phylogenetic relationships of Greek cyprinidae: Molecular evidence for at least two origins of the Greek cyprinid fauna. *Molecular Phylogenetics and Evolution* **13**, 122-131.

Zardoya R, Meyer A (1996) Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Molecular Biology and Evolution* **13**, 933-942.

Zhong L (1989) *Pond fisheries in China* Pergamon Press International Academic Publishers.

Zhou J, Wu Q, Wang Z, Ye Y (2004a) Genetic variation analysis within and among six varieties of common carp (*Cyprinus carpio* L.) in China using microsatellite markers. *Genetika* **40**, 1144-1148.

Zhou J, Wu Q, Wang Z, Ye Y (2004b) Molecular phylogeny of three subspecies of common carp *Cyprinus carpio*, based on sequence analysis of cytochrome b and control region of mtDNA. *Journal of Zoological Systematics and Evolutionary Research* **42**, 266-269.

Zhou JF, Wu QJ, Ye YZ, Tong JG (2003) Genetic divergence between *Cyprinus carpio carpio* and *Cyprinus carpio haematopterus* as assessed by mitochondrial DNA analysis, with emphasis on origin of European domestic carp. *Genetica* **119**, 93-97.

Zhou W, Chu XL (1986) Systematic study of the genus *Cyprinus* (Pisces: Cyprinidae) in Yunnan, China. *Zoological Research* **7**, 297-310 (in Chinese).

Zhu DQ, Jamieson BGM, Hugall A, Moritz C (1994) Sequence evolution and phylogenetic signal in control-region and Cytochrome-B sequences of Rainbow fishes (Melanotaeniidae). *Molecular Biology and Evolution* **11**, 672-683.

Zuckerkandl E, Pauling L (1962) Molecular disease, evolution, and genic heterogeneity. In: *Horizons in Biochemistry* (eds. Kasha M, Pullman B), pp. 189-225. Academic Press, New York.