Progress Toward a Global Genealogy of Common Carp (*Cyprinus carpio* L.) Strains Using Mitochondrial Nucleotide Sequences Data

B.T. Thai, T. A. Pham, U. D. Thai and C.M. Austin

**Abstract**

As part of a study of genetic variation in the Vietnamese strains of the common carp (*Cyprinus carpio* L.) using direct DNA sequencing of mitochondrial control and ATPase6/8 gene regions, samples from a number of other countries were analyzed for comparison. 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 while Koi and European carp are invariant. A genealogical analysis supports a close relationship among Vietnamese, Koi, 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, especially from China, in order to generate a comprehensive global genealogy of common carp strains.

**Introduction**

The sequence divergence of mitochondrial DNA (mtDNA) has been used to delineate the historical divergence of genetic groups and haplotype frequencies have revealed recent animal movements among local populations (Moritz 1994). MtDNA sequencing is also being increasingly used to investigate the relationships, origins and diversity of domesticated animal species, including rabbits, pigs, goats, buffalos and horses (Kierstein et al. 2004; Kim et al. 1999, 2002; Long et al. 2003; Manjunath et al. 2004; Wang et al. 2000). 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 most species. A notable exception to this is the common carp (*Cyprinus carpio* L.), the world’s oldest domesticated and arguably the most important aquaculture species (Yue et al. 2004).

In China, domestication of common carp commenced over 4000 years ago and nearly three million t are produced annually (FAO 2004). A wide range of molecular marker systems have been used to investigate genetic diversity in carp, including microsatellite (Crooijmans et al. 1997; Yue et al. 2004), randomly amplified polymorphic DNA (RAPDs) (Bartfai et al. 2003), amplified fragment length polymorphism (AFLPs) (David et al. 2001), allozymes (Brody et al. 1979; Kohlmann and Kersten 1999), restriction fragment length polymorphism (RFLPs) (Jian et al. 2003; Kohlmann et al. 2003; Riho et al. 2002) and direct sequencing of mtDNA fragments (Froufe et al. 2002; Wang and Li 2004). However, there have been no comprehensive studies of common carp that utilise nucleotide sequences.

This paper presents the results of sequencing the mtDNA ATPase6/8 and control region fragments for 89 individual common carp representing 32 strains or populations from around the world. The outcome demonstrates that a genealogical approach shows significant potential for furthering our understanding of the diversity of wild and domesticated carp populations including their origins, genetic history and relationships.

**Materials and Methods**

**Sampling**

Tissue samples of common carp were obtained from populations or strains from Vietnam, China, Japan, Indonesia, India, Hungary, Czech Republic, Israel and Australia. Strains, sample codes, and collection details are listed in Table 1.

**DNA extraction and amplification of mitochondrial DNA**

Total DNA was extracted from fin clip tissue by the protocol described by Crandall et al. (1999). Two to six individual fish were sequenced from
Table 1. Sample code, collecting locality and number of individuals sequenced (N) for each strain or population.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Sample code</th>
<th>Locality</th>
<th>N</th>
<th>Collector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xingguonensis XI</td>
<td>Jaing Xi, China</td>
<td>3</td>
<td>Li Sifa</td>
<td></td>
</tr>
<tr>
<td>Wananensis WN</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wuyuanensis WU</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color CL</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dor 70 D70</td>
<td></td>
<td>4</td>
<td>K. Ilan</td>
<td></td>
</tr>
<tr>
<td>Nasice NA</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White Koi WK</td>
<td></td>
<td>2</td>
<td>U. Lavi</td>
<td></td>
</tr>
<tr>
<td>Black Koi BAK</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kohaku Koi KK</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sanke Koi SK</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Showa Koi SHK</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red Koi KO</td>
<td>Komaki-shi, Japan</td>
<td>2</td>
<td>Y. Lemura</td>
<td></td>
</tr>
<tr>
<td>Blatna CZ</td>
<td>Czech Republic</td>
<td>5</td>
<td>K. Lukas</td>
<td></td>
</tr>
<tr>
<td>Hajduboszomeny HA</td>
<td></td>
<td>5</td>
<td>L. Varadi</td>
<td></td>
</tr>
<tr>
<td>Szeged SZ</td>
<td>Szarvas, Hungary</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tata TAT</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild Amur WA</td>
<td></td>
<td>3</td>
<td>G. C. Mair</td>
<td></td>
</tr>
<tr>
<td>Hungarian P3 P3</td>
<td></td>
<td>2</td>
<td>Y.Basavaraju</td>
<td></td>
</tr>
<tr>
<td>German X Bangkok LF</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhadra River LB</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majalaya MA</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rajadanu RJ</td>
<td>Sukamandi, Indonesia</td>
<td>2</td>
<td>N. Estu</td>
<td></td>
</tr>
<tr>
<td>Widan WI</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rockland Reservoir AU</td>
<td>Victoria, Australia</td>
<td>2</td>
<td>V. Versace</td>
<td></td>
</tr>
<tr>
<td>Yen Bai YB</td>
<td>Van Chan, Yen Bai, Vietnam</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuyen Quang TQ</td>
<td>Hoang Khai, Tuyen Quang, Vietnam</td>
<td>4</td>
<td>B. T. Thai U. D. Thai</td>
<td></td>
</tr>
<tr>
<td>Bac Kan BK</td>
<td>Bach Thong, Bac Kan, Vietnam</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vietnamese white VN</td>
<td>RIA 1, Bac Ninh, Vietnam</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian scale H</td>
<td>RIA 1, Bac Ninh, Vietnam</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesian yellow I</td>
<td>RIA 1, Bac Ninh, Vietnam</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GenBanK (X61010) GB</td>
<td>Taiwan ²</td>
<td>1</td>
<td>Y.Chang H.D. Huang</td>
<td></td>
</tr>
<tr>
<td>Goldfish (out group) GF</td>
<td>Japan</td>
<td>1</td>
<td>M. Murakami</td>
<td></td>
</tr>
</tbody>
</table>

1 Unique haplotypes within strains are presented by the number 1 and 2 immediately following the sample code (e.g., BK1 and BK2).
2 Location of strain is unknown.
RIA 1: Research Institute for Aquaculture No 1, Dinh Bang, Tu Son, Bac Ninh, Vietnam.

Each population or strain. The mitochondrial ATPase6/8 gene fragment was amplified using primers L8331 and H9236 (Lovette et al. 1998) and also the control region of primers designed from common carp sequences on GenBank (AC: X61010), Carp-Pro (5’ AACTCTCACCCCTG-GCTACCAAAG 3’) and Carp-Phe (5’ CTAGGACTCTCTTAGCATCTTCAGTG 3’). PCR conditions and procedures for visualization, purification and sequencing are given in Murphy and Austin (2002).

Data analysis and phylogenetic reconstruction

Sequences of common carp from GenBank (Chang and Huang 1994) were included in the data set for comparative purposes, and Carassius...
\textit{Auratus} sequences (GenBank accession number NC 002079) were used as the out group.

The DNA sequences from the control region were aligned using the program Clustal X (Thompson et al. 1997). The DNA sequences from ATPase6/8 regions were aligned by eye. Aligned sequences were imported into PAUP* 4.0.b.10 (Swofford 2000) for phylogenetic analysis. It is important to identify the most suitable model of nucleotide evolution for valid phylogenetic inference (Felsenstein 2004). The Modeltest 3.06 program (Posada and Crandall 1998) was used to test an alternative model of evolution. The best model was then used to calculate pair-wise sequence distances for constructing a neighbor-joining (NJ) tree. Confidence levels in the resulting relationship were assessed using the bootstrap procedure with 1000 replications. Nucleotide (\(\pi\)) and haplotype (\(h\)) diversity for specific grouping of samples were calculated using DNASP 4.10 (Rozas et al. 2003).

**Results**

A total of 820 bp of the ATPase6/8 genes and 745 bp of the control region were sequenced for 87 individuals. All sequences obtained in this study have been submitted to GenBank (accession numbers: AY597942-AY597985 and AY 600150-AY600241). The mean nucleotide composition is A = 31 per cent, T = 30 per cent, C = 23 per cent and G = 16 per cent. The most suitable evolution model identified for the nucleotide data is: Hasegawa-Kishino-Yano (HKY) + invariable site (I) + gamma distribution shape parameter (G); transition/ transversion ratio = 3.8306. The HKY model allows transition and transversion mutation occurring at different rates and base frequencies of nucleotide to vary (Hasegawa et al. 1985; Page and Holmes 1998). The level of nucleotide variation at both gene regions was found to be low, although the control region exhibited more variation than the ATPase6/8. Excluding the GenBank \textit{Cyprinus carpio} (which is of unknown origin) based on nucleotide substitution, the two most divergent ATPase6/8 haplotypes differed by 12 bp and the two most divergent control region haplotypes differed by 14 bp. The average divergence (~ 1 per cent) is low, given that the control region and ATPase6/8 genes are two of the fastest evolving regions of the mitochondrial genome. The average control region divergence (1.24 per cent) was higher than that of the average ATPase6/8 region (0.68 per cent) (data not shown). Despite this relatively limited variability, the two gene regions possessed significant phylogenetic signal and were largely consistent in the patterns of variation within and between samples and geographic groupings and the relationships amongst the haplotypes.

The nucleotide and haplotype diversity (range 0 to 1) in various geographic groupings is summarized in Figure 1. Although nucleotide diversity is low overall, differences are apparent between the Asian carp (0.008) and carp of European origin (0.000). These differences were most obvious in terms of haplotype diversity (\(h\)), which exceeded 0.93 for all carp of Southeast Asian origins but was 0.00 for European carp. Koi carp showed no diversity, were identical

![Figure 1](image-url)
to the Chinese Color (CL) carp and showed minimal divergence from Vietnamese carp. The European carp, while clearly distinguishable from all other carp, are closely related to the clade containing the Koi and Vietnamese carp. Significant haplotype diversity is apparent among the Indonesian and Chinese samples. The most divergent haplotype was represented by the Chinese Wuyuanensis strain (WU) that differs from between 6 to 20 bp from all other common carp haplotypes. In addition, the sample of Amur river carp (WA1 and WA2) is interesting as it contains two haplotypes, one close to Chinese Color carp and therefore also Vietnamese carp, and the other that is identical to European carp. The relationship of the carp strains is shown in Figure 2.

**Discussion**

**Mitochondrial variation**

This study has demonstrated that direct sequencing of variable mtDNA fragments can 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 in what are generally considered to be the most rapidly evolving vertebrate mtDNA fragments, the ATPase6/8 genes and the control region (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 localized loss of genetic variation (Balon 1995).

**Genetic relationships**

The origin of a number of carp strains is uncertain (Balon 1995). Koi carp for example, are thought to have been developed directly from common carp stocks in Japan, or Chinese Color carp or German pond carp (Balon 1995; Wang and Li 2004). The results of this study suggest that Koi carp have originated from Chinese Color carp which have a long history of domestication that can be traced back over 1200 years (Wang and Li 2004). This is consistent with the results of Wang and Li (2004) based on COII sequences, but contrary to those of Froufe et al. (2002) who sequenced 540 bp of the control region and found that their single sample of Japanese Koi (N=4) had a haplotype identical to European carp. This anomaly could be due to mixing of stocks with Koi carp acquiring the European mtDNA through cross breeding. 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 and Chinese Color carp having passed through a substantial bottleneck during the domestication process.

Carp from Europe or originating from Europe had the same haplotype, which was divergent from all other haplotypes detected. A remarkable feature of the European carp samples is the complete absence of any variation, which suggests a history of founder effects and small population size associated with translocation and domestication, like the Koi carp. In contrast, Chinese carp showed considerable haplotype diversity and divergence. While no differences were detected between the Xingguonensis (XI) and Wananensis (WN) strains, their haplotypes are highly differentiated from Chinese Color carp (CL) and Wuyuanensis (WU) carp with the latter showing the most distinctive common carp haplotype found so far. This suggests that carp have a longer evolutionary history in China, compared to other countries or geographic regions (i.e., Vietnam and Europe).

The genetic makeup of Amur river carp represents an anomaly in that, while they occur over 5000 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 another 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 attributed 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 reverse possibility, that European carp had been introduced to China, has also been proposed (Balon 1995) and is also consistent with the data.

**Taxonomic implications**

The taxonomy of common carp is highly contentious and confusing with over 30 synonyms of *C. carpio* listed on Fish-Base (www.fishbase.org/search.cfm) and a proliferation of named strains. Considering only relatively recent literature, Kirpichnikov (1967) recognised four subspecies: *Cyprinus carpio carpio* (Europe); *C. c. aralensis* (Central Asia); *C. c. haematopterus* (East Asia); and *C. c. viridiviolaceus* (Southeast Asia). In contrast, Balon (1995) recognized only two subspecies: *Cyprinus carpio carpio* (Europe) and
Cyprinus carpio haematopterus (East Asia). Subsequently, Kirpichnikov (1999) questioned the validity of C. c. viridiviolaceus whereas Li et al. (2001) recognised four morphologically distinctive red common carp strains from China which he referred to as C. c. xingguonensis, C. c. wannanensis, C. c. wuyanensis, C. c. color. In one of the most recent taxonomic treatments of common carp Kottelat (2001) considers the common carp cultured 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. The results of genetic studies using allozyme, microsatellite, and RFLP analysis of the mtDNA ND-3/4 and ND-5/6 regions by Kohlman et al. (2003) were interpreted to provide 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 form two distinct clades.

MtDNA sequence data, often used for clarifying species boundaries (Avise 2000), provides little support for any of the previous positions. While the results of this study are largely consistent with those of Kohlman et al. (2003) with 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 are highly divergent fails to support an independent origin for western and eastern common carp lineages and their recognition as distinct subspecies (or species). Indeed, this finding, combined with the deepest branches of the phylogenetic tree represented by samples of Chinese origin, suggests that common carp most probably originated in China.

While genetic distance is not always a reliable guide to taxonomic recognition, the levels of divergence observed for the control and ATPase6/8 regions (~1 per cent) are much less in carp compared to that usually observed among fish species and sub-species. For example, Hurt et al. (2001) observed between 15-18 per cent divergences in control region sequences among closely related species within the genus Acanthopagrus. It is essential that some stability to the taxonomy of the common carp is achieved and
we recommended that this study be extended to include more samples of carp, especially from China, with parallel studies of morphology, genetic variation at rapidly evolving nuclear loci (especially intron regions) and reproductive relationships.

The Development of a Global Genealogy of Common Carp Strains

The findings of this study indicate that the establishment of genealogical relations 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 (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 1995). It can also provide information on the historical origins of strains, whether they represent a composite of different strains, and if they have lost variation during the process of domestication.

Acknowledgement

The authors would like to thank the numerous fish farmers and researchers for providing common carp samples, particularly F.S. Li (China), G. C. Mair (AIT-Thailand), Y. Basavaraju (India), L. Varadi, Z. Jeney (Hungary), U. Lavi, I. Karpplus (Israel), E. Nugroho (Indonesia), K. Lukas (Czech Republic), Y. Lemura (Japan) and V. Versace (Australia). We also thank Dr. M.V. Gupta for his encouragement and support for this project. The authors are very grateful to the referees and editor for their helpful and valuable suggestions. This research was supported by funding from the Australian Agency for International Development (AusAID) through its Vietnam Capacity Building for Agriculture and Rural Development (CARD) pilot project program.

References


B. T. Thai, T. A. Pham and U. D. Thai are from Research Institute for Aquaculture No. I, Dinh Bang, Tu Son, Bac Ninh, Vietnam.

C. M. Austin is from the School of Science and Primary Industry, Charles Darwin University, Darwin NT 0909, Australia.

Corresponding author: C. M. Austin. Email: chris.austin@cdu.edu.au