GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

# Complete mitochondrial genome of a natural triploid crucian carp mutant, *Carassius auratus* var. *pingxiangnensis*, and phylogenetic analysis of different ploidies in crucian carp

**J.Q. Sheng[1], J.H. Wang[1], S.H. He[1], L.G. Zeng[2], K. Peng[1] and Y.J. Hong[1]**

[1]Aquaculture Science Department, College of Life Sciences, Nanchang University, Nanchang, China
[2]Nanchang Academy of Agricultural Science, Nanchang, China

Corresponding author: Y.J. Hong
E-mail: yijianghong@126.com

Genet. Mol. Res. 13 (3): 5849-5864 (2014)
Received July 23, 2013
Accepted November 18, 2013
Published August 1, 2014
DOI http://dx.doi.org/10.4238/2014.August.1.3

**ABSTRACT.** *Carassius auratus* var. *pingxiangnensis* is a natural triploid crucian carp mutant. In order to understand its placement and genetic background at the gene level, the characteristics of mitochondrial DNA sequences and phylogenetic relationship were examined. The results showed that the mitochondrial DNA is a circular double-stranded DNA molecule that is 16,576 bp in length with 13 protein-coding genes, 22 transfer RNA genes, 2 ribosomal RNA genes, and a non-coding control region. Mitochondrial genes overlapped by a total of 40 bp in 11 different locations from 1 to 14 bp. The base composition of the *C. auratus* mitogenome was estimated to be 29.70% A, 26.74% C, 15.35% G, and 28.21% T. The central conserved blocks and the conserved blocks were compared and were similar among *C. auratus* var. *pingxiangnensis* and six other cyprinids with different ploidies. The origin of light strand replication was similar to that of other vertebrates; it was 33 bp, but the characteristic sequence motif 5'-GCCGG-3'

 

at the base of the stem within $tRNA^{Cys}$ was mutated to 5'-GGCGG-3'. Our phylogenetic analysis based on whole mitogenome sequences indicated that *C. auratus* var. *pingxiangnensis* was clustered with *C. auratus* and then sister-grouped with *Carassius gibelio*. The systemic developmental tree of crucian carp with different chromosome ploidies showed that diploid *C. auratus auratus* was clustered with triploid *C. auratus auratus*, sister-grouped with tetraploid *C. auratus auratus*, and clustered with other diploid, triploid, and tetraploid *C. auratus*.

**Key words:** Complete mitochondrial DNA sequence; Ploidy; *Carassius auratus* var. *pingxiangnensis*; Molecular characteristics; Phylogenetic analysis

## INTRODUCTION

Vertebrate mitochondrial DNA (mtDNA) is generally a small double-stranded circular genome ranging in size from approximately 16 to 18 kb with no introns. A typical mitogenome contains a coding region with 37 genes: 2 ribosomal RNA genes (rRNAs), 22 transfer RNA genes (tRNAs), and 13 protein-coding genes, as well as non-coding regions (Anderson et al., 1981; Meng et al., 2012; Liu et al., 2013). The gene order is highly conserved in vertebrates with a few exceptions, such as amphibian species (Zhang et al., 2005) and certain fish (Miya and Nishida, 1999). Complete mitochondrial genomes from numerous vertebrate species have found widespread use in various studies on molecular ecology, population genetics, and evolutionary biology. An accumulating body of evidence demonstrates the usefulness of mitogenomic data in resolving several longstanding controversial phylogenetic relationships (Saitoh et al., 2006; Kawahara et al., 2008; Wang et al., 2008).

Pingxiang red-transparent crucian carp, which was named *Carassius auratus* var. *pingxiangnensis* by the National Fisheries thoroughbred Examining Committee of China in 2008, was found in Pingxiang district of Jiangxi Province in China, and the ratio of males to females was about 1:10 in natural water areas. Hong et al. (2005) concluded that the fish was a natural triploid crucian carp mutant based on DNA content measurement and chromosome analysis, and it had dual reproduction modes of gynogenesis and sexual reproduction. Such a vertebrate with an exceptional reproductive system is unusual in nature and is very attractive as a laboratory animal for genetic studies of genetic structure and phylogenetic relationships with respect to reproduction identification, management, and conservation. In order to know its placement and more genetic background at the gene level, we examined the characteristics of mtDNA sequences and phylogenetic relationship.

## MATERIAL AND METHODS

### mtDNA purification by long polymerase chain reaction (PCR)

All wild natural triploid crucian carp mutants were sampled from Pingxiang district of Jiangxi Province. The total genomic DNA was extracted from the tail fin samples using a DNA extraction kit (Takara Biotechnology Co., Ltd. Japan). The entire mitochondrial genomes of

 

the fishes were amplified using a long PCR (Mastercycler Personal, Eppendorf China Ltd.) technique (Miya and Nishida, 1999). Two sets of fish-versatile primer pairs (S-LA-16S-L+H12830-CYB and L12830-Leu+S-LA-16S-H, Table 1) were used to amplify almost the entire mitochondrial genome in two long PCRs (Miya and Nishida, 2000). The reactions were carried out with 30 cycles of a 25-μL reaction volume containing 1X LA PCR buffer (Takara), 1.0 mM dNTPs, 0.5 mM of each primer, 1 U LA Taq polymerase (Takara), 100 ng template DNA, and $ddH_2O$. The thermal cycle profile was "shuttle PCR": pre-denaturation at 94°C for 2 min, denaturation at 98°C for 10 s, annealing and extension combined at the same temperature (68°C) for 16 min, and 72°C for 5 min to denature the Taq polymerase. The long PCR products were electrophoresed on a 0.8% agarose gel, diluted 1:10, and used as templates for PCR except for a region between the two long PCR primers (S-LA-16S-L+H12830-CYB and L12830-Leu+S-LA-16S-H, Table 1), where total genomic DNA was used instead of long PCR products.

We used 26 different primer pairs that amplify contiguous, overlapping segments to obtain the entire mitochondrial genome of the fish. Some of these primers were versatile, based on the complete mitochondrial genome of six bony fish species, following Miya and Nishida (2000). The others were designed with reference to the complete mitochondrial genome sequences from *C. auratus* (GenBank Accession No. GQ903705), and all of the primers were listed in Table 1. PCRs were carried out in 50-μL reaction volumes containing 1X PCR buffer (Takara), 0.5 mM dNTPs, 0.5 mM of each primer, 2 U Ex Taq polymerase (Takara), 2.0 μL diluted long PCR products, and $ddH_2O$. The thermal cycle profile was pre-denaturation at 94°C for 5 min, denaturation at 94°C for 30 s, annealing at 52°C for 45 s, extension at 72°C for 1 min, and 72°C for 10 min to denature the Taq polymerase. The PCR products were electrophoresed on a 1.0% agarose gel.

Double-stranded PCR products that were purified by filtration through Millipore plates were subsequently used for direct cycle sequencing with dye-labeled terminators (Applied Biosystems Inc.). The primers used were the same as those for PCR. All sequencing reactions were performed according to manufacturer instructions. Labeled fragments were analyzed on a model 373S/377 DNA sequencer (Applied Biosystems Inc.). The complete mitogenome sequence of *C. auratus* var. *pingxiangnensis* has been deposited in GenBank under accession number HQ875340.

## Sequence assembly and analysis

Sequence alignment was performed initially using the DNASTAR software (DNASTAR, Madison, WI, USA) with default parameters and further adjusted manually. The locations of the 13 protein-coding genes were determined by comparisons of nucleotide and amino acid sequences of teleost mitochondrial genomes and SEQUIN10.30. The codon usage of the 13 protein-coding genes was analyzed with the MEGA 4.0 (Kumar et al., 2008) and Gene to Codon Usage program (http://www.entelechon.com). Most tRNA genes were identified by their proposed cloverleaf secondary structures using the web-based tRNAscan-SE v.1.21 software (http://lowelab.ucsc.edu/tRNAscan-SE/), which gave different cove score cut-off values to sequences (Lowe and Eddy, 1997). The remaining tRNA genes were determined by inspecting sequences for tRNA-like secondary structures and anticodons. The two rRNA genes were identified by sequence homology and proposed secondary structure (Gutell, 1993). The con-

trol region and putative origin of light strand replication ($O_L$) were identified by comparison with the homologous sequences of other Cypriniformes.

**Table 1.** Primers used for PCR amplification and analysis in the present study.

| No. | Primer name | Sequence |
|---|---|---|
| 1 | S-LA-16S-H | TGCACCATTRGGATGTCCTGATCCAACATC |
| | L12830-Leu | GGTCTTAGGAACCAAAAACTCTTGGTGCAA |
| 2 | L2897-16S | CTCGGCAAACATAAGCCTCGCCTGTTTACCAAAAAC |
| | H1435-CYB | GGCATAGTGGGGTATCTAATCCCAGTTTGT |
| 3 | L1031-12S | TACACATGCAAGTCTCCGCA |
| | H2405-16S | CTTAGGCAACCAGCTATCAC |
| 4 | L2325-16S | CGTCTCTGTGGCAAAAGAGTGG |
| | H3536-16S | TCCGGTCTGAACTCAGATCACGTA |
| 5 | L3387-16S | GGGATAACAGCGCAATC |
| | H4231-ND1 | CCTAAGCTGACTTCATATGAG |
| 6 | L4152-ND1 | TGAGCMTCWAATTCMAAATT |
| | H4925-MET | TTTAACCGWCATGTTCGGGGTATG |
| 7 | L4553-ND1 | TACAACAATCAACCTCATAA |
| | H5379-ND2 | AAKGGKGCKAGTTTTTGTCA |
| 8 | L5105-ND2 | CACCACCCWCGCGCAGTAGA |
| | H5847-ND2 | CGKAGRTAGAAGTAKAGGCT |
| 9 | L5297-ND2 | CACTTCTGAKTGCCCGARGT |
| | H6188-ASN | AACTAAGAGTTTGWAGGATCGAGGCC |
| 10 | L5593-ND2 | CATGACCTCCGCRGCATTCCCAAC |
| | H7074-COI-I | CCCCCGCCWGCGGGATCAAA |
| 11 | L6674-COI-I | GCATTCCCWCGAATAAATAA |
| | H7368-COI-I | AGTCAGCTAAAKACTTTTAC |
| 12 | L7205-COI-I | TATATAGGAATRGTMTGAGC |
| | H7999-SER | ATGTGGYTGGCTTGAAA |
| 13 | L7820-COI-I | GCCTTCGCKGCTAAACGAGAAGT |
| | H8641-COI-II | GTCGGYCTGGGACTGCGTCTA |
| 14 | L8365-COI-II | ATGGACGAAATTAATGATCC |
| | H9049-ATP 6 | AAGCTTAKTATCATGGTCAGT |
| 15 | L8823-Lys | AGCGTTGGCCTTTTAAGC |
| | H10569-Gly | CTTTCCTTGGGKTTTAACCAAG |
| 16 | L9708-ATP6 | AACGTTTAATGGCCCACCAAGC |
| | H12859-Leu | TTGCACCAAGAGTTTTTGGTTCCTAAGACC |
| 17 | L11577-ND4 | GCTCCTGMCCACACTACAAGCC |
| | H13378-ND5 | TGTGTTGGCATCTGKCCGTCC |
| 18 | L12848-Leu | CTCTTGGTGCAAMTCCAAGT |
| | H14649-ND5 | AGGTAKGTTTTAATTAKKCC |
| 19 | L14087-ND5 | TCCCACCTAAACGCCTGAGCCCT |
| | H1435-12S | AGCATAGTGGGGTATCTAATCCCAGTTTGT |
| 20 | L892-CR | AATCAGCCGCCACTCAACTAGAC |
| | H2212-16S | CGGCGATAACGAATGGTAGAAG |
| 21 | L1754-12S | GAAACTTTGGCTCTCTACTAGGA |
| | H2735-16S | AACAAGGAACAGTGCGAAAT |
| 22 | H1844-12S | CCGCATTCTCATCCGTTAC |
| | L2734-16S | AGAACAAGGAACAGTGCGAAATA |
| 23 | H2474-16S | AGTGGTATTTCCTATTTGCCTAC |
| | L3271-16S | GTTCATATTGTGCAACCC |
| 24 | H2724-16S | ATCCGTCCTATATTTCGCACTGT |
| | L3722-Leu | GCCAATTCGTACCTCAACTCGTT |
| 25 | H3290-16S | ATTACTGGCATCTGGTTCCTATT |
| | L4734-ND1 | GGGTTCAAAGGGACACTCTATT |
| 26 | H376-CR | TCAGCCGTAAACTTAGACATC |
| | L1523-12S | GCTTTAACGCTTTCTATCTAGGT |

L and H denote light and heavy strands, respectively. Positions with mixed bases are labeled with their IUBcodes: R = A or G; Y = C or T; K = G or T; M = A or C; S = G or C; W = A or T. Primer pairs No. 1 and 2 were used for LA-PCR; primer pairs No. 3-26 were used to amplify contiguous, overlapping segments of the entire mitochondrial genome.

## Phylogenetic analysis

The entire mtDNA sequences of *C. auratus* var. *pingxiangnensis* were used for comparative analyses and phylogeny. In addition, possible close outgroups in Cypriniformes were chosen to root phylogenetic trees (Fu et al., 2009). The mtDNA sequences of nine species of fishes were downloaded from GenBank, and *Myxocyprinus asiaticus* was used as an outgroup for phylogenetic analysis. Phylogenetic analyses were performed using maximum parsimony (MP) and neighbor joining (NJ) in MEGA 4.0 (Tamura et al., 2007).

# RESULTS

## Genome content and organization

The total length of the mitochondrial genome was determined to be 16,576 bp with a standard set of 22 tRNAs, 2 rRNAs, and 13 typical vertebrate protein-coding genes, as well as a putative non-coding control region (Figure 1, Table 2). Most of the genes were encoded on the heavy strand (H-strand), only ND6 and eight tRNA genes ($tRNA^{Phe}$, $tRNA^{Glu}$, $tRNA^{Ser}$, $tRNA^{Tyr}$, $tRNA^{Cys}$, $tRNA^{Asn}$, $tRNA^{Ala}$, and $tRNA^{Gln}$) were encoded on the light strand (L-strand). Salient features of the mitochondrial genes were listed in Table 2. Mitochondrial genes overlapped by a total of 40 bp in 11 different locations from 1 to 14 bp. Several overlaps were detected between protein-coding genes as shown in ATPase8-ATPase6, ATPase6-COI-III, ND4L-ND4, and ND5-ND6; sequences for $tRNA^{Ile}$ and $tRNA^{Gln}$, ND2 and $tRNA^{Trp}$, $tRNA^{Cys}$ and $tRNA^{Tyr}$, COI-II and $tRNA^{Lys}$, COI-III and $tRNA^{Gly}$, ND3 and $tRNA^{Arg}$, and $tRNA^{Thr}$ and $tRNA^{Pro}$ overlapped as well. In addition, 13 intergenic spacers were present in the *C. auratus* var. *pingxiangnensis* mitogenome and involved a total of 36 bp.

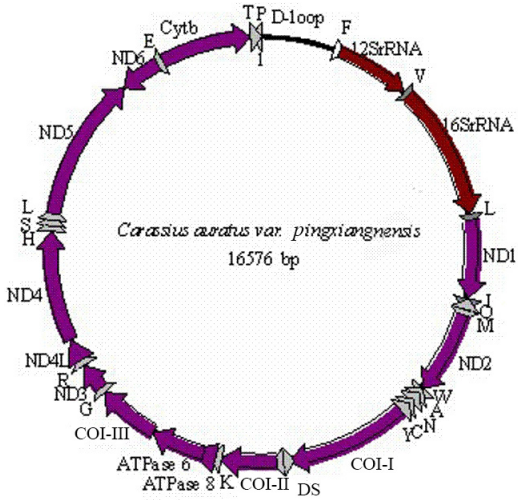


**Figure 1.** Gene organizations for the mitochondrial genome of *Carassius auratus* var. *pingxiangnensis*. All protein-coding genes are encoded by the heavy strand with the exception of ND6, which is coded by the light strand. Transfer RNA genes are designated by single-letter amino acid codes; ND1-6 and 4L = NADH dehydrogenase subunits 1-6 and 4L; COI-III = cytochrome c oxidase subunits I-III; ATP 6 and 8 = ATPase subunits 6 and 8; Cytb = cytochrome b; CR = control region.

**Table 2.** Characteristics of the mitochondrial genome of *Carassius auratus* var. *pingxiangnensis*.

| Gene | Position | | Size | | Codon | | Intergenic | Strand |
|---|---|---|---|---|---|---|---|---|
| | From | To | Nucleotide (bp) | Amino Acid | Start | Stop[a] | Nucleotide[b] | |
| $tRNA^{Phe}$ | 920 | 988 | 69 | | - | - | 0 | H |
| 12SrRNA | 989 | 1942 | 954 | | | | 0 | H |
| $tRNA^{Val}$ | 1943 | 2014 | 72 | | | | 0 | H |
| 16SrRNA | 2015 | 3695 | 1681 | | | | +1 | H |
| $tRNA^{Leu(UUR)}$ | 3697 | 3772 | 76 | | | | +1 | H |
| ND1 | 3774 | 4748 | 975 | 324 | ATG | TAA | +4 | H |
| $tRNA^{Ile}$ | 4753 | 4824 | 72 | | | | -2 | H |
| $tRNA^{Gln}$ | 4823 | 4893 | 71 | | | | +1 | L |
| $tRNA^{Met}$ | 4895 | 4963 | 69 | | | | 0 | H |
| ND2 | 4964 | 6010 | 1047 | 348 | ATG | TAG | -2 | H |
| $tRNA^{Trp}$ | 6009 | 6079 | 71 | | | | +2 | H |
| $tRNA^{Ala}$ | 6082 | 6150 | 69 | | | | +1 | L |
| $tRNA^{Asn}$ | 6152 | 6224 | 73 | | | | 0 | L |
| $O_L$ | 6225 | 6257 | 33 | | | | 0 | H |
| $tRNA^{Cys}$ | 6258 | 6326 | 71 | | | | -1 | L |
| $tRNA^{Tyr}$ | 6326 | 6396 | 71 | | | | +1 | L |
| COI-I | 6398 | 7948 | 1551 | 516 | GTG | TAA | 0 | H |
| $tRNA^{Ser(UCN)}$ | 7949 | 8019 | 71 | | | | +3 | L |
| $tRNA^{Asp}$ | 8023 | 8094 | 72 | | | | +12 | H |
| COI-II | 8107 | 8811 | 705 | 234 | ATG | TAA | -14 | H |
| $tRNA^{Lys}$ | 8798 | 8873 | 76 | | | | +1 | H |
| ATPase 8 | 8875 | 9039 | 165 | 54 | ATG | TAG | -7 | H |
| ATPase 6 | 9033 | 9716 | 684 | 227 | ATG | TAA | -1 | H |
| COI-III | 9716 | 10501 | 786 | 261 | ATG | TAA | -1 | H |
| $tRNA^{Gly}$ | 10501 | 10572 | 72 | | | | 0 | H |
| ND3 | 10573 | 10923 | 351 | 116 | ATG | TAG | -2 | H |
| $tRNA^{Arg}$ | 10922 | 10991 | 70 | | | | 0 | H |
| ND4L | 10992 | 11288 | 297 | 98 | ATG | TAA | -5 | H |
| ND4 | 11282 | 12662 | 1381 | 460 | ATG | T | 0 | H |
| $tRNA^{His}$ | 12663 | 12731 | 69 | | | | 0 | H |
| $tRNA^{Ser(AGY)}$ | 12732 | 12800 | 69 | | | | +1 | H |
| $tRNA^{Leu(CUN)}$ | 12802 | 12874 | 73 | | | | +3 | H |
| ND5 | 12878 | 14701 | 1824 | 607 | ATG | TAA | -4 | H |
| ND6 | 14698 | 15219 | 522 | 173 | ATG | TAG | 0 | L |
| $tRNA^{Glu}$ | 15220 | 15288 | 69 | | | | +5 | L |
| Cytb | 15294 | 16434 | 1141 | 380 | ATG | T | 0 | H |
| $tRNA^{Thr}$ | 16435 | 16506 | 72 | | | | -1 | H |
| $tRNA^{Pro}$ | 16506 | 16575 | 70 | | | | 0 | L |
| Control region | 1 | 919 | 919 | | | | 0 | H |

[a]TA and T represent incomplete stop codons. [b]Numbers correspond to the nucleotides separating adjacent genes. Negative numbers indicate overlapping nucleotides.

The overall base composition of *C. auratus* mitogenome was estimated to be 29.70% A, 26.74% C, 15.35% G, and 28.21% T, indicating an obvious anti-guanine bias that is commonly observed in fishes (Wang et al., 2008). The sequence analysis also showed a high A+T content with mean overall value of 57.91%, which was also reflected in the codon usage with A and T nucleotides being preferred over C and G nucleotides at the third codon position (Table 3).

## Protein-coding genes and usage

All of the 13 protein-coding genes that are found in other vertebrates were also present in the *C. auratus* var. *pingxiangnensis* mitogenome, including 3 subunits of cytochrome c oxidase (COI-III), 7 subunits of the NADH ubiquinone oxidoreductase complex (ND1-6, ND4L),

1 subunit of the ubiquinol cytochrome b oxidoreductase complex (Cytb), and 2 subunits of ATP synthases (ATP6 and ATP8). The total length of those genes was 11,429 bp, accounting for 68.95% of the whole mitogenome. In accordance with other bony fishes, overlapping protein-coding genes were found in the *C. auratus* var. *pingxiangnensis* mitogenome. There were three cases of reading-frame overlaps on the same strand (Table 2): the reading frames of ATP8-ATP6 and ND4L-ND4 each overlapped by seven nucleotides, and that of ATP6-COI-III overlapped by one nucleotide. The ND5-ND6 pair of genes, which were encoded on different strands, overlapped by four nucleotides.

**Table 3.** Base composition (%) at first, second and third codon positions of *Carassius auratus* var. *pingxiangnensis*.

| Codon position | A | | C | | G | | T | |
|---|---|---|---|---|---|---|---|---|
| | Number | Base composition (%) | Number | Base composition (%) | Number | Base composition (%) | Number | Base composition (%) |
| 1st | 1015 | 26.61 | 981 | 25.72 | 989 | 25.93 | 829 | 21.74 |
| 2nd | 709 | 18.60 | 1036 | 27.18 | 521 | 13.67 | 1545 | 40.54 |
| 3rd | 1674 | 43.93 | 1039 | 27.26 | 244 | 6.40 | 854 | 22.41 |
| All | 3394 | 29.70 | 3056 | 26.74 | 1754 | 15.35 | 3225 | 28.21 |
| tRNA | 482 | 38.8 | 312 | 29.70 | 381 | 7.50 | 393 | 24.00 |
| rRNA | 914 | 34.7 | 632 | 24.00 | 556 | 21.10 | 532 | 20.20 |

As shown in Table 4, the 12 genes encoded on the H-strand exhibited a marked similarity in nucleotide composition with an antiguanine bias, whereas ND6, encoded on the L-strand, exhibited an opposite trend (31.99% G). Among protein-coding genes, the A+T content varied from 54.02% (ND6) to 63.03% (ATPase8). The nucleotide composition reflects a vertebrate bias against G on the L-strand. In the protein-coding genes, an anti-G bias was observed in the third codon position (6.4%). Pyrimidines (C+T) were over-represented in the second codon position (67.7%), which has been noted for other vertebrate mitochondrial genomes because of the hydrophobic characteristic of the proteins.

**Table 4.** Base composition for 13 protein-coding gene of *Carassius auratus* var. *pingxiangnensis* mitogenome.

| Gene/region | Base composition (%) | | | | | Total number |
|---|---|---|---|---|---|---|
| | T | C | A | G | A+T | |
| ND1 | 26.15 | 28.92 | 29.64 | 15.28 | 55.79 | 975 |
| ND2 | 22.06 | 31.52 | 33.43 | 12.99 | 55.49 | 1047 |
| COI-I | 29.46 | 25.73 | 27.60 | 17.21 | 57.06 | 1551 |
| COI-II | 26.95 | 24.82 | 32.20 | 16.03 | 59.15 | 705 |
| ATPase 8 | 27.88 | 24.85 | 35.15 | 12.12 | 63.03 | 165 |
| ATPase6 | 29.39 | 25.73 | 32.16 | 12.72 | 61.55 | 684 |
| COI-III | 27.35 | 28.24 | 28.63 | 15.78 | 55.98 | 786 |
| ND3 | 31.62 | 26.21 | 27.35 | 14.81 | 58.97 | 351 |
| ND4L | 29.63 | 29.63 | 24.58 | 16.16 | 54.21 | 297 |
| ND4 | 27.83 | 26.88 | 31.74 | 13.55 | 59.57 | 1380 |
| ND5 | 27.69 | 26.64 | 32.79 | 12.88 | 60.47 | 1824 |
| Cytb | 28.83 | 28.13 | 28.22 | 14.81 | 57.02 | 1141 |
| ND6 | 40.61 | 13.98 | 13.41 | 31.99 | 54.02 | 522 |

ND6, which were encoded on the light strand, were converted to complementary strand sequences.

Start and stop codons were determined based on alignments with the corresponding genes and proteins of other fishes. As shown in Table 2, most protein-coding genes began with

an ATG codon; only the COI gene began with a GTG codon rather than an ATG codon. Three types of stop codons were used by the coding genes: TAA for ND1, COI-I, COI-II, ATPase6, COI-III, ND4L, and ND5; TAG for ND2, ATPase8, ND3, and ND6; and ND4 and Cytb did not have proper stop codons, but did have a terminal T. The use of an incomplete stop codon as a common mechanism for stopping protein translations was also observed in other fish species (e.g., Kartavtsev et al., 2007).

Codon usage patterns in crucian carp are shown in Table 5. Three other vertebrates (carp, frog, and human) are shown for comparison (Roe et al., 1985; Broughton et al., 2001). For amino acids with a fourfold degenerate third position, codons ending in A were always the most frequent in *C. auratus* var. *pingxiangnensis*, followed in frequency by codons ending in T or C. Among twofold degenerate codons, C appears to be used somewhat more than T. Consistent with the overall bias against G, G is the least common third position nucleotide in all categories except for glycine codons (where G is similar in frequency to T but still much less frequent than A and C). These patterns are generally similar across vertebrate groups, although the frog shows a tendency to use T more frequently than C (Roe et al., 1985).

## Mitochondrial rRNA and tRNA genes

The 12S and 16S rRNA genes were 954 and 1681 nucleotides long, respectively (Table 2), and the length between two rRNA subunits was found to vary from species to species (Kartavtsev et al., 2007). As in other vertebrates, they were located between genes of the $tRNA^{Phe}$ and $tRNA^{Leu(UUR)}$, being separated by the $tRNA^{Val}$ gene (Figure 1). The tRNA genes were slightly A+T-rich (55.9%), whereas rRNA genes had a high adenine content (34.7%) like they do in other bony fishes (Table 3).

The mitochondrial genome contained 22 typical tRNA genes interspersed between the rRNA and protein-coding genes. The tRNA genes range in size from 69 to 76 nucleotides (Table 2). As seen in other vertebrate tRNAs, numerous noncomplementary and U-G base pairs were found in tRNA stem regions. In $tRNA^{His}$, $tRNA^{Ser}$, and $tRNA^{Arg}$, A and C were not matched. In $tRNA^{Phe}$, U and U and A and A were not matched, and in $tRNA^{Thr}$, C and C were not matched. All other postulated cloverleaf structures contained 7 bp in the aminoacyl stem, 5 bp in the TΨC stem, 7 bp in the anticodon stems, and 4 bp in the DHU stem [3 bp in $tRNA^{Ser(AGY)}$].

## Non-coding regions

In a comparative alignment with the mitogenomes of other fishes, between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes, there was control region or displacement loop (D-loop) of 919 bp in length in the *C. auratus* var. *pingxiangnensis* mitogenome (Table 2). This non-coding sequence appears to correspond to the control region because it has conserved sequence blocks (CSBs) 1-3 and termination-associated sequence (TAS), and CSB-D, -E, and -F were also detected, which are typically present in the fish control region (Tables 6 and 7, Figure 2). Another relatively conserved element is the TAS located at the 5' end of the control region. Two putative TASs are identified in the control region. Moreover, the palindromic sequence motifs TACAT and ATGTA are repeated three and two times, respectively. An AT-repeat sequence was located at the 5' end of the control region.

**Table 5.** Comparison of codon usage (number of codons) among *Carassius auratus* var. *pingxiangnensis* and selected vertebrates.

| Amino acid | Codon | Crucian carp | Carp | Frog | Human |
|---|---|---|---|---|---|
| Lys | AAA | 69 | 73 | 77 | 85 |
| | AAG | 8 | 4 | 8 | 10 |
| Asn | AAC | 76 | 76 | 80 | 131 |
| | AAT | 48 | 45 | 70 | 33 |
| Thr | ACA | 150 | 146 | 145 | 133 |
| | ACG | 14 | 8 | 8 | 10 |
| | ACC | 96 | 104 | 71 | 155 |
| | ACT | 45 | 41 | 83 | 51 |
| Stop | AGA | 0 | 0 | 1 | 0 |
| | AGG | 0 | 0 | 0 | 0 |
| Ser | AGC | 35 | 46 | 42 | 39 |
| | AGT | 14 | 9 | 15 | 14 |
| Met | ATA | 126 | 126 | 161 | 167 |
| | ATG | 49 | 49 | 35 | 40 |
| Ile | ATC | 94 | 137 | 110 | 196 |
| | ATT | 191 | 157 | 228 | 125 |
| Gln | CAA | 94 | 98 | 93 | 81 |
| | CAG | 5 | 1 | 6 | 9 |
| His | CAC | 80 | 79 | 56 | 79 |
| | CAT | 25 | 23 | 40 | 18 |
| Pro | CCA | 119 | 105 | 121 | 52 |
| | CCG | 6 | 12 | 11 | 7 |
| | CCC | 58 | 52 | 20 | 119 |
| | CCT | 28 | 33 | 52 | 41 |
| Arg | CGA | 52 | 45 | 50 | 29 |
| | CGG | 4 | 10 | 1 | 2 |
| | CGC | 9 | 15 | 5 | 25 |
| | CGT | 11 | 9 | 10 | 7 |
| Leu | CTA | 266 | 297 | 141 | 276 |
| | CTG | 43 | 37 | 10 | 45 |
| | CTC | 83 | 97 | 54 | 167 |
| | CTT | 98 | 82 | 159 | 65 |
| Glu | GAA | 91 | 93 | 83 | 64 |
| | GAG | 10 | 8 | 14 | 24 |
| Asp | GAC | 55 | 64 | 38 | 51 |
| | GAT | 20 | 11 | 36 | 15 |
| Ala | GCA | 128 | 129 | 114 | 80 |
| | GCG | 6 | 10 | 6 | 8 |
| | GCC | 131 | 144 | 83 | 124 |
| | GCT | 69 | 55 | 73 | 43 |
| Gly | GGA | 136 | 137 | 113 | 67 |
| | GGG | 38 | 35 | 23 | 34 |
| | GGC | 46 | 44 | 40 | 88 |
| | GGT | 31 | 32 | 44 | 24 |
| Val | GTA | 112 | 118 | 75 | 70 |
| | GTG | 26 | 18 | 17 | 18 |
| | GTC | 25 | 30 | 25 | 49 |
| | GTT | 65 | 43 | 62 | 30 |
| Stop | TAA | 9 | 4 | 3 | 2 |
| | TAG | 4 | 1 | 0 | 2 |
| Tyr | TAC | 56 | 65 | 57 | 89 |
| | TAT | 59 | 50 | 60 | 46 |
| Ser | TCA | 86 | 82 | 123 | 83 |
| | TCG | 7 | 9 | 11 | 7 |
| | TCC | 57 | 57 | 44 | 99 |
| | TCT | 36 | 29 | 66 | 32 |
| Trp | TGA | 107 | 112 | 101 | 93 |
| | TGG | 13 | 8 | 16 | 11 |
| Cys | TGC | 20 | 20 | 18 | 17 |
| | TGT | 5 | 5 | 13 | 5 |
| Leu | TTA | 129 | 100 | 217 | 73 |
| | TTG | 11 | 11 | 20 | 16 |
| Phe | TTC | 118 | 146 | 103 | 141 |
| | TTT | 109 | 76 | 125 | 77 |

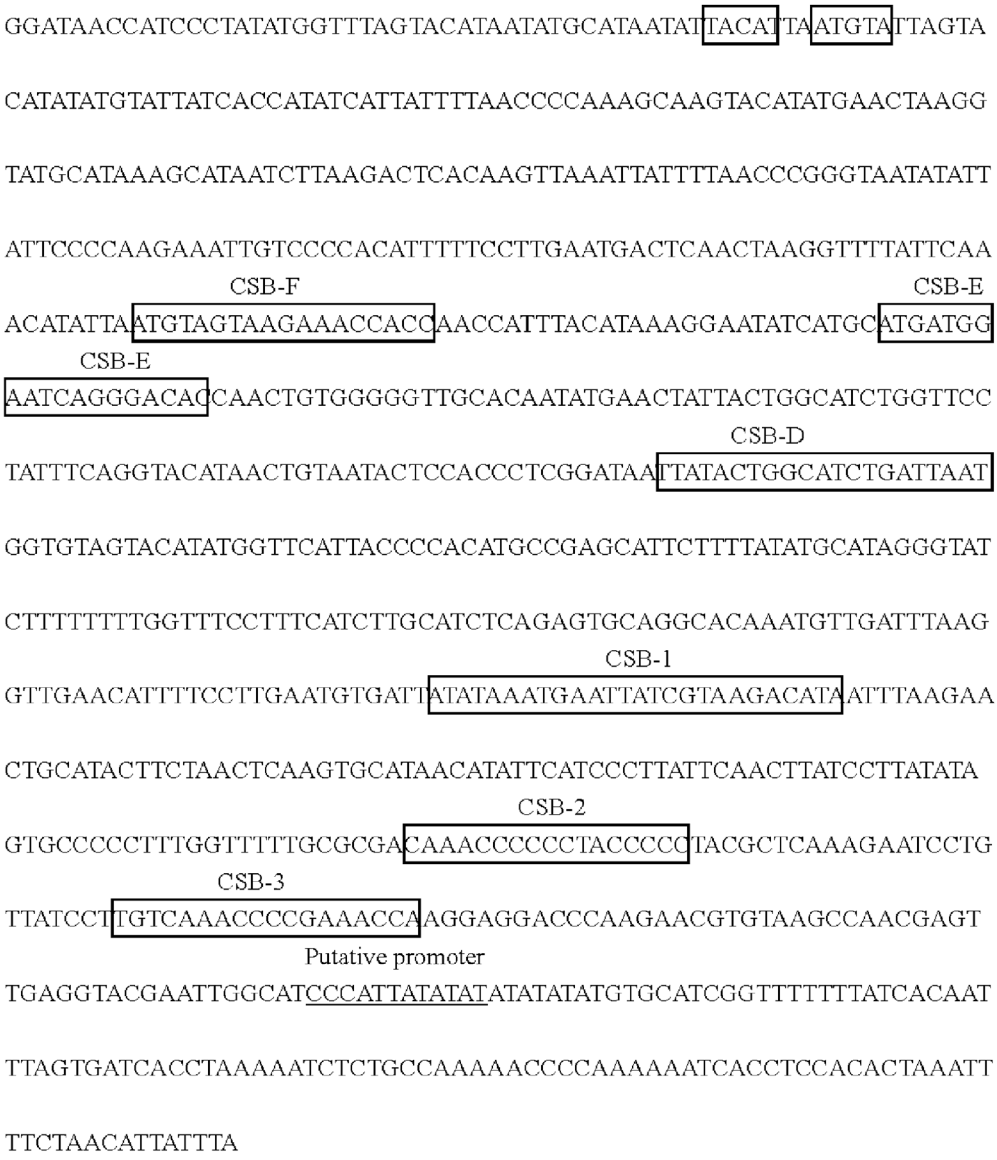


**Figure 2.** Complete sequences of the control region of *Carassius auratus* var. *pingxiangnensis*. The sequences are presented at the light-strand sequence from the 5' to 3'-end. In the control region, the putative conserved elements (palindromic sequence motifs, CSB-F, CSB-E, CSB-D, CSB-1, CSB-2, CSB-3) are boxed and marked.

Between $tRNA^{Asn}$ and $tRNA^{Cys}$, there is a 33-bp noncoding sequence that is similar to the $O_L$ in other vertebrates (Wong and Clayton, 1985). The region was predicted to be capable of folding into a stable stem-loop secondary structure with 8 bp in the stem and 14 bp in the loop. Interestingly, the characteristic sequence motif 5'-GCCGG-3' at the base of the stem within $tRNA^{Cys}$ (Elmerot et al., 2002) was mutated to 5'-GGCGG-3'.

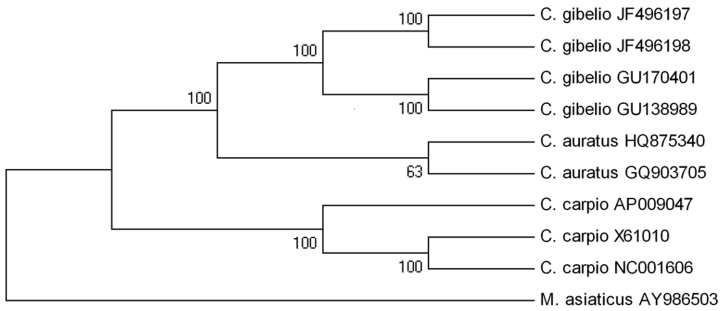


**Figure 3.** The consensus phylogenetic relationship of *Carassius auratus* var. *pingxiangnensis* with other crucian carp and *Cyprinus carpio*. *Myxocyprinus asiaticus* was used as outgroup.

## Phylogenetic analysis

Phylogenetic analysis was used to estimate the relationships among *C. auratus* var. *pingxiangnensis* and nine other types of fish (Figure 3). *M. asiaticus* (GenBank accession No. AY986503) was used as the outgroup. The results showed that *C. auratus* var. *pingxiangnensis* belongs to *Carassius* rather than *Cyprinus*; *C. auratus* var. *pingxiangnensis* was clustered with *C. auratus* (GenBank accession No. GQ903705) and then with *C. gibelio* (GenBank accession No. JF496197, JF496198, GU170401, and GU138989).

Systemic developmental trees of the crucian carp with different chromosome ploidies were built (Figure 4), and the results showed that diploid *C. auratus auratus* (GenBank accession No. GU086395) was clustered with triploid *C. auratus auratus* (GenBank accession No. GU086396), then sister-grouped with tetraploid *C. auratus auratus* (GenBank accession No. GU086397), and finally clustered with other diploid, triploid, and tetraploid *C. auratus* (GenBank accession No. GQ303444, GQ903705, AY714387, and HQ875340).

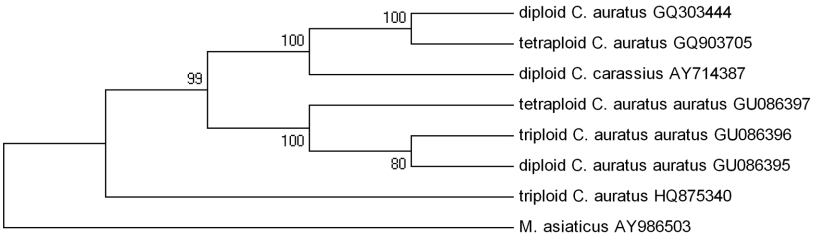


**Figure 4.** Neighbor-joining tree of diploid, triploid and tetraploid the crucian carp with other crucian carp. *Myxocyprinus asiaticus* was used as outgroup.

# DISCUSSION

## General features of the *C. auratus* var. *pingxiangnensis* mitogenome

The complete mitogenome sequence of *C. auratus* var. *pingxiangnensis* was 16,576 bp in length, making it smaller than all other available sequenced species within Cyprinidae (Wang

et al., 2008; Komiyama et al., 2009). The length variation of mitogenomes in these species is predominately due to the number and size of non-coding regions. The structural organization of the *C. auratus* mitogenome (Figure 1) was similar to earlier described patterns for other vertebrates. As in most vertebrates, the gene organization of *C. auratus* var. *pingxiangnensis* (Figure 1) is identical to the other crucian carp (GenBank accession No. GU086395, GU170401, and AY714387). The characteristics of 13 protein-coding genes are almost identical to those of other crucian carp (GenBank accession No. GU086395, GU170401, and AY714387). The reason could be that mitochondrial genomes are highly conserved in vertebrates with a few exceptions.

Besides identical gene order and content among *C. auratus* var. *pingxiangnensis* and other Cyprinidae, the unequal codon usage and patterns of strand-specific nucleotide bias were observed in the fish. We concluded that a common feature for the protein-coding genes in *C. auratus* var. *pingxiangnensis* is that NNR and NNN codon families end mostly in A, and NNY synonymous codon families end mostly in C, which was similar to what was observed in zebrafish and other vertebrates (Broughton et al., 2001). These observations suggest that the general codon usage bias pattern was also found in the triploid crucian carp. The reasonable explanation for the unequal codon usage in the mitochondrial protein genes is that codon usage is generally biased toward the available tRNA-specific nucleotide bias in mtDNA (Broughton et al., 2001). There are only 22 tRNAs in the mitochondria; therefore, there is only one specific tRNA species for most amino acids except leucine and serine. In *C. auratus* var. *pingxiangnensis*, the most frequently used codons for most amino acids are those with matching tRNAs; this was also true in zebrafish (Broughton et al., 2001). The reason for this may be that there is some advantage to match codons and anticodons in protein translation.

An orthodox initiation codon, ATG, was used for most protein-coding genes; only COI used GTG as the initiation codon. A diverse pattern of codon usage within stop codons seems to be a common tendency in fish mitogenomes. The codon usage pattern in the 13 protein-coding genes of the *C. auratus* mitogenome was similar to that observed in bony fishes (e.g., *Collichthys niveatus*; Cheng et al., 2012).

The complete set of 22 tRNA genes that are usually found in metazoans is present in the *C. auratus* mitogenome. Typical cloverleaf secondary structures were predicted for 21 tRNA genes with the exception of $tRNA^{Ser(AGY)}$, although G-U wobbles and other atypical pairings were allowed in the stem regions. Stem mismatches seem to be a common phenomenon for mitochondrial tRNA genes and are probably repaired via a post-transcriptional editing process (Lavrov et al., 2000). The $tRNA^{Ser(AGY)}$ found in the *C. auratus* mitogenome had no recognizable DHU stem, which is similar to almost all vertebrate mitogenomes (e.g., Cui et al., 2007; Zhou et al., 2009; Cheng et al., 2011). Although this $tRNA^{Ser(AGY)}$ has been described as a pseudo-gene to explain its unusual structure, Ohtsuki et al. (2002) suggested that it could also perform its function by adjusting its structural conformation to fit the ribosome in a similar way to that of usual tRNAs in the ribosome. As in all other mitogenomes described so far, two rRNA genes were identified in *C. auratus* var. *pingxiangnensis*, a small (12S) subunit of rRNA comprising 954 bp and a large (16S) subunit of 1681 bp. Their locations in the genome were identical to those of other teleosts (Elmerot et al., 2002).

## Comparative analysis of control region features in cyprinids with different ploidy levels

The mitochondrial control region is the only major non-coding segment in the vertebrate

mitogenome. The control region was 919 bp in length in the *C. auratus* var. *pingxiangnensis* mitogenome (Table 2), which was typical of lengths found for other bony fishes (781-1200 bp). This AT-rich region usually evolves relatively fast as a result of few selective constraints and is identified as the source of size variation in the whole mitogenome, whereas its control elements that are related to regulatory functions are known to be highly conserved (Cui et al., 2007).

Three central CSBs (CSB-D, CSB-E, and CSB-F) can be detected in fishes examined to date (Lee et al., 1995). In comparative studies of the control region in cyprinids with different ploidy levels, CSB-F, CSB-E, and CSB-D were also identified in the 5' ends of the control regions (Table 6). Compared with the flanking domains, the three central domains are also highly conserved. Sequences of CSB-F in *C. auratus* var. *pingxiangnensis* were the same as those of diploid Japanese crucian carp and triploid allotriploid carp. Sequences of CSB-D in seven different ploidy cyprinids were the same. In the CSB domain, CSB-1, CSB-2, and CSB-3 were identifiable in seven different cyprinids (Table 7). They showed strong similarity to CSBs that were identified in other vertebrate sequences (Roe et al., 1985).

**Table 6.** Comparison of the central conserved blocks in the red crucian carp (RC), common carp (CC), allotetraploid (AT), Japanese crucian carp (JC), allotriploid carp (TC), allotriploid common carp (TO), and natural red crucian carp (NRC).

| Species | Central conserved blocks | | |
|---|---|---|---|
| | CSB-F | CSB-E | CSB-D |
| RC (diploid) | ATGTAG-TAAGAGACCACC | ATG-ATAGAATCAGGGACAT | TTATAC-TGGCATCTGATTAA |
| CC (diploid) | ******G********** | **********A*****A | ******-**************** |
| AT (tetraploid) | ******- ********** | ******************** | ******-**************** |
| JC (diploid) | ******-*****A****** | ******G******A*****C | ******-**************** |
| TC (triploid) | ******-*****A****** | ******G******A*****C | ******-**************** |
| TO (triploid) | ******-********** | **********A*****A | ******-**************** |
| NRC (triploid) | ******-*****A****** | ******G**********C | ******-**************** |

Data of RC, CC, AT, JC, TC, and TO are quoted from Guo et al. (2003).

**Table 7.** Comparison of the conserved blocks in the red crucian carp (RC), common carp (CC), allotetraploid (AT), Japanese crucian carp (JC), allotriploid crucian carp (TC), allotriploid common carp (TO), and natural triploid crucian carp (NRC).

| Species | Conserved blocks | | |
|---|---|---|---|
| | CSB1 | CSB2 | CSB3 |
| RC (diploid) | ATATAAATGAATTATCGTAAGACATA | CAAACCCCCCTACCCCC | TGTCAAACCCCGAAACCA |
| CC (diploid) | *****T**T* ******** **** **** | A********T******* | ****************** |
| AT (tetraploid) | ********* ********** **** | ****************** | ****************** |
| JC (diploid) | ********* ********** **** | ********T******* | ****************** |
| TC (triploid) | ********* ********** **** | ****************** | ****************** |
| TO (triploid) | *****T**T* ******** **** **** | ********T******* | ****************** |
| NRC (triploid) | ********* ********** **** | ****************** | ****************** |

Data of RC, CC, AT, JC, TC, and TO are quoted from Guo et al. (2003).

## Phylogenetic relationships of *C. auratus* var. *pingxiangnensis*

The tree topologies based on complete mtDNA sequences in this study were identical and were statistically well supported by high bootstrap and posterior probability values. The mi-

togenomic data provided strong support that *C. auratus* var. *pingxiangnensis* was clustered with *C. auratus* (GenBank accession No. GQ903705) and then sister-grouped with *C. gibelio* (GenBank accession No. JF496197, JF496198, GU170401, and GU138989), which were consistent with previous findings using the Cytb gene (Fu et al., 2007). The closest phylogenetic relationship was between *C. auratus* var. *pingxiangnensis* and the common crucian carp.

Polyploidization is well known as an important evolutionary force in plants, and indications of its importance in the evolution of animals are constantly accumulating (Le Comber and Smith, 2004; Slechtová et al., 2006). Examples of fish groups in which changes in ploidy level have been already identified as key events in their evolution include Acipenseridae (Ludwig et al., 2001), Catostomidae (Ueno et al., 1988), Salmonidae (Crespi and Fulton, 2004), and Cyprinidae (Alves et al., 2001; David et al., 2003). Yang et al. (2009) reported the evolutionary relationship between natural diploid and tetraploid weather loach *Misgurnus anguillicaudatus* and reported that the tetraploid loaches were at the derivative clade, while diploid loaches were at the basal clade of MP and NJ trees. Yan et al. (2009) analyzed the phylogenetic relationship of different polyploid blunt snout bream and concluded that these artificial polyploid hybrids possess the mtDNA of the maternal parent regardless of their ploidy level. In this study, the phylogenetic relationships among branching lineages of ploidy showed that the same ploidy crucian carp did not form a monophyletic lineage; this result was different from that of Yang et al. (2009) and was similar to the conclusion of Yan et al. (2009).

## ACKNOWLEDGMENTS

We are grateful to Pingxiang Fisheries Research Institute for providing the samples. We also appreciate anonymous reviewers for providing valuable comments on the manuscript. Research supported by the National Natural Science Foundation of China (NSFC #31260282 and #31040082) and the Modern Agricultural Industry Technology System of Jiangxi (#JXARS-02).

## REFERENCES

Alves MJ, Coelho MM and Collares-Pereira MJ (2001). Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. *Genetica* 111: 375-385.

Anderson S, Bankier AT, Barrell BG, de Bruijn MH, et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.

Broughton RE, Milam JE and Roe BA (2001). The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA. *Genome Res.* 11: 1958-1967.

Cheng J, Ma GQ, Song N and Gao TX (2012). Complete mitochondrial genome sequence of bighead croaker *Collichthys niveatus* (Perciformes, Sciaenidae): a mitogenomic perspective on the phylogenetic relationships of Pseudosciaeniae. *Gene* 491: 210-223.

Cheng Y, Wang R and Xu T (2011). The mitochondrial genome of the spinyhead croaker *Collichthys lucida*: genome organization and phylogenetic consideration. *Mar. Genomics* 4: 17-23.

Crespi BJ and Fulton MJ (2004). Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. *Mol. Phylogenet. Evol.* 31: 658-679.

Cui P, Ji R, Ding F, Qi D, et al. (2007). A complete mitochondrial genome sequence of the wild two-humped camel (*Camelus bactrianus ferus*): an evolutionary history of camelidae. *BMC Genomics* 8: 241.

David L, Blum S, Feldman MW, Lavi U, et al. (2003). Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. *Mol. Biol. Evol.* 20: 1425-1434.

Elmerot C, Arnason U, Gojobori T and Janke A (2002). The mitochondrial genome of the pufferfish, *Fugu rubripes*, and ordinal teleostean relationships. *Gene* 295: 163-172.

Fu XQ, Zeng LG, Hong YJ and Sheng JQ (2009). Sequence analysis of mitochondrial cytochrome b gene from natural triploid mutant, *Carassius auratus* var. *pingxiangnensis*. *Chinese J. Zool.* 44: 97-101.

Guo XH, Liu SJ and Liu Y (2003). Comparative analysis of the mitochondrial DNA control region in cyprinids with different ploidy level. *Aquaculture* 224: 25-38.

Gutell RR (1993). The Simplicity Behind the Elucidation of Complex Structure in Ribosomal RNA. In: The Translational Apparatus (Nierhaus KH, ed.). Plenum Press, New York, 477-488.

Hong YJ, Yu ZJ, Zhou L and Gui JF (2005). A population of red-transparent, triploid *Carassius auratus*. *J. Fish Biol.* 67: 1139-1143.

Kartavtsev YP, Jung SO, Lee YM and Byeon HK (2007). Complete mitochondrial genome of the bullhead torrent catfish, *Liobagrus obesus* (Siluriformes, Amblycipididae): genome description and phylogenetic considerations inferred from the Cytb and 16S rRNA genes. *Gene* 396: 13-27.

Kawahara R, Miya M, Mabuchi K, Lavoue S, et al. (2008). Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): a new perspective based on whole mitogenome sequences from 75 higher teleosts. *Mol. Phylogenet. Evol.* 46: 224-236.

Komiyama T, Kobayashi H, Tateno Y, Inoko H, et al. (2009). An evolutionary origin and selection process of goldfish. *Gene* 430: 5-11.

Kumar S, Nei M, Dudley J and Tamura K (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* 9: 299-306.

Lavrov DV, Brown WM and Boore JL (2000). A novel type of RNA editing occurs in the mitochondrial tRNAs of the centipede *Lithobius forficatus*. *Proc. Natl. Acad. Sci. U. S. A.* 97: 13738-13742.

Le Comber SC and Smith C (2004). Polyploidy in fishes: patterns and processes. *Biol. J. Linn. Soc.* 82: 431-442.

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

Liu YG, Kurokawa T, Sekino M, Tanabe T, et al. (2013). Complete mitochondrial DNA sequence of the ark shell *Scapharca broughtonii*: an ultra-large metazoan mitochondrial genome. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 8: 72-81.

Lowe TM and Eddy SR (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955-964.

Ludwig A, Belfiore NM, Pitra C, Svirsky V, et al. (2001). Genome duplication events and functional reduction of ploidy levels in sturgeon (Acipenser, Huso and Scaphirhynchus). *Genetics* 158: 1203-1215.

Meng XP, Zhao N, Shen X and Hao J (2012). Complete mitochondrial genome of *Coelomactra antiquate* (Mollusca: Bivalvia): the first representative from the family Mactridae with novel gene order and unusual tandem repeats. *Comp. Biochem. Phys. D* 7: 175-179.

Miya M and Nishida M (1999). Organization of the mitochondrial genome of a deep-sea fish, *Gonostoma gracile* (Teleostei: Stomiiformes): first example of transfer RNA gene rearrangements in bony fishes. *Mar. Biotechnol.* 1: 416-0426.

Miya M and Nishida M (2000). Use of mitogenomic information in teleostean molecular phylogenetics: a tree-based exploration under the maximum-parsimony optimality criterion. *Mol. Phylogenet. Evol.* 17: 437-455.

Ohtsuki T, Kawai G and Watanabe K (2002). The minimal tRNA: unique structure of *Ascaris suum* mitochondrial tRNA(Ser)(UCU) having a short T arm and lacking the entire D arm. *FEBS Lett.* 514: 37-43.

Roe BA, Ma DP, Wilson RK and Wong JF (1985). The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *J. Biol. Chem.* 260: 9759-9774.

Saitoh K, Sado T, Mayden RL, Hanzawa N, et al. (2006). Mitogenomic evolution and interrelationships of the Cypriniformes (*Actinopterygii: Ostariophysi*): the first evidence toward resolution of higher-level relationships of the world's largest freshwater fish clade based on 59 whole mitogenome sequences. *J. Mol. Evol.* 63: 826-841.

Slechtová V, Bohlen J, Freyhof J and Rab P (2006). Molecular phylogeny of the Southeast Asian freshwater fish family Botiidae (Teleostei: Cobitoidea) and the origin of polyploidy in their evolution. *Mol. Phylogenet. Evol.* 39: 529-541.

Tamura K, Dudley J, Nei M and Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24: 1596-1599.

Ueno K, Nagase A and Ye YJ (1988). Tetraploid origin of the karyotype of the Asian sucker, *Myxocyprinus asiaticus*. *Jpn. J. Ichthyol.* 34: 512-514.

Wang C, Chen Q, Lu G, Xu J, et al. (2008). Complete mitochondrial genome of the grass carp (*Ctenopharyngodon idella*, Teleostei): insight into its phylogenic position within Cyprinidae. *Gene* 424: 96-101.

Wong TW and Clayton DA (1985). *In vitro* replication of human mitochondrial DNA: accurate initiation at the origin of light-strand synthesis. *Cell* 42: 951-958.

Yan JP, Guo XH, Liu SJ and Xiao J (2009). Maternal inheritance in polyploid fish inferred from mitochondrial ATPase genes analysis. *Prog. Nat. Sci.* 19: 693-698.

Yang CT, Cao L, Wang WM and Yang Y (2009). Comparative and evolutionary analysis in natural diploid and tetraploid weather loach *Misgurnus anguillicaudatus* based on cytochrome b sequence data in central China. *Environ. Biol. Fish.* 86: 145-153.

Zhang P, Zhou H, Liang D, Liu YF, et al. (2005). The complete mitochondrial genome of a tree frog, *Polypedates megacephalus* (Amphibia: Anura: Rhacophoridae), and a novel gene organization in living amphibians. *Gene* 346: 133-143.

Zhou Y, Zhang JY, Zheng RQ, Yu BG, et al. (2009). Complete nucleotide sequence and gene organization of the mitochondrial genome of *Paa spinosa* (Anura: Ranoidae). *Gene* 447: 86-96.