Molecular phylogeny of the Romanian cyprinids from the Danube River

CĂTĂLINA LUCA*, ADINA MANEA*, ADRIAN LUCA, ANCA DINISCHIOTU*, MARIETA COSTACHE***

**University of Bucharest, Department of Biochemistry and Molecular Biology, 91-95, Spl. Independentei, 050095 Bucharest, Romania*

***Faculty of Electronics, Telecommunications and Information Technology, University Politehnica of Bucharest*

Abstract

The family Cyprinidae is one of the largest families of fish in the world and a well-known component of the East Asian freshwater fish fauna. The majority of research on animals has used single mitochondrial DNA genes to assess population or low-level taxonomic relationships. Phylogenetic relationships between Romanian cyprinids (Teleostei: Cypriniformes: Cyprinidae) from the Danube river were investigated comparing cox2 mitochondrial gene sequences from ten species (Barbus barbus, Arischthys nobilis Hypophthalmichthys molitrix, Scardinius erytrophtalmus, Tinca tinca, Rutilus rutilus, Cyprinus carpio, Abramis brama, Carassius auratus gibelio, Carassius carassius). Scardinius erytrophtalmus was newly sequenced. Phylogenetic analysis indicated that there are two principal lineages in Cyprinidae: cyprinine and leuciscine. The cyprinine group includes the Cyprinus, Carassius, Barbus and Tinca genera. The taxonomic position of genus Tinca is controversial, because the neighbor-joining and maximum parsimony analysis place it in the cyprinine group alongside with the Carassius genus and in the maximum likelihood tree it appears as a paraphyletic group with the leuciscine lineage. The leuciscine group is divided into two clades. One includes the genera Abramis, Scardinius and Rutilus and the other includes the Arischthys and Hypophthalmichthys species which originated from East Asia. Cobitis danubialis was used as an outgroup species.

Keywords: cyprinids, phylogeny, mitochondrial DNA, cox 2 gene

Introduction

Cyprinidae, the largest fish family, comprises approximately 210 recognized genera and 2010 species [11] widely distributed in Eurasia, East Indian Island, Africa, and North America [9]. It is difficult to build a comprehensive phylogeny of *Cyprinidae* due to the large number of genera and species. Previous systematic analyses have focused on morphology or, more recently, mitochondrial DNA sequences - mtDNA [1; 3; 8; 19; 16; 13].

Most animal mitochondrial genomes contain 37 genes, including 13 protein-coding genes, 2 ribosomal RNAs (rRNA) and 22 transfer RNAs (tRNA) necessary for translation of the proteins encoded by the mtDNA [2]. They also possess a major non-coding control region that contains the initial sites for mtDNA replication and mtRNA transcription. The mitochondrial genome generally evolves at elevated rates (5–10 times) compared to single copy nuclear genes, however its gene order often remains unchanged over long periods of evolutionary time, with some exceptions [2]. The genetic code of mitochondrial genomes is more degenerated and thus less constrained than the universal eukaryotic nuclear code [10].

Mitochondrial DNA-derived markers have become popular for evolutionary studies, as the data obtained by their analysis may yield significant insights into the evolution of both the organisms and their genomes [2,4].

Phylogenetic relationships between Romanian cyprinids (Teleostei: Cypriniformes: Cyprinidae) from the Danube river were investigated comparing cox2 mitochondrial gene sequences from ten species (*Barbus barbus, Arischthys nobilis Hypophthalmichthys molitrix, Scardinius erytrophtalmus, Tinca tinca, Rutilus rutilus, Cyprinus carpio, Abramis brama, Carassius auratus gibelio, Carassius carassius*). *Scardinius erytrophtalmus* was newly sequenced and *Cobitis danubialis* was designed as an outgroup species, considering that it is a species appropriate to the cypriniform taxa.

Materials and methods

DNA extraction

Total cellular DNA was extracted from the muscle of the Romanian fish species (*Cobitis danubialis, Barbus barbus, Arischthys nobilis Hypophthalmichthys molitrix, Scardinius erytrophtalmus, Tinca tinca, Rutilus rutilus, Cyprinus carpio, Abramis brama, Carassius auratus gibelio, Carassius carassius*) following the protocol Wizard Genomic DNA Purification Kit (Promega).

PCR Amplification and Sequencing

Polymerase chain reaction (PCR) was used to amplify the partial sequence of the cox2 mitochondrial gene (302pb). According to complete *cox 2* genes sequences of the common carp (*Cyprinus carpio*) and goldfish (*Carassius auratus*), primer set COX2-F (5`-AGG ACA CCA ATG ATA CTGA AG-3`) /COX2-R (5`-GTT TAA AGT CTC GTA ACA GGC-3`) were designed for the amplification of a fragment from *cox 2* gene. PCR products were sequenced using the ABI PRISM ® BigDye ^{TM} Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems) on an automated DNA sequencer (Applied Biosystems 310) following the manufacturer's instructions. DNA sequences of both strands were obtained using the same primer set as the PCR amplification (forward and reverse).

Sequence alignment and phylogenetic analysis

Multiple alignments were performed using CLUSTAL X [15] and then manually refined. Aligned sequences were analyzed by the maximum parsimony (MP; heuristic searches, TBR branch-swapping algorithm) [5; 7], maximum likelihood (ML; setting the HKY+Γ model: Hasegawa, Kishino, and Yano's model with a gamma distribution of substitution rate among different nucleotide sites) [6] and neighbor-joining (NJ; setting the Hasegawa-Kishino-Yano the 85 parameter model -HKY85) [12]. MP, ML, and NJ analyses were performed using PAUP* 4.0b10 [14]. The internal stability of the inferred MP and NJ trees was measured by bootstrapping using 1000 replications. *Cobitis danubialis,* fish species of the *Cobitidae* family, were used as a root for the *Cyprinidae* phylogeny.

Results and discussions

The PCR reactions have led to the amplification of a 302bp ADN fragment from the cox2 gene of all fish species included in the study and have been checked by agarose gel electrophoresis. All the sequences obtained for these species were compared with those which are corresponding with from GenBank (Table 1). The resulted identity was greater than 95%, denoting a close phylogenetic relationship between analyzed species.

The sequence obtained for the *Scardinius erytrophtalmus* species was introduced in the GenBank (www.ncbi.nlm.nih.gov.url/ GenBank) and received the following accession number: EF112529.

Tabel 1. The comparison between our nucleotide cox1 sequence and nucleotide collection from GenBank

Species	**Identities (%)**	**Database species and accession number**
Carassius auratus gibelio	100	*Carassius auratus gibelio /* AY704452
Carassius carassius	95/	*Carassius carassius/* AY714387
	95	Carassius auratus auratus/ AB111951
Hypophthalmichthys molitrix	100	*Hypophthalmichthys molitrix/* AY704457
Rutilus rutilus	98/	*Scardinius erythrophthalmus/* EF112529
	98	*Rutilus rutilus lacustris*/ AY704466
Barbus barbus	98	*Barbus barbus*/ AB238965
Tinca tinca	100	*Tinca tinca/* AY7022656
Abramis brama	98/	*Scardinius erythrophthalmus/* EF112529
	95	*Rutilus rutilus lacustris*/ AY704466
Cyprinus carpio	100	*Ciprinus carpio/* EU260040
Scardinius erythrophthalmus	100	*Scardinius erythrophthalmus*/ EF112529
Arischthys nobilis	100	*Hypophthalmichthys nobilis*/ EU343733

The sequences were aligned with the Clustal X program (see Figure 1). The alignment results were used to build the phylogenetic trees through three different methods: neighbor-joining method, parsimony method and likelihood method. Of a total of 306 characters analyzed 253 were constant, 15 were variable (parsimony-uninformative) and 38 were parsimony-informative. The combined dataset resulted in the best likelihood score (-lnL=891.55993) for the HKY+Γ model. The estimated nucleotide empirical frequencies were: A=0.25780, C=0.27235, G=0.21102 and T=0.25884. An overall Ts/Tv ratio of 4.53 was estimated for this dataset. Base composition was calculated across all taxa for 1st, 2nd, and 3rd codon positions and all codon positions combined. The divergence between the two species was estimated based on the HKY85 test (table 2). Pair-wise sequence divergence between taxa varied from 0.6% (between *Scardinius erythrophthalmus* and *Rutilus rutilus*) to 26.6% (between *Cobitis danubialis* and *Abramis brama*).

Table 2. Levels of nucleotide divergence within and between ten cyprinid species, together with those for the outgroup *C. danubialis*. The estimates were based on HKY85 model.

	1	2	3	4	5	6	7	8	9	10	11
1 Cobitis danubialis	-										
2 Carassius auratus	0.200	-									
3 C. carassius	0.244	0.051	-								
4 H. molitrix	0.199	0.160	0.176	-							
5 A. nobilis	0.208	0.153	0.186	0.044	-						
6 Rutilus rutilus	0.264	0.144	0.136	0.130	0.131	-					
7 Sc. erytrophtalmus	0.264	0.152	0.145	0.138	0.139	0.006	-				
8 Abramis brama	0.266	0.146	0.146	0.139	0.140	0.032	0.039	-			
9 Tinca tinca	0.263	0.137	0.161	0.160	0.137	0.151	0.158	0.138	-		
10 Cyprinus carpio	0.186	0.156	0.156	0.144	0.186	0.203	0.212	0.198	0.161	-	
11 Barbus barbus	0.209	0.154	0.174	0.159	0.203	0.221	0.219	0.215	0.219	0.122	-

```
Carassius_auratus_Gibelio        ----------------------------TAGTTGTCCCAATAGAGTCCCC
Carassius_carassius              ----------------------------TAGTTGTCCCAATGGAGTCCCC
Hypophthalmichthys_molitrix      ----------------------------TAGTAGTCCCCATAGAATCGCC
Arischthys_nobilis               ----------------------------TAGTAGTCCCCATAGAATCGCC
Rutilus_rutilus                  ------------------------------AGTAGTTCCGATAGAGTCACC
Scardinius_erytrophtalmus        ------------------------------AGTAGTTCCGATAGAGTCACC
Abramis_brama                    --------------------------------GTTGTTCCGATAGAATCACC
Tinca_tinca                      ------------------------------AGTTGTCCCAATAGAATCACC
Cyprinus_carpio                  ----------------------------TAGTTGTTCCAATAGAATCCCC
Barbus_barbus                    TTTCGACTACTAGAAACAGATCATCGAATAGTTGTTCCAATAGAATCCCC
Cobitis_danubialis               --------------------------AATAGTAGTTCCAATAGAATCACC
** ** ** ** ** ** **

Carassius_auratus_Gibelio        AGTCCGTGTCTTAGTATCCGCTGAAGACGTACTACACTCCTGAGCCGTTC
Carassius_carassius              AGTCCGTGTTTTAGTATCCGCTGAAGACGTACTACACTCCTGAGCTGTTC
Hypophthalmichthys_molitrix      AGTTCGTGTTCTAGTATCCGCCGAAGATGTATTACACTCTTGAGCCGTTC
Arischthys_nobilis               AGTTCGTGTTCTAGTATCCGCCGAAGATGTATTACACTCCTGAGCCGTTC
Rutilus_rutilus                  AGTTCGTGTTTTAGTATCCGCAGAAGACGTATTACACTCTTGAGCAGTCC
Scardinius_erytrophtalmus        AGTTCGTGTTTTAGTGTCCGCAGAAGACGTATTACACTCTTGAGCAGTCC
Abramis_brama                    AGTTCGTGTTTTAGTATCCGCAGAAGACGTATTACACTCTTGAGCCGTCC
Tinca_tinca                      AGTACGTGTTCTAGTATCCGCTGAAGACGTCTTACACTCCTGAGCCGTCC
Cyprinus_carpio                  AGTCCGTGTCCTAGTATCTGCTGAAGACGTGCTACATTCTTGAGCTGTTC
Barbus_barbus                    AATTCGTATCCTGGTCTCTGCTGAAGACGTACTACACTCTTGAACTGTTC
Cobitis_danubialis               AATCCGTGTACTAGTCTCTGCTGAAGATGTACTTCATTCTTGAGCCGTTC
* * *** *  * ** ** ** ***** **  * ** ** *** * ** *

Carassius_auratus_Gibelio        CATCCTTAGGTGTAAAAATAGACGCAGTCCCAGGCCGACTAAATCAAACT
Carassius_carassius              CATCTTTAGGTGTAAAAATAGACGCAGTCCCCGGACGACTAAATCAAACT
Hypophthalmichthys_molitrix      CATCCCTAGGCGTAAAAATGGACGCAGTACCAGGACGACTTAACCAAACT
Arischthys_nobilis               CATCCCTGGGCGTAAAAATGGACGCAGTACCAGGACGATTAAACCAAACT
Rutilus_rutilus                  CATCTTTGGGCGTAAAAATAGACGCAGTACCAGGACGATTAAATCAAACT
Scardinius_erytrophtalmus        CATCTTTGGGCGTAAAAATAGACGCAGTACCAGGACGATTAAATCAAACT
Abramis_brama                    CATCTTTGGGCGTAAAAATAGACGCAGTGCCAGGACGATTAAATCAAACT
Tinca_tinca                      CATCATTAGGAGTTAAAATGGACGCAGTCCCAGGACGATTAAATCAAGCC
Cyprinus_carpio                  CATCCCTTGGCGTAAAAATGGACGCAGTCCCAGGACGACTGAATCAAGCC
Barbus_barbus                    CATCCCTAGGCGTAAAAATGGATGCGGTCCCAGGACGACTTAATCAAACC
Cobitis_danubialis               CGGCCCTTGGTATTAAAATAGACGGTGTCCCAGGCCGCCTAAACCAAACT
*  *  * **  * ***** ** *  ** ** ** **  * ** *** *

Carassius_auratus_Gibelio        GCTTTCATCGCCTCACGCCCAGGAG-TATTC-TACGGACA-TG-------
Carassius_carassius              GCCTTCATCGCCTCACGCCCAGGAG-TGTTC-TACGGACA-TG-------
Hypophthalmichthys_molitrix      GCCTTTATTGCCTCACGCCCAGGCG-TATTT-TACGGACA-TG-------
Arischthys_nobilis               GCTTTTATTGCCTCGCGCCCAGGCG-TATTC-TACGGACA-TG-------
Rutilus_rutilus                  GCCTTCATCGCCTCCCGCCCAGGCG-TATTTCTACGGACA-TG-------
Scardinius_erytrophtalmus        GCCTTCATCGCCTCCCGCCCAGGCG-TATTT-TACGGACAATG-------
Abramis_brama                    GCCTTCATCGCCTCCCGCCCAGGCGCTACTTCTACGGACA-TG-------
Tinca_tinca                      GCTTTTATCGCCTCGCGCCCAGGTG-TATTT-TACGGACAATG-------
Cyprinus_carpio                  GCCTTTATTGCCTCACGCCCAGGGG-TGTTT-TACGGACA-TG-------
Barbus_barbus                    GCCTTCATTGCCTCACGCCCAGGAG-TATTC-TACGGACAATGCTCTGAA
Cobitis_danubialis               GCCTTTATTGCCTCTCGCCCAGGGG-TATTC-TACGGACA----------
** ** ** ***** ******** * *  *   ********
```

Figure 1. CLUSTAL X multiple cox2 sequence alignment

All the trees obtained are widely based on the same topology. The robustness of the tree was corroborated with bootstrap analyses. The cyprinid species were grouped into two assemblages: *Cyprininae* and *Leuciscinae.* The *Cyprininae* group is represented by the species *Carassius carassius, Carassius auratus gibelio, Barbus barbus* and forms a separate branch in the MP and ML trees (fig.2A and fig.3). The cyprinine group was represented differently in the NJ tree (fig.2B): a branch contains *Barbus barbus* and *Cyprinius carpio*, while the Carassius genus and *Tinca tinca* are placed on another branch. This topology could be explained by the uncertain position of the *Tinca* species within the cyprinid family. Most researchers [17; 18] placed this species inside the leuciscine group, as it is represented in the ML tree (fig.3). The *Leuciscinae* group includes the *Hypophthalmichthyinae* clade: *Arischthys sp.* and *Hypophthalmichthys sp.* and abramini clade: *Abramis sp., Rutilus sp., Scardinius sp. Cobitis danubialis* was an outgroup species in our phylogenetic trees.

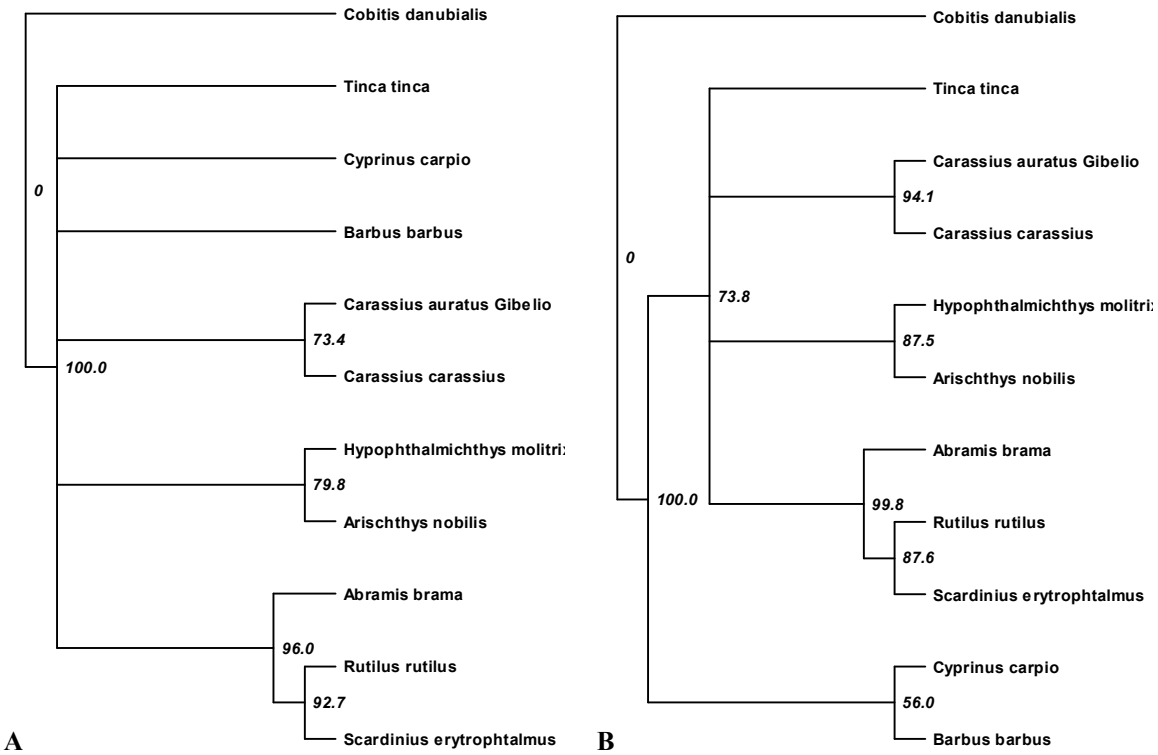

Figure 2. Phylogenetic trees built with the Maximum Parsimony (A) and Neighbor-joining (B) methods based on sequenced cox2 gene fragments. Tree length =110, Consistency index (CI) = 0.6455, Homoplasy index (HI) = 0.3545, Retention index (RI) = 0.6214, Rescaled consistency index (RC) = 0.4011. Numbers at nodes represent percentage recovery in bootstrap analysis (100 replicates).

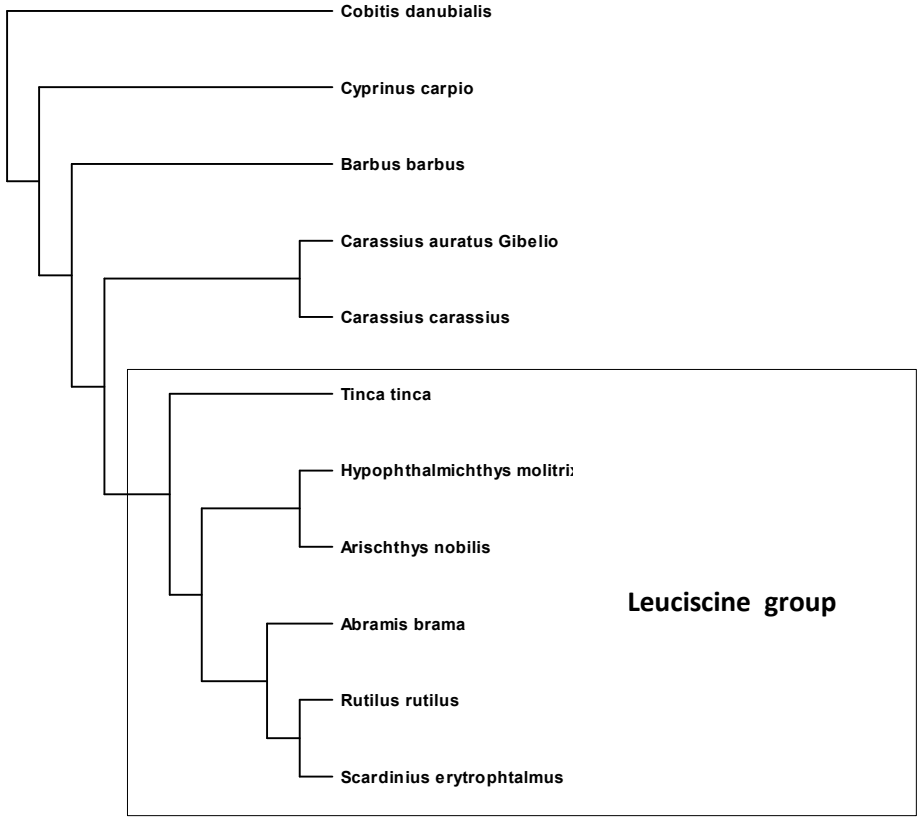

Figure 3. Phylogenetic trees built with the Maximum Likelihood method based on sequenced cox2 gene fragments. The analysis was based on the HKY+Γ model. The estimated parameters for likelihood analysis: –Ln L =891.55993; ratio Ti/Tv=4.53; Kappa= 9.044; α= 0.0.1492.

Conclusions

The topologies of neighbor-joining, maximum parsimony and maximum likelihood trees based *on cox2* sequences have enabled us to identify two major lineages in cyprinids: cyprinine (including the barbine lineage) and leuciscine (including abramini and *Hypophthalmichthyinae clade*). The taxonomic position of genus *Tinca* is controversial, because the neighbor-joining and maximum parsimony analyses place it in the cyprinine group alongside with *Carassius* genus and in the maximum likelihood tree it appears as a paraphyletic group with the leuciscine lineage. The outgroup species *Cobitis danubialis* is identified separately in all trees, as this species belongs to another fish family: *Cobitidae.*

References

1. ARAI R.,1982, A chromosome study on two Cyprinid Fishes, *Acrossocheilus labiatus* and *Pseudorasbora pumila pumila,* with notes on Eurasian Cyprinids and their karyotypes. Bull. Natn. Sci. Mus., Tokyo Ser. A 8: 131–152.
2. BOORE JL, 1999, Animal mitochondrial genomes, Nucleic Acids Res., 27:1767-1780.
3. CHEN, X., YUE, P., LIN, R., 1984, Major groups within the family Cyprinidae and their phylogenetic relationships, Acta Zootaxonomica Sinica (in Chinese), 9(4): 424-440.
4. CUROLE AP, KOCHER TD, 1999, Mitogenomics: digging deeper with complete mitochondrial genomes, Trends Ecol. Evol., 14:394-398.
5. FARRIS J.S., 1970. Methods for computing Wagner trees. Syst. Zool. 18, 374–385.
6. FELSENSTEIN J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.
7. HENDY, M.D., PENNY, D., 1982. Branch and bound algorithms todetermine minimal evolutionary trees. Math. Biosci. 59, 277–290.
8. HOWES G. J., 1991, Systematics and biogeography: An overview, in Cyprinid Fishes Systematics, Biology and Exploitation (eds. Winfield, I., Nelson, J.), New York: Chapman & Hall, 1-54.
9. MAYDEN, R.L., 1991, New world cyprinids. In: Winfield, I.J., Nelson, J.S. (Eds.), Biology of Cypinids. Chapman and Hall Ltd., London, pp. 240–263.
10. MEYER A., 1993, Evolution of mitochondrial DNA in fishes. In Biochemistry and molecular biology of fishes, edited by: Hochachka PW and Mommsen TP. Amsterdam, Elsevier Science Publishers, 1-38.
11. NELSON J.S., 1994, Fishes of the World. John Wiley and Sons, Inc., New York.
12. SAITOU, N., NEI, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
13. SHUNPING HE, RICHARD L. MAYDEN, XUZHENG WANG, WEI WANG, KEVIN L. TANG, WEI-JEN CHEN and YIYU CHEN, 2008, Molecular phylogenetics of the family *Cyprinidae* (*Actinopterygii: Cypriniformes*) as evidenced by sequence variation in the first intron of S7 ribosomal protein-coding gene: Further evidence from a nuclear gene of the systematic chaos in the family, Molecular Phylogenetics and Evolution 46,818–829.
14. SWOFFORD D., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Ver. 4.0b10. Sinauer Associates, Sunderland, MA
15. THOMPSON, J.D., GIBSON, T.J., PLEWNIAK, F., JEANMOUGIN, F., HIGGINS, D.G., 1997. The Clustal X windows interface: flexible strategies for multiple sequences alignment aided by quality analysis tools. Nucleic Acids Res.25, 4876–4882.
16. TSIGENOPOULOS C.S., P. RÀB, D. NARAN and P. BERREBI, 2002, Multiple origins of polyploidy in the phylogeny of southern African barbs (*Cyprinidae*) as inferred from mtDNA markers, Heredity 88, 466–473
17. XUZHEN WANG, JUNBING LI, SHUNPING HE, 2007, Molecular evidence for the monophyly of East Asian groups of Cyprinidae (Teleostei: Cypriniformes) derived from the nuclear recombination activating gene 2 sequences, Molecular Phylogenetics and Evolution 42, 157–170.
18. ZARDOYA R., and DOADRIO I., 1998, Phylogenetic relationships of Iberian cyprinids: Systematic and biogeographical implications, Proc. R. Soc. Lond. B 265: 1365–1372.
19. ZARDOYA, R., DOADRIO, I., 1999, Molecular evidence on the evolutionary patterns of European Cyprinids. J. Mol. Evol. 49, 227–237.