



Mitochondrial DNA variation in wild populations of *Leporinus elongatus* from the Paraná River basin

Cesar Martins, Adriane P. Wasko, Claudio Oliveira and Fausto Foresti

Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista, Botucatu, SP, Brazil.

Abstract

Leporinus elongatus, a fish species widely distributed throughout the Paraná River basin in South America, is an important fishery resource and a valuable species in aquaculture programs. Despite its great economic importance, several wild populations have been suffering a drastic reduction. The comprehension of its population structure represents an important step for the conservation of these organisms in natural environments, and also for the selection of wild stocks to be used in hatchery programs. In order to understand the genetic-population structure of *L. elongatus*, D-loop mitochondrial DNA analyses were applied in six wild populations of the species. The results were used to estimate the levels of within and among population genetic variability. Although the D-loop variations could not be correlated to the geographic distribution of these organisms, it was possible to detect high levels of genetic variability within each population and the occurrence of exclusive population haplotypes, which suggests a partial genetic differentiation among them. The obtained data can be useful in selecting fish stocks that preserve a better genetic diversity of *L. elongatus* for use in conservation and/or hatchery programs.

Key-words: conservation genetics, D-loop, genetic variability, *Leporinus elongatus*, mtDNA.

Received: May 6, 2002; accepted: January 27, 2003.

Introduction

The Anostomidae family (Characiformes), comprised of twelve genera thus far recognized, is an important freshwater fish group distributed throughout the Neotropical region (Géry, 1977). *Leporinus*, the most diversified genera of Anostomidae, is distributed from Central to South America and several species, such as *L. macrocephalus*, *L. obtusidens* and *L. elongatus*, represent important fishery resources for many South American communities. Despite the great economic importance of *L. elongatus* (piapara), a migratory species widely distributed throughout the Paraná River basin, a decrease in the amount of collected individuals in the last years has been reported (fishermen of Guaíra, Paraná State, Brazil, personal communication), probably due to habitat alterations. Surprisingly, no information is available on the genetic diversity of natural populations of this species, even though such data may be critical to their survival and conservation. Since *L. elongatus* represents an endangered fishery resource and a promising species to be cultivated in fish farms, the careful selection of appropriate wild stocks, based on genetic crite-

ria, can offer greater potential for success in species-recovery and maintenance programs and support for aquaculture activities in order to increase fish production.

In recent years, different molecular techniques, using nuclear or mitochondrial DNA (mtDNA), have provided new information concerning the genetic variability of wild and cultivated populations of several fish species (Bielawski and Pumo, 1997; Nielsen *et al.*, 1995; Perez-Enriquez *et al.*, 1999; Koh *et al.*, 1999; Iguchi *et al.*, 1999; Reilly *et al.*, 1999; Sivasundar *et al.*, 2001; Yoon and Park, 2002; Was and Wenne, 2002; among others), permitting the delineation of management units and allowing assessment of conservation priorities. The mitochondrial DNA can be an efficient tool for gathering information for population genetic analyses since it has unique characteristics due to its maternal transmission, rapid rate of evolutionary changes, transmission without recombination, and haploid inheritance (Avice, 1994). Therefore, mtDNA variations can be extremely useful for identifying and managing stocks of fish species (Grewe and Hebert, 1988; Billington *et al.*, 1992).

To date, there has been no literature on genetic stock characterization of *Leporinus elongatus* populations. Focusing on the genetic diversity of natural populations of *L. elongatus*, the aim of the present study was to present a pre-

liminary assessment of the genetic variability of six wild populations, based in the nucleotide sequence of a segment of the non-coding control region (D-loop) of the mitochondrial genome. The results were useful not only to characterize *L. elongatus*, but also to give support to recovery efforts and to the biodiversity maintenance of this fish species.

Material and Methods

Sample collection and DNA extraction

Specimens of *Leporinus elongatus* were obtained from six wild populations from the Paraná River basin (Figure 1 and Table 1). In order to compare to the results of *L. elongatus*, another four Anostomidae fish species (*Leporinus* cf. *elongatus*, *Leporinus obtusidens*, *Leporinus friderici* and *Leporinus desmotes*) and one Parodontidae species (*Parodon tortuosus*) were also analyzed. Total genomic DNA was obtained from fish caudal or anal fin tissue samples, in order to prevent the sacrifice of the animals, using a TNES-urea-digestion buffer (10 mM Tris-HCl pH 8.0; 125 mM NaCl; 10 mM EDTA pH 8.0; 0.5% SDS; 4 M urea), according to A.P. Wasko, C.A. Martins, C. Oliveira and F. Foresti (unpublished data). Further DNA purification was performed with phenol-chloroform (Sambrook and Russell, 2001).

PCR amplification of the D-loop region of the mtDNA

A segment of approximately 1.300 base pairs (bp) that corresponds to the D-loop region of the mtDNA was amplified by PCR using the primers Dloop2F (5' - GCC TAA GAG CAT CCG TCT TGT AA - 3') and Dloop2R (5' - GTC AGG ACC ATG CCT TTG TG - 3') (R.A. Torres, personal communication). Amplifications were carried out using 0.4 mM of each primer, 100 ng of total genomic DNA, 1x *Taq* buffer, 200 μ M of each dNTP and 2U of *Taq* DNA polymerase in a final reaction volume of 50 μ L. Cycling times were as follow: 94 °C (5 min) of denaturation, 35 cycles of 1 min at 92 °C, 1 min at 50 °C and 1.5 min at 70 °C. A final 7-minute extension step was done at 72 °C. The PCR-amplified products were analyzed in 1% agarose gels in 1x TAE buffer (Tris-Acetic acid-EDTA). Amplified

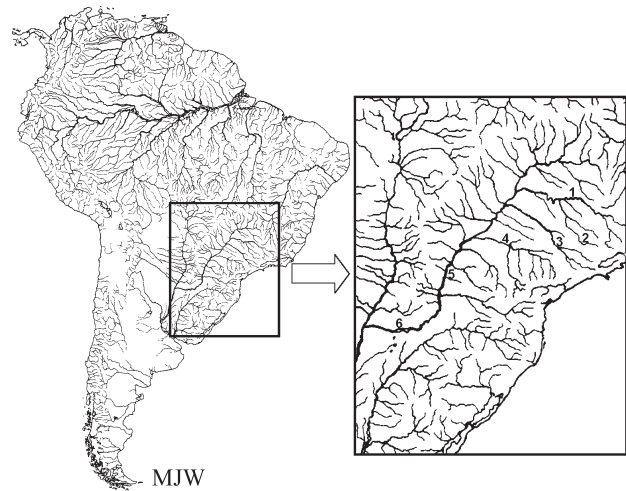


Figure 1 - South America geographic map indicating sampling localities of *Leporinus elongatus* from the Paraná River basin. (1) Grande River, (2) Mogi-Guaçu River, (3) Piracicaba/Tietê Rivers, (4) Paranapanema River, (5) Paraná River, Municipality of Guaira, and (6) Paraná River, Municipality of Ayolas. (1-5) Brazil, (6) Paraguay.

fragments were visualized with Ethidium Bromide staining (Sambrook and Russell, 2001) under UV illumination and recorded with the Kodak Digital Science 1D computer program. Molecular weights were estimated using a standard DNA marker.

DNA sequencing and D-loop sequence analyses

PCR fragments were purified using the GFX™ PCR DNA and Gel Band Purification (Amersham Pharmacia Biotech) kit and sequenced on the OpenGene Automated DNA Sequencing System I (Visible Genetics Inc.) with a Thermo Sequenase Cy 5.5 Terminator Cycle Sequencing kit (Amersham Pharmacia Biotech). Nucleic acid sequences were subjected to BLASTn (Altschul *et al.*, 1990) searches at the National Center for Biotechnology Information (NCBI), website (<http://www.ncbi.nlm.nih.gov/blast>) and they were aligned using ClustalW software (Thompson *et al.*, 1994), web site (<http://www.ebi.ac.uk/clustalw/>). Sequences of the mtDNA control region of *Leporinus elongatus* were deposited in GenBank, under the accession

Table 1 - *Leporinus elongatus* collection localities and sample sizes from the Paraná River basin.

Sampled Localities	mtDNA samples
Brazil	
1 Grande River, Municipality of São José da Barra, Minas Gerais State	09
2 Mogi-Guaçu River, Municipality of Pirassununga, São Paulo State	12
3 Paranapanema River, Municipality of Salto Grande, São Paulo State	06
4 Piracicaba River and Tietê River, Municipality of Santa Maria da Serra, São Paulo State	10
5 Paraná River, Municipality of Guaira, Paraná State	03
Paraguay	
1 Paraná River, Municipality of Ayolas, Misiones	03

numbers AY100330-AY100372. Kimura 2-parameter genetic distances (Kimura, 1980) were determined using the computer program Molecular Evolutionary Genetics Analysis (MEGA Version 2.1). The number of haplotypes, haplotype diversity, and nucleotide diversity were calculated using the program DNA SP version 3.0 (Rozas and Rozas, 1999).

Results and Discussion

Previous studies have shown that the D-loop region, *i.e.* a non-coding control region of the mtDNA, corresponds to a DNA segment with rapid evolution in the mitochondrial genome (Saccone *et al.*, 1993). Although the main domain of this region exhibits a great nucleotide similarity among different species and presents a divergence index similar to the one observed in the mitochondrial genes that code for proteins, the peripheral regions that flank the main domain present a rapid level of evolution, which generates size and base composition heterogeneity (Sbisà *et al.*, 1997). Therefore, in the present study, we intended to analyze the nucleotide sequence of one of the peripheral domains of the D-loop region of *Leporinus elongatus* in order to achieve a preliminary assessment of the genetic variability in wild populations of the species.

The PCR amplification of the total genomic DNA of several individuals of *Leporinus elongatus*, using primers D-loop2F e D-loop2R, generated a fragment of approximately 1.300 base pairs (bp) that corresponds to the mtDNA D-loop region (Figure 2). No size or heteroplasmy polymorphism was observed within or among individuals from the six sampled *L. elongatus* populations. However, distinct amplified D-loop size-fragments were detected in different *Leporinus* species, which permitted the identification of a possible species-specific mtDNA band patterns for some species of the genus, such as *L. obtusidens* and *L. desmotes* (Figure 2). A D-loop fragment of approximately 1.300bp could also be identified for *Parodon tortuosus* (Figure 2), a fish species closely related to the Anostomidae group.

After purifying the 1.300bp-amplified DNA fragment from agarose gels, the D-loop fragments of several *L.*

elongatus individual were sequenced using the primer D-loop2F. A D-loop segment of approximately 300 base pairs, flanked by the proline RNAt gene, was obtained for 43 individuals of the species sampled from the Grande River, Mogi-Guaçu River, Piracicaba/Tietê Rivers, Paranapanema River, and Paraná River populations (Table 1).

At present, all the analyzed organisms, except primates, show an AT-content higher than the GC-content in the entire D-loop region. In primates, the AT-rich regions can be identified in the peripheral domains of the D-loop region, while the GC-content is higher in the main domain (Sbisà *et al.*, 1997). The fish species *Leporinus elongatus* also presented a higher proportion of AT nucleotides in the D-loop mtDNA. The obtained D-loop nucleotide sequences of the sampled individuals of the species were 70.20% AT-rich and 29.80% GC-rich. Moreover, the amount of nucleotide bases of the D-loop region of *L. elongatus* followed the rule $A > T > C > G$, that also seems to be a common feature in the peripheral domains of the light-strand of the D-loop region of distinct organisms (Sbisà *et al.*, 1997).

The *L. elongatus* mtDNA sequences presented 198 variable sites (63.46%) and 109 conserved sites (34.93%). The conserved positions were observed at a higher frequency in the initial portion of the nucleotide sequence (nucleotides 1 to +/-170) and the variable regions were mostly visualized between nucleotides 170 and 312. Figure 3 presents the nucleotide sequence alignment of the observed variable sites of the 43 analyzed *L. elongatus* individuals sampled from distinct populations. Although a few base substitutions were identified in the D-loop sequences of the species, short insertions/deletions were more frequent. Similarly, the occurrence of short base insertions and deletions has been observed in the mtDNA sequences of other fish, such as salmonid species (Shedlock *et al.*, 1992). Traces of imperfect tandem repeats were also observed in the *L. elongatus* D-loop sequences. These tandem repeats, typically found in the 5'-end or in the 3'-end of the D-loop segment, where the DNA replication initiates and ends, respectively (Lee *et al.*, 1995; Nesbo *et al.*, 1999), can be correlated to intra and inter-individual size variations in the D-loop region (Brzuzan, 2000).

Results of the inferences of the Kimura 2-parameter genetic distances (Kimura, 1980), based on nucleotide divergence, showed that the nucleotide divergence values varied from 0.000 (distance between identical nucleotide sequences) to 0.165, with an average value equal to 0.037. The average within and among population genetic distance values for *L. elongatus* were also obtained (Table 2). The population from the Paraná River-Guaíra showed the lowest genetic distance index (average value of 0.016) in relation to the other sampled locations, indicating a higher genetic similarity among the individuals of this region. The Paranapanema River and the Paraná River-Ayolas popula-

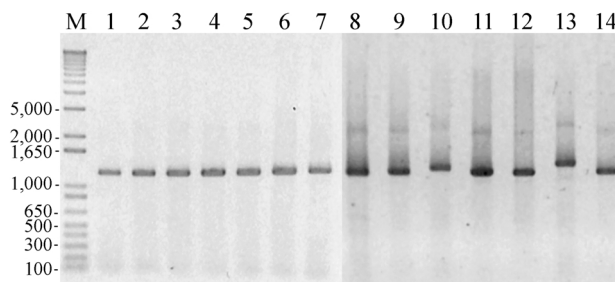


Figure 2 - 1% agarose gel of PCR-generated D-loop mtDNA fragments of *Leporinus elongatus* (1-8), *Leporinus cf. elongatus* (9), *Leporinus obtusidens* (10), *Leporinus friderici* (11-12), *Leporinus desmotes* (13), and *Parodon tortuosus* (14). M indicates 1Kb ladder for molecular weight markers.

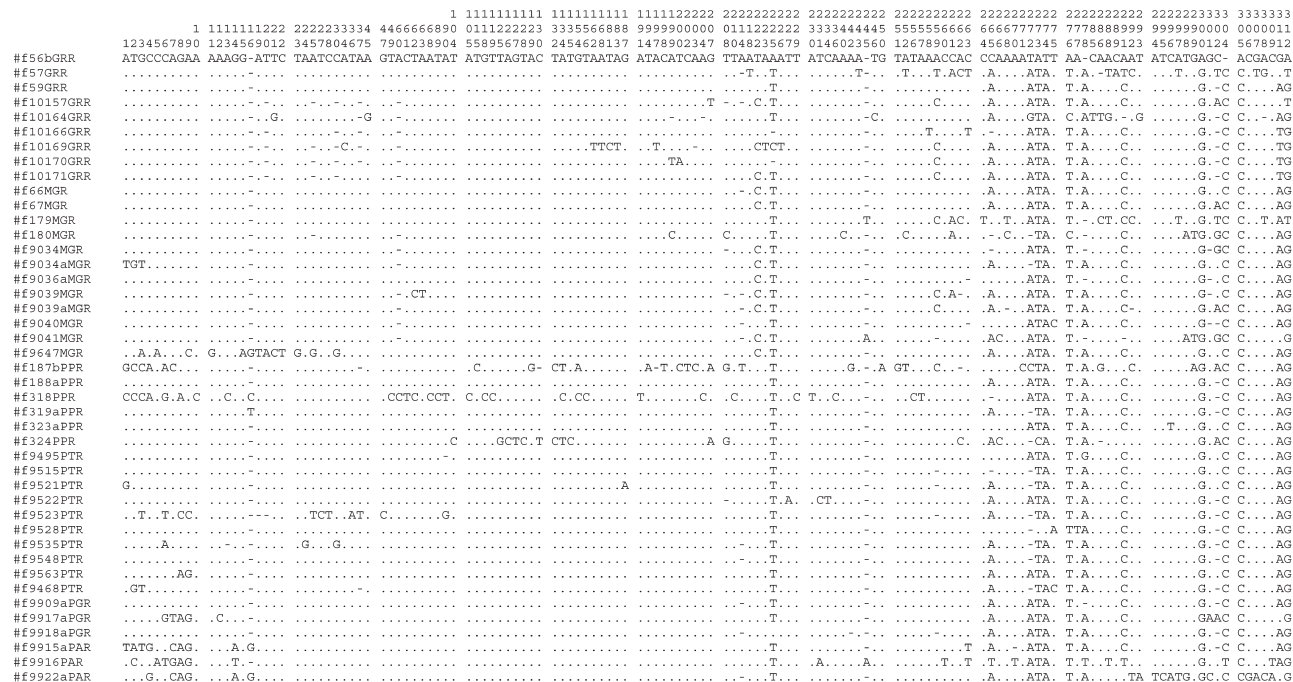


Figure 3 - Nucleotide sequence alignment of the observed variable sites in the D-loop 5'-region of 43 *Leporinus elongatus* individuals sampled from distinct populations from Grande River (GRR), Mogi-Guaçu River (MGR), Paranapanema River (PPR), Piracicaba/Tietê Rivers (PTR), Paraná River-Guaíra (PGR), and Paraná River-Ayolas (PAR). Hyphens indicate gaps and dots represent nucleotide identity.

Table 2 - Nucleotide diversity among sampled populations of *L. elongatus* (above the diagonal, columns 1-6) and average genetic distances, based on nucleotide divergence, within (column 7) and among distinct populations of *L. elongatus* (below the diagonal).

Sampled populations	1	2	3	4	5	6	7
Grande River	-	0.050	0.058	0.049	0.050	0.078	0.034
Mogi-Guaçu River	0.029	-	0.044	0.020	0.020	0.041	0.021
Paranapanema River	0.060	0.053	-	0.027	0.027	0.052	0.077
Piracicaba/Tietê Rivers	0.029	0.022	0.049	-	0.019	0.056	0.019
Paraná River-Guaíra	0.026	0.021	0.047	0.017	-	0.045	0.016
Paraná River-Ayolas	0.064	0.057	0.082	0.052	0.046	-	0.060

tions showed a greater genetic variability since higher genetic distance indices (0.077 and 0.060, respectively) were obtained for the sampled individuals from these regions. Based on the obtained D-loop nucleotide divergence indices, it was also possible to infer that the *Leporinus elongatus* populations from the Paraná River-Guaíra and from the Piracicaba/Tietê Rivers presented a higher genetic similarity, while the individuals sampled at the Paraná River-Ayolas represented the most divergent population of the species, since they presented higher genetic distance indices compared to the other analyzed populations (Table 2). A D-loop nucleotide sequence of another fish species, *Prochilodus lineatus* (Sivasundar et al., 2001), was also included in the genetic distance analysis and, as expected, it presented a lower genetic similarity with all the sampled *L. elongatus* populations, with an average nucleotide diversity of 0.540.

A total of 32 distinct D-loop mtDNA haplotypes could be identified in the six sampled populations of *L. elongatus* (Table 3). A common haplotype was not observed in any the populations of the species and several haplotypes could be considered exclusive to a specific population, suggesting a reduced gene flow among them. Differences in the haplotype number and frequency led to a high haplotype diversity that ranged from 0.848 to 1.000, with an average value equal to 0.958 (Table 3). The elevated nucleotide diversity observed in *L. elongatus* also reflected high levels of haplotype diversity. The percentage of nucleotide diversity ranged from 1.78 in the samples from the Piracicaba/Tietê Rivers to 7.70 in the Paranapanema River (Table 3).

Although the obtained data proved ineffective in separating the sampled *Leporinus elongatus* populations into distinct groups, it was possible to detect high levels of ge-

Table 3 - Sample sizes, number of D-loop mtDNA haplotypes, haplotype diversity, and nucleotide diversity, obtained for the sampled populations of *Leporinus elongatus*.

Sampled populations	Sample sizes	Number of haplotypes	Haplotype diversity	Nucleotide diversity (%)
Grande River	09	09	1.000	3.83
Mogi-Guaçu River	12	08	0.848	2.21
Paranapanema River	06	05	0.933	7.70
Piracicaba/Tietê Rivers	10	09	0.978	1.78
Paraná River-Guaíra	03	03	1.000	1.82
Paraná River-Ayolas	03	03	1.000	3.73
Total	43	32	0.958	3.51

netic variability within each population and the occurrence of exclusive population haplotypes, which suggest a partial genetic differentiation among the analyzed populations. High levels of genetic variability in the D-loop nucleotide sequence have been commonly identified in wild populations of other fish species (Igushi *et al.*, 1999). Therefore, the analysis of mtDNA haplotypes represents an important tool for the characterization of distinct populations, for the evaluation of the genetic variability, and for the identification of genetic markers, which could be extremely useful in conservation and hatchery programs of fish species.

Despite the great economic importance of *Leporinus elongatus* as a fishery resource and a valuable species in aquaculture programs, several wild populations of the species have been suffering a drastic reduction in the last years. The careful selection of appropriate natural stocks, based on genetic criteria, can offer greater potential for success in species-recovery and maintenance programs (Quattro and Vrijenhoek, 1989), and the present data on the *L. elongatus* mitochondrial D-loop could be extremely useful to choose founder wild stocks of the species that retain more DNA variability. Moreover, it would be instructional to continue the genetic characterization and monitoring of natural and captive stocks of the species in coming years, using other molecular markers such as RAPD and microsatellites, since further genetic data results will be of great value to establish a clear scene of the distribution and dynamics of the *Leporinus elongatus* populations in the Paraná River basin.

Acknowledgements

C.M and A.P.W. were supported by fellowships from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação para o Desenvolvimento da UNESP (FUNDUNESP).

References

- Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410.
- Avise JC (1994) Molecular markers, natural history and evolution. Chapman & Hall, London, 551 pp.
- Bielawski JP and Pumo DE (1997) Randomly amplified polymorphic DNA (RAPD) analysis of Atlantic Coast striped bass. *Heredity* 78:32-40.
- Billington N, Barrette RJ and Hebert PDN (1992) Management implications of mitochondrial DNA variation in walleye stocks. *North Am J Fish Manag* 12:276-284.
- Brzuzan P (2000) Tandemly repeated sequences in mtDNA control region of whitefish, *Coregonus lavaretus*. *Genome* 43:584-587.
- Géry J (1977) Characoids of the world. T.F.H. Publications, Neptune City, 672 pp.
- Grewe PM and Hebert PDN (1988) Mitochondrial DNA diversity among broodstocks of the lake trout, *Salvelinus namaycush*. *Can J Fish Aquat Sci* 45:2114-2122.
- Iguchi K, Watanabe K and Nishida M (1999) Reduced mitochondrial DNA variation in hatchery populations of ayu (*Plecoglossus altivelis*) cultured for multiple generations. *Aquaculture* 178:235-243.
- Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111-120.
- Koh TL, Khoo G, Fan LQ and Phang VPE (1999) Genetic diversity among wild forms and cultivated varieties of *Discus (Symphysodon ssp.)* as revealed by random amplified polymorphic DNA (RAPD) fingerprinting. *Aquaculture* 173:485-497.
- Lee WJ, Conroy J, Howell WH and Kocher TD (1995) Structure and evolution of teleost mitochondrial control regions. *J Mol Evol* 41:54-66.
- Nesbo CL, Fossheim T, Vollestad LA and Jakobsen KS (1999) Genetic divergence and phylogeographic relationships among European perch (*Perca fluviatilis*) populations reflect glacial refugia and postglacial colonization. *Mol Ecol* 8:1387-1404.
- Nielsen JL, Gan CA, Wright JM, Morris DB and Thomas K (1995) Biogeographic distributions of mitochondrial and nuclear markers for southern steelhead. *Mol Mar Biol Biotech* 3:281-293.
- Perez-Enriquez R, Takagi M and Taniguchi N (1999) Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. *Aquaculture* 173:413-423.

- Quattro JM and Vrijenhoek RC (1989) Fitness differences among remnant populations of endangered Sonoran topminnow. *Science* 245:976-978.
- Reilly A, Elliott NG, Grewe PM, Clabby C, Powell R and Ward RD (1999) Genetic differentiation between Tasmanian cultured Atlantic salmon (*Salmo salar* L.) and their ancestral Canadian population: comparison of microsatellite DNA and allozyme and mitochondrial DNA variation. *Aquaculture* 173:459-469.
- Rozas J and Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15:174-175.
- Saccone C, Lanave C, Pesole G and Sbisà E (1993) Peculiar features and evolution of mitochondrial genome in mammals. In: DiMauro S and Wallace DC (eds) *Mitochondrial DNA in human pathology*. Raven Press, New York, NY.
- Sambrook J and Russell DW (2001) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.
- Sbisà E, Tanzariello F, Reyes A, Pesole G and Saccone C (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205:125-140.
- Shedlock AM, Parker JD, Crispin DA, Pietsch TW and Burner GC (1992) Evolution of the salmonid mitochondrial control region. *Mol Phylogenet Evol* 1:179-192.
- Sivasundar A, Bermingham E and Orti G (2001) Population structure and biogeography of migratory freshwater fishes (Prochilodus: Characiformes) in major South American rivers. *Mol Ecol* 10:407-417.
- Thompson JD, Higgins DG and Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-4680.
- Was A and Wenne R (2002) Genetic differentiation in hatchery and wild sea trout (*Salmo trutta*) in the Southern Baltic at microsatellite loci. *Aquaculture* 204:493-506.
- Yoon JM and Park HY (2002) Genetic similarity in the cultured and wild Crucian carp (*Carassius carassius*) estimated with random amplified polymorphic DNA. *Asian-Australasian J Animal Sci* 15:470-476.

Associate Editor: Yatiyo Yonenaga-Yassuda