© 2014 Blackwell Verlag GmbH

ISSN 0175-8659





Received: June 8, 2013 Accepted: December 9, 2013 doi: 10.1111/jai.12423

Sequence variability at the mitochondrial ND1, ND6, cyt b and D-loop segments in tench (*Tinca tinca* L.)

By R. Lo Presti¹, K. Kohlmann², P. Kersten², C. Lisa¹ and L. Di Stasio¹

¹Department of Agriculture, Forest and Food Sciences, University of Turin, Grugliasco, Italy; ²Department of Ecophysiology and Aquaculture, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

Summary

The authors' previous PCR-RFLP studies on ND1, ND6, cyt b and D-loop segments of mtDNA in tench revealed the existence of nine composite haplotypes, named H1-H9. As cyt b is the only mtDNA segment whose variability has been widely studied in tench, the same four segments were sequenced in 29 individuals from 17 populations covering all nine haplotypes, with the aim of providing additional data on tench mtDNA variation. Ninety-six polymorphic sites were identified and all differences were single nucleotide substitutions, except for six indels found in the D-loop. The observed polymorphisms gave origin to five different ND1 sequences, seven ND6 sequences, six cyt b sequences and 17 D-loop sequences. Comparison with data in GenBank revealed that none of the five ND1 sequences were as yet reported, nor the one for cyt b or sixteen for the D-loop; concerning ND6, the data were the first contribution on the variability of this segment. The combination of the haplotypes at the single segments determined 20 composite haplotypes: seven (H1a-H1g) had been classified as H1 and six (H2a-H2f) as H2 with the PCR-RFLP analysis, while for the other haplotypes previously identified (H3-H9) no additional variability was found. The analysis of phylogenetic trees showed that each segment was informative enough to clearly identify the two highly divergent haplogroups already reported in tench, but only the D-loop gave a good resolution of the tree branching because of its high variability. The Median-Joining network, constructed using the composite haplotypes, indicated H1a as the founder node for haplogroup A and the haplotypes H2a-H2f as ancestral with respect to H7–H9. The new data on the variability of mtDNA can contribute to a better understanding of the between and within population diversity in tench.

Introduction

Mitochondrial DNA (mtDNA) proved to be very effective for detecting genetic diversity of animal populations, including several fish species (Briolay et al., 1998), due to its peculiar characteristics, including a high rate of mutation. However, there was limited information on tench mtDNA, whose complete sequence was published by Saitoh et al. (2006).

In the following years, the cyt b gene, together with three nuclear genes, was analysed to study the phylogeography of the tench, leading to the discovery of two geographical clades

Table 1 Origins of sequenced tench individuals

Population	Code	Status	Geographical location	Haplotypes found with the PCR-RFPL procedure				
Badajoz	BAD	Cultured	Badajoz, Spain	H2				
Döllnsee	DÖL	Wild	Döllnsee Lake, Germany	H1				
Felchowsee	FEL	Wild	Felchowsee Lake, Germany	H1, H2, H6, H9				
Königswartha	KÖW	Cultured	Königswartha, Germany	H1				
Golden	GOL	Cultured	Vodňany, Czech Republic	H9				
Marianske Lazne	MAL	Cultured	Marianske Lazne, Czech Republic	H8				
Velke Mezirici	VEM	Cultured	Velke Mezirici, Czech Republic	H1, H2				
Vodňany 1998	VOD	Cultured	Vodňany, Czech Republic	H2				
Valagola	VAL	Wild	Valagola Lake, Italy	H1, H5				
Pianalto	PIA	Cultured	Poirino highland, Italy	H1, H2				
Trasimeno	TRA	Wild	Trasimeno Lake, Italy	H1, H3, H4				
Bolsena	BOL	Wild	Bolsena Lake, Italy	H1, H3				
Alcantara	ALC	Wild	Alcantara River, Italy	H1, H3				
Hungary	HUN	Cultured	Hungary	H2				
Romania	ROM	Cultured	Romania	H1, H2				
Turkey	TUR	Wild	Sapanca Lake, Turkey	H2, H7				
China	CHI	Cultured	Wuhan, China	H2				

R. L. Presti et al.

Table 2 Polymorphic sites and resulting composite haplotypes of tench mtDNA. a) coding segments; b) D-loop

	ND	1															ND	6					
(a) Base position	2 9 0 7	2 9 1 2	2 9 1 6	2 9 6 0	3 0 1 7	3 1 5 8	3 2 4 8	3 3 2 3	3 4 9 1	3 6 0 5	3 7 0 1	3 7 2 5	3 7 3 7	3 7 8 3	3 7 8 8	38 42	1 3 8 1 7	1 3 8 2 3	1 3 8 3 2	1 3 8 7 0	1 3 8 8 6	1 3 8 8 9	1 3 9 9 8
RefSeq H1a H1b H1c H1d H1e H1f H1g H2a H2b H2c H2d H2c H2d H2f H3 H4 H5 H6 H7 H8	G	T	G	A	A	G	G	T	G	G	A	C	A	C	G	A	G	G	T	G	A	C	T
		D-1	oop																				
(b) Base position		1 5 6 8 8	1 5 6 8 9	1 5 7 0 5	1 5 7 6 5	1 5 7 8 4	1 5 8 0 9	1 5 8 1 3	1 5 8 1 4	1 5 8 2 2	1 5 8 3 9	1 5 8 5 1	1 5 8 5 8	1 5 8 7 4	1 5 8 7 9	1 5 9 1 1	1 5 9 4 8	1 6 0 7 1	1 6 0 8 5	1 6 1 3 5	1 6 1 9 4	1 6 2 0 4	1 6 2 1 4
RefSeq H1a H1b H1c H1d H1e H1f H1g H2a H2b H2c H2d H2b H2c H2d H2d H2e H2f H3 H4 H5 H6 H7 H8		G	A	C	C	C	A	A	G A A A A A A A A A A A A A A A A A A A	A	A		T	C	A	A	A	A	T	A	T	A	A

(Eastern and Western), possibly developed in response to recurrent isolation in glacial refugia during the Pleistocene (Lajbner et al., 2007). Within the Eastern phylogroup, the analysis of cyt *b* also allowed the identification of

populations distinct from the major Eastern clade in the Anzalee lagoon of the Caspian Sea in Iran and in the Iskar River of the Danube River drainage in Bulgaria (Lajbner et al., 2011).

						cyth)																	—
1 4 0 1 3	1 4 0 7 8	1 4 0 8 7	1 4 2 1 0	1 4 2 2 5	1 4 2 7 6	14 49 5	1 4 5 4 6	1 4 7 2 6	1 4 8 3 4	1 4 9 4 8	1 4 9 6 1	1 4 9 6 3	1 5 0 3 2	1 5 1 1 0	1 5 2 3 4	1 5 2 5 4	1 5 3 0 8	1 5 3 8 3	1 5 3 9 0	1 5 4 1 0	1 5 4 4 6	1 5 4 6 2	1 5 4 8 8	15 48 9
G		G	G	T	A	A	A	A	C	G	A	A	T	T	C	A	C	A	G	T	T	C	T	A
1 6 2 2 2 2	1 6 2 3 8	1 6 2 5 5	1 6 2 5 8	1 6 2 6 5	1 6 2 7 8	1 6 2 9	1 6 3 1 9	1 6 3 2 7	1 6 3 2 8	1 6 3 3 1	1 6 3 3 6	1 6 3 4 6	1 6 3 4 7	1 6 3 5 4	1 6 3 5 5	1 6 4 6 8	1 6 5 0 5	1 6 5 0 6	1 6 5 2 6	1 6 5 3 2	1 6 5 3 9	1 6 5 6 6	1 6 5 7 4	1 6 5 7 7
A		C	C	G A A A A A A A A A A A A A A A A A A A	G	C	T	A	C	C	G			C		G	A	T	A		T	C	G	A

PCR-RFLP analysis of ND1, ND6, cyt *b* and D-loop segments was performed to investigate the tench genetic diversity within and between populations. The study of 19 populations detected ten polymorphic sites originating nine haplotypes, and revealed considerable haplotype and

nucleotide diversity in nine of the examined populations (Lo Presti et al., 2010, 2012).

As cyt *b* is the only mtDNA segment intensively studied in tench, with 21 sequences already deposited in GenBank (EU856058, AJ555552, Y10451, HM560230, HM167941-57),

R. L. Presti et al.

the four segments previously studied by PCR-RFLP were sequenced with the aim of providing additional data on mtDNA variation in tench.

Materials and methods

Samples

Based on the results of the PCR-RFLP analyses of tench mtDNA conducted by Lo Presti et al. (2012), 29 tench individuals belonging to 17 populations and representing all observed populations by composite haplotype combinations (Table 1) were selected for sequencing the ND1, ND6, cyt *b* and D-loop segments whenever possible (i.e. more than one individual available) in a random manner. Total genomic DNA was extracted from muscle or fin using the NucleoSpin Tissue kit (Macheray-Nagel, Düren, Germany).

Sequencing

The amplification of the four segments was performed using the primers described in Lo Presti et al. (2010), however, due to their length, ND1, cyt *b* and D-loop segments were split into two overlapping fragments, designing internal primers with the PRIMER3, v. 0.4.0 software (http://bioinfo.ut.ee/primer3/) based on reference sequences derived from the complete tench mitochondrial genome (GenBank accession n. NC008648):

ND1 - fragment 1: F: cccagttcatgctaaacactt, R: tagggtatatc ccccggaaa; fragment 2: F: cgagcagtagcccaaacaat, R: aaagtgg tccctaggcatt

Cyt b - fragment 1: F: aacaataatggcaagcctacga, R: ggggtg gaaggagattttgt; fragment 2: F: cgattcttcgcattccactt, R: gctcatt tcaatgctttattttcc

D-loop – fragment 1: F: cgcccagaaaaaggagatt, R: atgcaaaa tgaaaggcaacc; fragment 2: F: ggccctttaatgaattattacttgc, R: ttggactttttagcattaagaaattg.

The PCR reaction mixes for amplification of ND1 contained 5.0 µl of 10x PCR buffer (MBI-Fermentas), 4.0 µl of 25 mm MgCl2, 4.0 μ l of 1.25 mm dNTPs, 1.0 μ l of each primer (10 pmol μl^{-1}), 2.0 μl template DNA, 0.1 μl of TaqDNA-polymerase (5 units μl^{-1} ; MBI-Fermentas) and sterile water up to a final volume of 50.0 μ l and those for amplification of ND6, cyt b and D-loop were composed of 5.0 μ l of 10× PCR buffer (MBI-Fermentas), 2.0 μ l BSA $(20 \mu g \mu l^{-1})$, 5.0 μl of 25 mm MgCl2, 2.5 μl of 1.25 mm dNTPs, 1.0 μ l of each primer (10 pmol μ l⁻¹), 1.0 μ l template DNA, 0.2 μ l of Taq DNA-polymerase (5 units μ l⁻¹; MBI-Fermentas) and sterile water up to a final volume of 50.0 μl. The hot start PCR program for amplification of ND1 consisted of an initial denaturation at 94°C for 3 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 55°C for 30 s, extension at 72°C for 1 min, and a final extension at 72°C for 10 min; that for amplification of ND6, cyt b and D-loop consisted of an initial denaturation at 95°C for 1 min followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 59°C for 30 s, extension at 72°C for 1 min, and a final extension at 72°C for 10 min.

All PCR products were purified using the peqGOLD Cycle-Pure Kit (Peqlab Biotechnologie), and DNA concentrations were measured with a BioPhotometer (Eppendorf). Cycle sequencing was performed using the CEQ DTCS - Quick Start Kit (Beckman Coulter) according to manufacturer instructions. Forward and reverse sequences of all PCR products were recorded on a CEQ 8000 (Beckman Coulter), and aligned and edited manually using the Genetic Analysis System v.7.0, CEQuence Investigator module (Beckman Coulter) and the reference sequences mentioned above. With the ND1, cyt *b* and D-loop the two overlapping sequence fragments were assembled manually into one sequence using the MEGA4 software (Tamura et al., 2007).

Data analysis

The 29 composite sequences and the reference sequence (NC008648) were aligned using the Clustal alignment editor in MEGA4 software (Tamura et al., 2007) and compared to identify the different haplotypes. The same software was used to align the sequences obtained for the single segments with the sequences already available in GenBank (1 for ND1, 21 for cyt *b* and 4 for D-loop). The MEGA4 software was also used to construct Neighbor-Joining trees for each segment separately, with transition/transversion ratios estimated from the data. Bootstrap values were based on 1000 replicates.

Network analysis was performed using the Median-Joining network method implemented in the NETWORK 4.6 software (http://fluxus-engineering.com). Considering that frequently changing sites are less valuable for network construction, the weight for indels, which are less likely to occur, was doubled, as suggested in the user guide.

Results

The lengths of the sequenced tench mtDNA segments were 1019 bp for ND1, 576 bp for ND6, 1146 bp for cyt b, and 996-1001 bp for D-loop, resulting in a total of 3743 bp determined, representing about 22.5% of the total mtDNA length. The analysis of the base composition revealed a general low G content in all of the examined segments (mean 13.3%), compared to T (26.6%), C (26.1%), and especially A (33.8%). Similar results had been reported for cyt b in *Cyprinidae* (Briolay et al., 1998).

The analysis of the obtained sequences detected 96 polymorphic sites compared to the reference sequence (Table 2), corresponding to an average of 2.6% of the total determined

Table 3
Type and number of differences from tench mtDNA reference sequence (NC008648)

	ND1	ND6	Cyt b	D-loop
Transition Transversion Indel Percentage of polymorphic sites	14 2 0 1.6	13 1 0 2.4	16 3 0 1.7	31 10 6 4.7

Table 4 Non-synonymous base substitutions in the three coding tench mtDNA segments

Segment	SNP position	AA change
ND1	2907 G>A	Ala/Thr
	2916 G>A	Val/Ile
	3787 C>A	Leu/Met
ND6	13998 T>C	Asp/Gly
	14013 G>A	Ala/Val
	14013 G>C	Ala/Gly
Cyt b	14961 A>G, 14963 A>C	Thr/Ala
•	15390 G>A	Val/Met
	15489 A>G	Thr/Ala

Table 5 Occurrence of composite mtDNA haplotypes in examined tench populations

Haplotype	Population
H1a	DÖL, KÖW, VAL
Hlb	FEL
H1c	ROM
H1d	VEM
Hle	PIA
H1f	ALC, BOL
Hlg	TRA
H2a	FEL, HUN, VOD
H2b	BAD
H2c	CHI
H2d	TUR
H2e	ROM
H2f	PIA, VEM
H3	ALC, BOL, TRA
H4	TRA
H5	VAL
H6	FEL
H7	TUR
H8	MAL
H9	FEL, GOL

nucleotides. As expected, most of the observed differences were nucleotide substitutions (94%), mainly transitions (82%) (Table 3). One position (14013 in ND6) displayed a double variation: the transition G/A and the transversion G/C. In the D-loop segment indels were also observed. Considering the number of polymorphic sites in relation to the segment length, the highest degree of polymorphism was found in the D-loop. Non-synonymous nucleotide substitutions induced three amino acid changes in each of the coding sequences (Table 4).

The observed polymorphisms gave origin to five different ND1 sequences (deposited in GenBank under accession numbers JX974526-JX974530), seven ND6 sequences (Accession no. JX974548-JX974554), six cyt b sequences (Accession no. JX974520-JX974525) and 17 D-loop sequences (Accession no. JX974531-JX974547). The comparison with data in GenBank, at least for the nucleotides in common, revealed that five of the six cyt b sequences were already described: JX974520 corresponded to HM167950; JX974522 HM167951; JX974523 to to AJ555552. HM560230 and HM167941; JX974524 to HM167943; JX974525 to HM167946. As for D-loop, JX974541 was the same as DQ296148.

The combination of the haplotypes at the single segments allowed the identification of 20 distinct composite haplotypes (Table 2). Seven (H1a-H1g) had been classified as H1 and six (H2a-H2f) as H2 with the PCR-RFLP analysis, while for the other previously identified (H3-H9) haplotypes no additional sequence variability was found. The occurrence of haplotypes in the examined populations is reported in Table 5.

The phylogenetic trees constructed to verify the discriminatory power of the four segments showed that each of them was informative enough to clearly identify (bootstrap value of 100%) the two highly divergent haplogroups already reported in tench (Lajbner et al., 2007, 2011; Kohlmann et al., 2010; Kocour and Kohlmann, 2011; Lo Presti et al.,

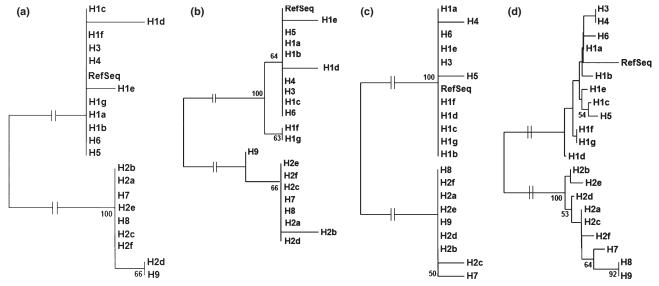


Fig. 1. NJ tree constructed with data of the single segments: a) ND1, b) ND6, c) cyt b, d) D-loop.

R. L. Presti et al.

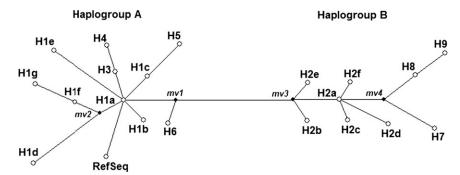


Fig. 2. Median-Joining network of the tench mtDNA composite haplotypes. Median vectors (mv) represent hypothesized sequences required to connect existing sequences.

2012). However, only the D-loop gave a good resolution of the tree branching, while the tree constructed with the data of the cyt *b* was the most poorly resolved (Fig. 1). The Median-Joining network (Fig. 2) showed that all the haplotypes included in the haplogroup A (or Western phylogroup, according to Lajbner et al., 2007), except for H6, clustered around H1a, which could represent the founder node. The results for the haplogroup B (or Eastern phylogroup, according to Lajbner et al., 2007) indicated the haplotypes H2a-H2f as ancestral with respect to H7-H9.

Discussion

The sequencing of the four mtDNA segments of 29 tench individuals provided additional information on mtDNA variability in tench. In fact, the data on ND6 were the first contribution on this segment and new haplotypes were identified for the others.

The D-loop was the most variable, for both the percentage of polymorphic sites, which was even twice that of the other segments, and for the type of variation, which included a higher percentage of transversions and the presence of indels, absent in the coding sequences. The higher quantitative and qualitative variability observed in the D-loop compared to the coding sequences might depend on the absence of purifying selection, which has been shown to act upon the proteincoding regions in mtDNA genome of several species (Ho et al., 2011; Soares et al., 2013). In rainbow trout (Oncorhynchus mykiss) indels were also found in coding regions, where they induced frame shifts affecting the amino acid sequences (Brown et al., 2006). The present data on tench revealed for each coding sequence base differences resulting in three amino acid substitutions. As some of the changes involved amino acids with different structure, studies on their possible physiological effects might deserve attention.

The high degree of polymorphism observed in the tench D-loop is quite peculiar: in other fish species the D-loop variability was lower than that observed for coding sequences (Apostolidis et al., 1997); in common carp (*Cyprinus carpio*) a low variability of the control region, with 1.9% of polymorphic sites, was also reported (Thai et al., 2004). The short history of domestication and/or lower intensities of artificial selection of tench compared to the above mentioned species could have contributed to the maintenance of the high level of variation.

Apart from being the most variable, the D-loop was also the most informative; this segment alone was able to discriminate all but three existing haplotypes that were identified for variations at ND1 and cyt *b*. This situation was reflected by the different resolution of the phylogenetic trees obtained with the data of the single segments.

However, much more variability for cyt *b* has been reported by Lajbner et al. (2011) in a wider study, so that additional data on the other segments would be needed to test the D-loop superiority.

More in general, these preliminary results, including also less investigated mtDNA segments, underline that the informativeness of different regions depends not on their length or percentage of polymorphism, but on the presence of variants specific for a given haplotype. Therefore, the most appropriate segments to be studied have to be selected in a species-specific manner and considering the purpose of the study.

The present sequencing analysis revealed that the haplotyes H1 and H2 detected by PCR-RFLP are very heterogeneous, while all H3 to H9 haplotypes did not show any additional variability. The Median-Joining network, illustrating the evolutionary branching of the observed haplotypes and the potential connecting nodes, indicated for the haplogroup A a star-like phylogeny typical of genetic lines that underwent a rapid expansion perhaps following the colonization of new areas. The limited variation between the haplotypes derived from H1a (one or few substitutions) could suggest a recent differentiation. The haplogroup B showed a branched shape and included haplotypes with a much wider geographical distribution, from Spain to China. However, additional individuals and populations clearly need to be sequenced for the four segments to draw a reliable picture of the composite haplotype relationships and distribution in tench.

In conclusion, the new data on the variability of mtDNA in tench provided by this study add further information for a better understanding of the between and within population diversity. In this context, the results on ND6, ND1 and D-loop segments are of special interest, due to the limited data previously available. The high informativeness highlighted for the D-loop segment underlines the need for further investigation to exploit the potential of mtDNA markers for population studies in tench.

References

Apostolidis, A. P.; Triantaphyllidis, C.; Kouvatsi, A.; Economidis, P. S., 1997: Mitochondrial DNA sequence variation and

- phylogeography among *Salmo trutta* L. (Greek brown trout) populations. Mol. Ecol. **6**, 531–542.
- Briolay, J.; Galtier, N.; Brito, R. M.; Bouvet, Y., 1998: Molecular phylogeny of Cyprinidae inferred from Cytochrome b DNA sequences. Mol. Phylogenet. Evol. 9, 100–108.
- Brown, K. H.; Drew, R. E.; Weber, L. A.; Thorgaard, G. H., 2006: Intraspecific variation in the rainbow trout mitochondrial DNA genome. Comp. Biochem. Physiol. Part D Genomics Proteomics D1, 219–226.
- Ho, S. Y. W.; Lanfear, R.; Bromham, L.; Phillips, M. J.; Soubrier, J.; Rodrigo, A. G.; Cooper, A., 2011: Time-dependent rates of molecular evolution. Mol. Ecol. 20, 3087–3101.
- Kocour, M.; Kohlmann, K., 2011: Growth hormone gene polymorphisms in tench, *Tinca tinca* L. Aquaculture **310**, 298–304.
- Kohlmann, K.; Kersten, P.; Panicz, R.; Memis, D.; Flajšhans, M., 2010: Genetic variability and differentiation of wild and cultured tench populations inferred from microsatellite loci. Rev. Fish Biol. Fish. 20, 279–288.
- Lajbner, Z.; Linhart, O.; Kotlik, P., 2007: Molecular phylogeography of the tench *Tinca tinca* (Linnaeus, 1758). Proc. 12th Eur. Congr. Ichthyology, Cavtat, Croatia. p. 35 (abstr.).
- Lajbner, Z.; Linhart, O.; Kotlik, P., 2011: Human-aided dispersal has altered but not erased the phylogeography of the tench. Evol. Appl. 4, 545–561.
- Lo Presti, R.; Gasco, L.; Lisa, C.; Zoccarato, I.; Di Stasio, L., 2010: PCR-RFLP analysis of mitochondrial DNA in tench *Tinca tin-ca*. J. Fish Biol. **76**, 401–407.

- Lo Presti, R.; Kohlmann, K.; Kersten, P.; Gasco, L.; Lisa, C.; Di Stasio, L., 2012: Genetic variability in tench (*Tinca tinca* L.) as revealed by PCR-RFLP analysis of mitochondrial DNA. Ital. J. Anim. Sci. 11, 103–108.
- Saitoh, K.; Sado, T.; Mayden, R. L.; Hanzawa, N.; Nakamura, K.; Nishida, M.; Miya, M., 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.
- Soares, P.; Abrantes, D.; Rito, T.; Thomson, N.; Radivojac, P.; Li, B.; Macaulay, V.; Samuels, D. C.; Pereira, L., 2013: Evaluating purifying selection in the mitochondrial DNA of various mammalian species. PLoS ONE 8, e58993.
- Tamura, K.; Dudley, J.; Nei, M.; Kumar, S., 2007: MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
- Thai, B. T.; Burridge, C. P.; Pham, T. A.; Austin, C. M., 2004: Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.). Anim. Genet. 36, 23–28.
- Author's address: Liliana Di Stasio, University of Turin, Department of Agriculture, Forest and Food Sciences, Via Leonardo da Vinci 44, I-10095 Grugliasco, Italy. E-mail: liliana.distasio@unito.it