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

Rossella Lo Presti[1], Klaus Kohlmann[2], Petra Kersten[2], Claudio Lisa[1] and Liliana Di Stasio[1]

[1]Department of Agriculture, Forest and Food Sciences, University of Turin, Grugliasco, Italy

[2]Department of Ecophysiology and Aquaculture, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

Corresponding author's address:
Liliana Di Stasio
University of Turin, Department of Agriculture, Forest and Food Sciences
Via Leonardo da Vinci 44, 10095 Grugliasco, Italy
e-mail: liliana.distasio@unito.it

Summary

The authors' previous PCR-RFLP studies on ND1, ND6, cyt *b* and D-loop segments of mtDNA in tench revealed the existence of 9 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 9 haplotypes, with the aim of providing additional data on tench mtDNA variation. Ninety-six polymorphic sites were identified and all the 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. The comparison with data in GenBank revealed that all the five ND1 sequences were not yet reported, as well as one for cyt *b* and sixteen for 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, thanks to 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, limited information existed on tench mtDNA, whose complete sequence has been published by Saitoh *et al*. in 2006.

In the following years, 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 (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).

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 9 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), 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 population 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), but, 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://frodo.wi.mit.edu/primer3/) based on reference sequences derived from the complete tench mitochondrial genome (GenBank accession n. NC008648):

ND1 - fragment 1: F: cccagttcatgctaaacactt, R: tagggtatatcccccggaaa; fragment 2: F: cgagcagtagcccaaacaat, R: aaagtggtccctaggcatt

Cyt *b* - fragment 1: F: aacaataatggcaagcctacga, R: ggggtggaaggagattttgt; fragment 2: F: cgattcttcgcattccactt, R: gctcatttcaatgctttattttcc

D-loop – fragment 1: F: cgcccagaaaaaggagatt, R: atgcaaaatgaaaggcaacc; fragment 2: F: ggccctttaatgaattattacttgc, R: ttggacttttagcattaagaaattg.

The PCR reaction mixes for amplification of ND1 contained 5.0 µl of 10x PCR buffer (MBI-Fermentas), 4.0 µl of 25 mM $MgCl_2$, 4.0 µl of 1.25 mM dNTPs, 1.0 µl of each primer (10 pmol/µl), 2.0 µl template DNA, 0.1 µl of *Taq* DNA-polymerase (5 units/µl; MBI-Fermentas) and sterile water up to a final volume of 50.0 µl and those for amplification of ND6, cyt *b* and D-loop were composed of 5.0 µl of 10x PCR buffer (MBI-Fermentas), 2.0 µl BSA (20 µg/µl), 5.0 µl of 25 mM $MgCl_2$, 2.5 µl of 1.25 mM dNTPs, 1.0 µl of each primer (10 pmol/µl), 1.0 µl template DNA, 0.2 µl of *Taq* DNA-polymerase (5 units/µl; 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, and 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. In case of 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 1,000 replicates.

Network analysis was performed using the Median-Joining network method implemented in the Network 4.6 software (http://fluxus-engineering.com). To take into account 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 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 also indels were 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 (acc. n. JX974548-JX974554), six cyt *b* sequences (acc. n. JX974520-JX974525) and 17 D-loop sequences (acc. n. 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 to HM167951; JX974523 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 haplotypes previously identified (H3-H9) 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*., 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 protein-coding 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 also was the most informative: this segment alone was able to discriminate all the existing haplotypes but three, 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 superiority of D-loop.

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 seqencing analysis revealed that the haplotyes H1 and H2 detected by PCR-RFLP are very heterogeneous, while all the H3 to H9 haplotypes did not show 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 more investigations 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.

~~Bandelt, H. J.; Lahermo, P.; Richards, M.; Macaulay, V., 2001: Detecting errors in mtDNA data by phylogenetic analysis. Int. J. Legal Med. 115, 64-69~~.

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 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.

~~Kumar, S.; Tamura, K.; Nei, M., 2004: MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. Bioinform. 5, 150-163~~.

Lajbner, Z.; Linhart, O.; Kotlìk P., 2007: Molecular phylogeography of the tench *Tinca tinca* (Linnaeus, 1758). Proc. 12th Eur. Congr. Ichthiology, Cavtat, Croatia, p. 35 (abstr.).

Lajbner, Z.; Linhart, O.; Kotlìk, 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 tinca*. 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., Pererira, L., 2013. Evaluating purifying selection in the motochondrial 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.

List of Tables and Figures.

Table 1. Origin of the 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	Vodñany, Czech Republic	H9
Marianske Lazne	MAL	Cultured	Marianske Lazne, Czech Republic	H8
Velke Mezirici	VEM	Cultured	Velke Mezirici, Czech Republic	H1, H2
Vodñany 1998	VOD	Cultured	Vodñany, 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

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

a)

	ND1																ND6													cytb																		
base position	2907	2912	2916	2960	3017	3158	3248	3323	3491	3605	3701	3725	3737	3783	3788	3842	13817	13823	13832	13870	13886	13889	13998	14013	14078	14087	14210	14225	14276	14495	14546	14726	14834	14948	14961	14963	15032	15110	15234	15254	15308	15383	15390	15410	15446	15462	15488	15489
RefSeq	**G**	**T**	**G**	**A**	**A**	**G**	**G**	**T**	**G**	**G**	**A**	**C**	**A**	**C**	**G**	**A**	**G**	**G**	**T**	**G**	**A**	**C**	**T**	**G**	**A**	**G**	**G**	**T**	**A**	**A**	**A**	**A**	**C**	**G**	**A**	**A**	**T**	**T**	**C**	**A**	**C**	**A**	**G**	**T**	**T**	**C**	**T**	**A**
H1a	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1b	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1c	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1d	A	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1e	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1f	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1g	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H2a	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H2b	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	G	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H2c	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	A	C	.	.	C	G
H2d	.	C	A	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H2e	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H2f	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
H5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H7	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	T	C	G
H8	.	C	.	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	G	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G
H9	.	C	A	G	G	A	A	G	A	A	G	T	G	.	A	.	A	A	C	A	.	T	C	C	.	A	.	.	G	G	.	G	T	C	G	C	C	C	T	G	G	G	.	C	.	.	C	G

b)

base position	15688	15689	15705	15765	15784	15809	15813	15814	15822	15839	15851	15858	15874	15879	15911	15948	16071	16085	16135	16194	16204	16214	16222	16238	16255	16258	16265	16278	16291	16319	16327	16328	16331	16336	16346	16347	16354	16355	16468	16505	16506	16526	16532	16539	16566	16574	16577
	D-loop																																														
RefSeq	**G**	**A**	**C**	**C**	**C**	**A**	**A**	**G**	**A**	**A**	**C**	**T**	**C**	**A**	**A**	**A**	**A**	**T**	**A**	**T**	**A**	**A**	**A**	**A**	**C**	**C**	**G**	**G**	**C**	**T**	**A**	**C**	**C**	**G**	**-**	**-**	**C**	**-**	**G**	**A**	**T**	**A**	**-**	**T**	**C**	**G**	**A**
H1a	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1b	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1c	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1d	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	G	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1e	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.
H1f	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H1g	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	-	-	.	.	.	.	.	.
H2a	A	T	T	T	.	.	G	.	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	.
H2b	A	T	T	T	.	.	G	.	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	.	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	.
H2c	A	T	T	T	.	.	G	.	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	.
H2d	A	T	T	T	.	.	G	A	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	.
H2e	A	T	T	T	.	.	G	.	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	.	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	C	T	A	.
H2f	A	T	T	T	T	.	G	.	G	T	T	.	.	C	G	G	G	G	.	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	.
H3	.	.	.	.	.	T	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H4	.	.	.	.	.	T	.	A	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H5	.	.	.	.	.	.	.	A	.	.	T	.	.	.	.	.	.	.	.	.	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
H6	.	.	.	.	.	.	.	A	.	.	T	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	A	T	.	.	.	.	.	.	.	.	.	.	.
H7	A	T	T	T	.	.	G	.	G	T	T	.	T	C	G	G	G	G	G	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	.	.	T	A	.
H8	A	T	T	T	.	.	G	.	G	T	T	A	.	C	G	G	G	G	G	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	G
H9	A	T	T	T	.	.	G	.	G	T	T	A	.	C	G	G	G	G	G	C	G	.	G	C	T	T	A	A	T	C	G	A	A	A	A	T	T	.	A	.	.	T	T	.	T	A	G

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

	ND1	ND6	Cyt *b*	D-loop
Transition	14	13	16	31
Transversion	2	1	3	10
Indel	0	0	0	6
Percentage of polymorphic sites	1.6	2.4	1.7	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 the composite mtDNA haplotypes in the examined tench populations.

Haplotype	Population
H1a	DÖL, KÖW, VAL
H1b	FEL
H1c	ROM
H1d	VEM
H1e	PIA
H1f	ALC, BOL
H1g	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

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

a)

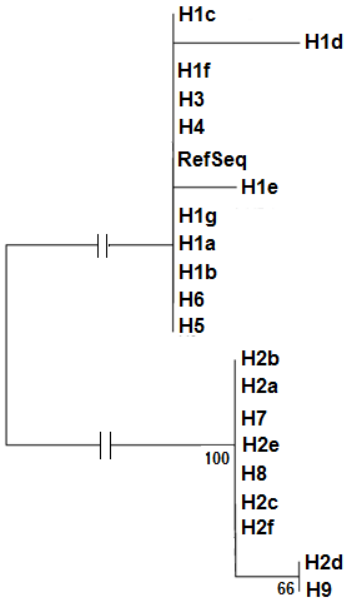

Fig 1

b)

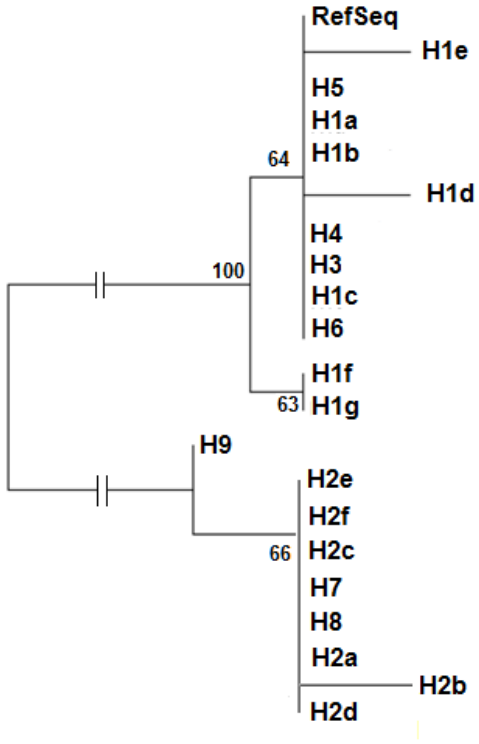
RefSeq
H1e
H5
H1a
H1b
64
H1d
H4
H3
100
H1c
H6
H1f
63
H1g
H9
H2e
H2f
66
H2c
H7
H8
H2a
H2b
H2d

Fig 1

c)

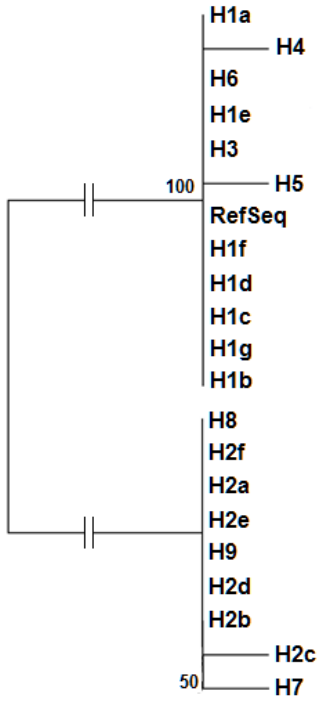
H1a
H4
H6
H1e
H3
100
H5
RefSeq
H1f
H1d
H1c
H1g
H1b
H8
H2f
H2a
H2e
H9
H2d
H2b
H2c
50
H7

Fig 1

d)

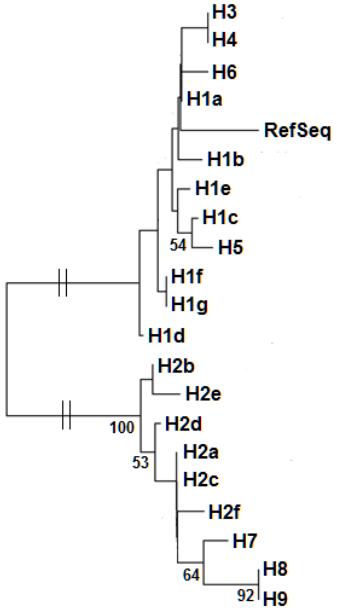
H3
H4
H6
H1a
RefSeq
H1b
H1e
H1c
54
H5
H1f
H1g
H1d
H2b
H2e
100
H2d
53
H2a
H2c
H2f
H7
64
H8
92
H9

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