



Genetic diversity of the cyprinid fish *Capoeta trutta* (Heckel, 1843) populations from Euphrates and Tigris rivers in Turkey based on mtDNA COI sequences

ARIF PARMAKSIZ AND ESRA EKSI

Harran University, Faculty of Science and Art, Department of Biology, Sanliurfa - 631 00, Turkey
e-mail: aprmksiz@gmail.com

ABSTRACT

In this study, genetic diversity of *Capoeta trutta* (Heckel, 1843) populations from Euphrates and Tigris rivers in Turkey was evaluated based on gene sequence analysis of mitochondrial DNA cytochrome c oxidase subunit I (mtDNA COI) locus. Six polymorphic sites and seven haplotypes were detected in 47 samples collected from four populations viz., Adiyaman, Birecik, Bismil and Batman. The mean haplotype diversity (h) and nucleotide diversity (π) were calculated as $h = 0,6420$ and $\pi = 0,00138$ respectively. Pairwise F_{ST} statistics of different populations were found to be negative, low and were insignificant, indicating gene flow. AMOVA analysis showed $F_{ST} = 0.09865$ and $p = 0.00489$, indicating that the populations were isolated. The results of Neutrality tests showed an increase in Adiyaman, Birecik and Bismil populations and a decline in Batman population, all values being statistically insignificant ($p > 0.05$). Three haplotypes determined for mtDNA COI locus in the present study form important data set for genetic diversity of this species.

Keywords: *Capoeta trutta*, Euphrates River, Genetic diversity, mtDNA COI, Tigris River

Introduction

Capoeta trutta (Heckel, 1843) (Family : Cyprinidae) is a fish species having economic importance with wide distribution in Turkey, Iran, Iraq and Syria (Gunduz *et al.*, 2014) which is dominantly thriving in both the Euphrates and Tigris river systems (Geldiay and Balik 2007). The species is locally called as karabalik, karaca and sirazin in Turkey.

Ozdemir and Kabukcu (1982) studied the length-weight relationship, condition factor and the reproductive biology of populations in Keban Dam Lake. Comparative age determination from operculum, otoliths and scales of populations in Keban Dam Lake were conducted by Ozdemir and Sen (1983). Unlu (1991) investigated on the growth rates, reproductive age, reproductive cycle and egg production of populations in the Tigris River while Bozkurt (1998) studied age, height-weight relationships, condition factor and reproductive characteristics in Ataturk Dam Lake. Other studies include, investigations on the effect of water quality on ovary of *C. trutta* in Keban Dam Lake (Ural, 2004); characteristics of growth and reproduction in Karakaya Dam Lake (Kalkan, 2008); reproductive biology and histological changes in the gonads in Ataturk Dam Lake (Oymak *et al.*, 2008); determination of fish freshness in Ataturk Dam Lake (Sarac, 2011); comparison of the reproductive period in populations of Keban, Karakaya

and Ataturk Dam Lakes (Dusukcan and Calta, 2012) and studies on the reproductive biology in the Tigris River (Bilici *et al.*, 2016).

For the management and conservation of fish species with economic importance, it is important to have in depth understanding about the genetic diversity and population structure (Ward 2000; Ortega-Villaizan Romo *et al.*, 2006). However, there is no information available on the genetic diversity of *C. trutta* living in the Tigris and Euphrates river systems. The goal of the present study was to investigate the genetic diversity of *C. trutta* populations in the Euphrates and Tigris rivers in Turkey by gene sequence analysis of mitochondrial DNA cytochrome c oxidase subunit I (mtDNA COI) locus. Because mtDNA is maternally inherited and has high mutation ratio, it has been considered as an ideal marker for population genetics studies (Avise *et al.*, 1987). Besides distinguishing similar species, mtDNA COI locus is one of the most used molecular markers for determination of the differences between populations of the same species. (Croos and Palsson 2010; Keskin and Atar, 2012).

Materials and methods

Sample collection and DNA extraction

A total of 47 samples from 4 populations comprising 29 individuals from 2 populations in the Euphrates

River (Adiyaman:19, Birecik:10), 18 individuals from 2 populations in the Tigris River (Bismil:11, Batman:7), were collected and transferred in ice to the laboratory. Basal muscle tissues of pectoral or dorsal fin were dissected from each sample, immediately placed in micro-centrifuge tubes with 95% ethanol and stored at 4°C until DNA isolation. Tissue samples were digested for 3h at 55°C in 200 µl DNA extraction buffer containing 20 µl proteinase K. Genomic DNA was extracted following the protocol of GeneJET Genomic DNA Purification Kit (Thermo Scientific) with minor modifications.

PCR amplification and sequencing

Primers used for amplification of mtDNA COI locus in study by Darabi (2014) *viz.*, COI-625F:5'-TC AACCAACCACAAAGACATTGGCAC-3' and COI-625 R:5'-GACTTCTGGGTGGCCAAA-GAATCA-3' were used in the present study.

The PCR amplification was carried out using BIO-RAD T100™ Thermal Cycler device under the following conditions: initial denaturation for 3 min at 95°C, followed by 35 cycles of denaturation at 95°C for 30 sec; annealing at 62°C for 30 sec and extension at 72°C for 45 sec followed by a final extension at 72°C for 10 min. All PCR reactions were performed at a total volume of 25 µl containing 0.5 mM of each primer, 0.2 mM of each dNTP, 1x PCR buffer, 2.5mM MgCl₂, 1 unit taq polymerase and 90 ng of template DNA. A total 3 µl of each PCR product was checked by agarose (1.5%) gel electrophoresis using a 100 bp DNA ladder (Gene ruler Plus DNA ladder). After recovery and purification steps, the PCR products were sequenced on an 3500 XL Genetic Analyser (Thermo Fisher Scientific).

Sequence analysis

Data of mtDNA sequences were analysed using Chromas Pro software (<http://www.technelysium.com.au/ChromasPro.html>) and converted into the FASTA format. FASTA file of sequences was edited and aligned using ClustalW Multiple Alignment Tool and trimmed. Data file was prepared for further statistical analysis using BioEdit software version 7.2.5 (Hall, 1997-2013). mtDNA COI 625 sequence was evaluated in terms of number of polymorphic sites and haplotypes, haplotype diversity (Hd) and nucleotide diversity (π) for each population using DnaSP 5.10.01 (Rozas *et al.*, 2003). The phylogenetic relationships between haplotypes were identified by Network version 4.6 software (<http://www.fluxus-engineering.com>; Rohlf, 1999). Tajima D and Fu's statistics were calculated using the software ARLEQUIN 3.5 (with 1000 permutations) in order to analyse if populations were subjected to any selection in the past (Tajima 1989; Fu 1997).

Results and discussion

Genetic variation

An average of 600 bp fragments of mtDNA COI 625 locus from 47 *C. trutta* specimens was sequenced; totally 6 variable sites and 7 haplotypes were detected. Nucleotide variations of this region are shown in Table 1. Haplotype diversity (h), nucleotide diversity (π) and the neutrality tests for each population are given in Table 2.

Table 1. Haplotypes and nucleotide variations of mtDNA COI 625

Haplotypes	151	293	445	451	464	526
H1	C	G	C	C	T	C
H2	T	T
H3	T
H4	T	.	T	.	.	.
H5	T	.	.	.	C	.
H6	T	.	.	T	.	.
H7	T	A

Table 2. Genetic diversity and neutrality tests of *C. trutta* populations (n = number of individuals, Nh: number of haplotypes, h: haplotype diversity, π : nucleotide diversity)

Locality	n	Nh	h	π	Tajima's D	Fu's Fs
Adiyaman	19	6	0,708	0,00150	-1.17758	-2.720
Birecik	10	3	0,378	0,00070	-1.40085	-1.164
Bismil	11	5	0,782	0,00178	-0.83418	-1.929
Batman	7	2	0,476	0,00080	0.55902	0.589
Total	47	7	0,6420	0,00138	-1.08945	-2.946

Mean haplotype diversity (h) was calculated as 0.6420 and mean nucleotide diversity (π) as 0.00138. Bismil population was found to have the highest haplotype (h = 0.782) and nucleotide diversity (π = 0.00178) while the lowest haplotype (h = 0.378) and nucleotide diversity (π = 0.00070) were observed in Birecik population (Table 2).

From seven haplotypes detected in total; six haplotypes (H1, H2, H3, H4, H5 and H7) were observed in samples from Adiyaman, three (H1, H2 and H3) in samples from Birecik, five (H2, H3, H4, H6 and H7) in samples from Bismil and two (H3 and H4) in samples from Batman. H3 was found to be the common haplotype with 27 individuals in all localities. H6 haplotype was represented with only one individual in Bismil. H5 haplotype was detected in only two individuals from Adiyaman while H1 haplotype was found in three individuals from Adiyaman and Birecik localities.

Population genetic structure

Genetic structure of *C. trutta* populations was analysed using pairwise F_{ST} values and molecular variance analysis (AMOVA). Pairwise F_{ST} values belonging to 4 populations ranged from -0.08700 to 0.10547 (Table 4).

Table 3. Relative haplotype frequencies between four populations of *C. trutta*

Haplotype	Adiyaman (n = 19)	Birecik (n = 10)	Bismil (n = 11)	Batman (n = 7)
H1	0,105263	0,1	0	0
H2	0,157895	0,1	0,090909	0
H3	0,526316	0,8	0,363636	0,714286
H4	0,052632	0	0,363636	0,285714
H5	0,105263	0	0	0
H6	0	0	0,090909	0
H7	0,052632	0	0,090909	0

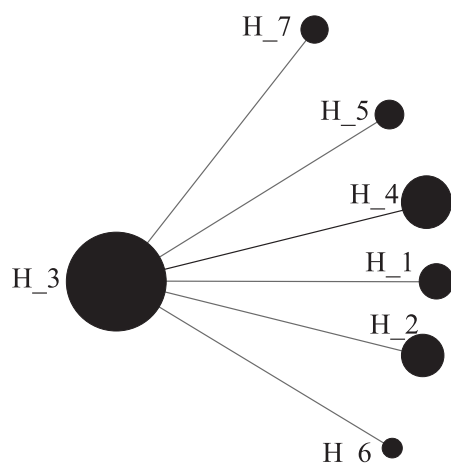
Table 4. Fixation index (F_{ST}) between localities

Locality	Adiyaman	Birecik	Batman	Bismil
Adiyaman	0.00000			
Birecik	-0.05126	0.00000		
Batman	0.03369	0.10287	0.00000	
Bismil	0.06422	0.10547	-0.08700	0.00000

AMOVA analysis conducted based upon the haplotype frequencies indicated 90.14% genetic variation within localities and 9.86% genetic variation among localities. The mean fixation index was calculated as 0.09865 ($p < 0.05$).

Seven haplotypes were identified in Median-Joining Network of haplotypes created for the 47 samples of *C. trutta* analysed in the present study. Resulting network showed existence of a central haplotypes (H3) which indicates an evolutionary link. This was found to be the most prevalent haplotype in all localities. It is also possible to speculate that H3 might be linked to all other haplotypes (Fig. 1).

Tajima's D (1996) and Fu's FS (Fu, 1997) belonging to neutrality tests were evaluated which revealed presence of selection for an allele in the population (Tajima 1989; Fu, 1997). Tajima's D value had positive value (0.55902) in Batman population and negative in other populations

Fig. 1. Haplotype network of *C. trutta* Neutrality tests

while the sum total was negative (-1.08945) and found to be statistically insignificant ($p > 0.05$). In Fu's F_s tests, Batman populations had positive values (0.589), the values were, negative for other populations with the sum total being negative (-2.946) and therefore found to be statistically insignificant ($p > 0.05$).

In this study, genetic diversity of populations was investigated by sequence analysis of mtDNA COI 625 locus from 47 samples of *C. trutta* in total, including 29 individuals (Adiyaman:19, Birecik:10) from 2 populations in the Euphrates River, 18 individuals (Bismil:11, Batman:7) from 2 populations in the Tigris River. We determined six polymorphic sites and seven haplotypes on that locus. Six haplotypes (H1, H2, H3, H4, H5 and H7) were observed in samples from Adiyaman, three (H1, H2 and H3) from Birecik, five (H2, H3, H4, H6 and H7) from Bismil and two (H3 and H4) from Batman. H3 was found to be the common haplotype with 27 individuals in all localities and therefore we can assume that this haplotype could be an ancestral one as it is the most abundant one. In addition, H1 haplotype was found in only three individuals from Adiyaman (2) and Birecik (1) localities and two individuals with H5 haplotype was seen only in Adiyaman which indicates that H1 and H5 haplotypes are indigenous to the Euphrates River. H6 haplotype was represented with only 1 individual from Bismil.

Related to this locus, there are five different haplotypes (H1, H2, H3, H5, H8) in NCBI. Three new haplotypes (H4, H6, H7) have been identified in the present study. In all the haplotypes, H3 is most common haplotype which implied that it is the ancestral haplotype in *C. trutta* (Table 5).

Mean haplotype diversity and nucleotide diversity were calculated as: $h=0.6420$ and $\pi = 0.00138$ respectively. Bismil population was found to be the highest in both haplotype diversity ($h=0.782$) and nucleotide diversity ($\pi=0.00178$) amongst the populations studied. On the other hand, Birecik population had the lowest haplotype diversity ($h=0.378$) and nucleotide diversity

Table 5. Total haplotypes of *C. trutta* in the present study and GenBank

Haplotype	GenBank data
H1	Present study and KU892584.1
H2	Present study and KM590422.1
H3	Present study, KM590421.1, KU312383.1, KU312382.1, KU312351.1, KU899129.1, KU899128.1, KU899125.1, KU899118.1 and KU948086.1
H4	Present study
H5	Present study and KU899116.1
H6	Present study
H7	Present study
H8	KU312386.1

($\pi = 0.00070$). In the study by Darabi (2014), the same primers (mtDNA COI 625) were used for *Barbus sharpeyi* species and they determined nucleotide diversity of three different populations as: 0,0286; 0,0785 and 0,0072 respectively. Compared to these, the values in our study were found to be lower. The diversity of nucleotide is a precise method used for genetic analysis of the populations (Nei and Li, 1979). Genetic diversity can be affected by life period in time, character, population size and environmental conditions of the populations (Ne, 1987; Avis, 2000).

In the current study, the haplotype ratios in populations of the Euphrates and Tigris rivers were high but the values of nucleotide diversity were low. This is an indicator that the species had a population growth following a decline of population size in the past (Saraswat *et al.*, 2013). Low population diversity and genetic diversity are influenced by many factors, such as habitation, anthropogenic activity, founder effects and bottleneck effects (Fennando *et al.*, 2000; Ma *et al.*, 2010). These factors would have caused harmful effects on the population structure by decreasing genetic diversity and population levels (Saraswat *et al.*, 2013).

In Median Joining Network analysis, we saw that H3 haplotype was in the center of network and dominant and also all of the haplotypes consisted of H3 (Fig. 1).

Pairwise F_{ST} values between populations of Adiyaman-Birecik (-0.05126) and Batman-Bismil (-0.08700) were negative and were found to be lower for other locations. Because the ratio of similarity in populations living in the same river is higher, we can assume this was predictable. Adiyaman and Birecik populations live in the Euphrates River; Bismil and Batman populations dwell in the Tigris River. The value between Birecik and Bismil populations (0.10547), which is higher than the others, also resulted from the fact that they live in different river systems, which indicates there is an isolation between these two populations.

AMOVA analysis conducted based on the haplotype frequencies revealed that the ratios among localities and within localities genetic variation was 9.86 and 90.14%, respectively and mean fixation index was calculated as 0.09865 ($p < 0.05$). Darabi (2014), determined in their study among localities and within localities genetic variation and fixation index for *Barbus sharpeyi* to be 85.46, 14.54% and 0.14535 respectively. Our results are comparable with the results of the above study. Low genetic variation among localities is an indicator of the fact that there is a high gene flow between populations or these populations were the last ones which were isolated.

Tajima D and Fu's tests results were negative in populations of Adiyaman, Birecik and Bismil and positive in Batman population, and not all of the values were statistically significant ($p > 0.05$). Negative values in populations of Adiyaman, Birecik and Bismil are considered to be associated with the growth of population while positive value in Batman population would have resulted from the decline in population size.

Factors like construction of dams, excessive fishing, alien fish grafting and pollution which have played a major role in the destruction of the freshwater fish habitat, are thought to cause reduction of genetic diversity. Haplotypes determined for mtDNA COI 625 locus are new results for the literature and it would be even more explanatory in terms of population genetics of this species to use more populations and different genetic markers in further studies. The results of the study are assumed to be useful for planning effective strategies for the conservation and management of fisheries in the two river systems.

Acknowledgements

This study was funded by Harran University Research Fund (Project No: 16041). The authors report no conflicts of interest.

References

- Avis, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. and Saunders, N. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.*, 18: 489-522.
- Avis, J. C. 2000. *Phylogeography*. Harvard University Press, London.
- Bilici, S., Unlu, E., Cicek, T. and Satici, O. 2016. The reproductive biology of *Carasobarbus luteus* and *Capoeta trutta* in the Tigris River, Turkey. *Cybius*, 40(2): 147-153.
- Bozkurt, R. 1998. *Ataturk Baraj Golu'ndeki Acanthobrama marmid (Heckel, 1843), C. Trutta (Heckel, 1843) ve C. luteus (Heckel, 1843)'un biyolojik ozellikleri uzerinde arastirmalar*. Ph. D. Thesis, Harran University, 130 pp.
- Croos, D. M. and Palsson, S. 2010. Mitochondrial DNA variation and population genetic structure of white shrimp *Fenneropenaeus indicus* along the coastal belt of Sri Lanka. *Aquat. Living Resour.*, 23: 315-32. doi: 10.1051/alr/2010027.
- Darabi, A. R., Kashan, N., Fayazi, J., Aminafshar, M. and Chamani, M. 2014. Investigation of phylogenetic relationship among two *Barbus* species (Cyprinidae) populations with mitochondrial DNA using PCR sequencing. *IJBPAS*, 4(2): 302-311. doi: 10.12692/ijb/5.2.41-46

- Dusukcan, M. and Calta, M. 2012. *C. trutta* (Heckel, 1843) 'nınKeban, Karakayave Ataturk Baraj Golleri'ndeki populasyon larını nuremedon emlerininkarsilastirilmasi. *Firat Univ. Fen Bil. Derg.*, 24(2): 57-61.
- Fennando, P., Pfrender, M. E., Encalada, S. E. and Lande, R. 2000. Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. *Heredity*, 84: 362-72.
- Fu, Y. X. 1997. Statistical tests of neutrality against population growth, hitchiking and background selection. *Genetics*, 147: 915-925.
- Geldiay, R. and Balik, S. 2007. Turkiye Tatlisu Balıkları, DersKitabi, *Ege Universitesi Basım Evi.*, 97: 250.
- Gunduz, F., Coban, M. Z., Yuksel, F., Demiroglu, F., Kurtoglu, M. and Yıldız, N. 2014. Uzuncayır Baraj Golu'ndeki (Tunceli) *Capoeta trutta* (Heckel, 1843) 'nın Bazı Populasyon Parametreleri. *Yunus Araştırma Bülteni*, 2: 3-14. doi: 10.17693/yunusae.vi.235393
- Hall, T. A. 1997. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids Symposium Series*, 41: 95-98.
- Hall, T. A. 2013. *BioEdit v 7.2.3*. Biological sequence alignment editor for Win 95/98/NT/2K/XP7.
- Kalkan, E. 2008. Growth and reproduction properties of *Capoeta trutta* (Heckel, 1843) in Karakaya Dam Lake. *Turk. J. Zool.*, 32: 1-10.
- Keskin, E. and Atar, H. H. 2012. Turkiye'nin Akdeniz Kıyısındaki Mavi Yengec (*Callinectes sapidus* Populasyonları Arasındaki Genetik Farklılığın COI Gen Dizileri Kullanılarak Degerlendirilmesi. *J. Fish. Sci.*, 6(2): 125-131.
- Ma, C., Cheng, Q., Zhang, Q., Zhuang, P. and Zhao, Y. 2010. Genetic variation of *Coilia ectenes* (Clupeiformes: Engraulidae) revealed by the complete cytochrome b sequences of mitochondrial DNA. *J. Exp. Mar. Bio. Ecol.*, 385: 14-19.
- Nei, M. and Li, W. H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci.*, 76: 5269-5273.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia University Press, New York.
- Ortega-Villaizan, R. M., Suzuki, S., Nakajima, M. and Taniguchi, N. 2006. Genetic evolution of inter individual relatedness for broodstock management of the rare species barfin flounder *Verasper moseri* using microsatellite DNA markers. *Fish. Sci.*, 72: 33-39. doi: 10.1111/j.1444-2906.2006.01113.x.
- Oymak, S. A., Musa, D. and Unlu, E. 2008. Ataturk baraj Golu'ndeyasayankarabalıkların *C. trutta* (Heckel, 1843) uremebiyolojisi ve gonad larında kihistolojik degisimler. *Istanbul Univ. Su Urun. Derg.*, 23(2): 1-11.
- Ozdemir, N. and Kabukcu, M. A. 1982. Keban Baraj Golu'nde bulunan *C. Trutta* (Heckel, 1843) 'nın boy-agırlık ilişkisi, kondusyon faktoru ve ureme periyodu uzerine arastirmalar. *Firat Univ.Vet. Fak. Derg.*, 7: 139-150.
- Ozdemir, N. and Sen, D. 1983. Keban Baraj Golu'nde bulunan *C. Trutta* (Heckel, 1843) 'nın pul, otolit ve operkulumundan karsilastirmalı yas tayini calısmaları. *Et ve Bal. End. Derg.*, 6(35): 15-22.
- Rohl, A. 1999. *NETWORK: A program package for calculating phylogenetic network*. Version 2.0b. Available from: <http://www.fluxus-engineering.com>.
- Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X. and Rozas, R. 2003. DnaSP DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 19: 2496-2497.
- Sarac, A. 2011. *Ataturk Baraj Golu'nde yasayan C. luteus ve C. trutta' da balık tazeliginin tespiti*. M. Sc. Thesis, Harran University, 62 pp.
- Tajima, F. 1996. The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics*, 143: 1457-1465.
- Ural, M. 2004. *Keban Baraj Golu'nun Kockale ve Pertek bolgelerindeki su kalitesinin C. Trutta (Heckel, 1843) 'ninvaryumuna olan etkisinin arastırılması*. Ph. D. Thesis, Firat University, 52 pp.
- Unlu, E. 1991. Dicle Nehri'nde yasayan *C. Trutta* (Heckel, 1843) 'nın biyolojik ozellikleri uzerinde calısmalar. *Doga Turk. Zool. Derg.*, 15: 22-38.
- Ward, R. D. 2000. Genetics in fisheries management. *Hydrobiologia*, 420: 191-201.