Identification and Molecular Investigation of *Diplostomum* in *Capoeta umbla* Caught from Freshwater Sources, Turkey

Sibel Doğan Barata¹,*, Mustafa Dörücü², Murad Gürses²

¹Fisheries Faculty, Fırat University, 23110, Elazığ, Turkey.
²Faculty of Veterinary Medicine, Balıkesir University, 10145, Balıkesir, Turkey.

How to cite

Abstract
In this study, identification of the cytochrome c oxidase subunit I (COI) gene region in mitochondrial DNA of the *Diplostomum* specimens, which are parasitic in the eyes of *Capoeta umbla* fish naturally distributed in Karakaya Dam Lake, Keban Dam Lake and Lake Hazar, (Elazığ province) was performed. A total of 8 haplotypes were determined using sequence analysis of the mtCOI gene region and phylogenetic trees were constructed by determining the genetic distance based on these haplotypes by the Maximum Likelihood (ML), Neighbor Joining (NJ), Minimum Evolution (ME) and Bayes methods. As a result, the lowest genetic distance was 0,00209 and the highest genetic distance was 0,01264. Phylogenetic analysis revealed a closer relationship between haplotype 3 and haplotype 4. It has been evaluated that all the samples studied belong to *Diplostomum spathaceum* and that the parasite may have been spread by the main hosts between study sites.

Introduction
Industrial and agricultural activities generate a variety of chemical and physical agents that continuously damage the natural environment (Barim et al., 2009; Barim & Karatepe, 2010). Developing and changing technological innovations and rapidly increasing world population leads to a gradual decrease in natural resources. This situation brings with it a balanced diet problem. The limited amount of protein sources of animal origin, which are indispensable for people to eat a balanced diet, has led to a search for new protein sources. Therefore, besides meeting the animal protein demand, the fisheries sector also constitutes an important export item with its contribution to the country's economy (Öge, 1999). However, in order to minimize product loss in the fisheries sector, it is necessary to know the disease-causing factors and to undertake scientific studies (Öge, 1999; Barata & Dörücü, 2014). Parasitic diseases occurring from time to time emerge as an important factor in fish (Aksoy, 1999).

Morphometric and morphological studies are not comprehensive and have been mostly performed on only local populations. Therefore, they are insufficient for the evaluation of populations at a subspecific level (Seçkinozan Şeker, 2018). Today, sufficient knowledge about the distribution, biology and identification of fish parasites has not been reached yet (Dörücü et al., 2008).

The larvae of the *Diplostomum* species cause degeneration in the eye of the fish and form the "Larval Eye Diplostomidosis", known as eye worm disease. *Diplostomum* sp. their metacerceariae primarily settle in the eye and other body parts (fins, skin, etc.). In dead
fish, it is possible to encounter metacercariae in the brain. These metacercariae can cause opacity in the humor, swelling of the cornea (keratoglobus), cloudiness of the lens (parasitic cataract), intraocular pressure increase, exophthalmos, corneal rupture and consequently can blind fish. Fish that are blind can not feed effectively and start to weaken because and they easily caught by birds due to swimming disorders. *Diplod stomum* sp. infection caused by metacercariae, there are bloody foci at the beginning of the fish (Öge, 1999; Markevic, 1951; Bykhouskaya-Pavlouskaya, 1964; Tiğin et al., 1992; Tokşen et al., 1996; Erer 2002; Arda et al., 2005). 

There are studies showing that mt COI sequences give more accurate results than other DNA gene regions in the separation studies of Digenea. Désilets et al. (2013) found 5 different species in their study on *Diplod stomum* species. Blasco-Costa et al. (2014) found 17 species in the yet unknown species diversity study of *Diplod stomum* in Iceland. Faltinková et al. (2014) diagnosed 6 *Diplod stomum* species using molecular methods in their study in Iceland. Locke et al. (2015) found 52 Diplod stomum species using the DNA barcoding method. 

In this study, *Diplod stomum* species found in *Capoeta umbra*, which are naturally distributed in Karakaya Dam Lake, Keban Dam Lake and Hazar Lake within the borders of Elazig province, were identified based on cytochrome c oxidase I (mtCOI) gene region together with morphological data. The aim is to describe the species taxonomically in the light of sequence analysis data. In addition, it has been investigated whether there is a genetic difference between the samples of *Diplod stomum* species collected from different localities, whether environmental factors (different dam lakes or rivers or different regions of the same river) have an effect on the genetic structure of the parasite. Thus, detecting the species and subspecies of the parasite (if any) will allow them to be separated from each other according to the host they live in and to develop appropriate drugs, vaccines and biological substances against the parasite in the fight against parasites.

**Material and Methods**

**Sample Collection**

Fish samples used in the study were carried out on a total of 150 *Capoeta umbra* (Heckel, 1843) caught from Karakaya Dam Lake (50 individuals), Keban Dam Lake (60 individuals) and Hazar Lake (50 individuals) between May-August 2017.

The study was reviewed by the local Ethics Committee (Fırat University Animal Experiments Local Ethics Committee Presidency) who sanctioned the work (Protocol No: 2016/19; Decision No: 69).

The eyes of the fish were cut with the help of forceps and scalpel and the eyepiece was removed and left in physiological saline (0.9%). Then, *Diplod stomum* metacercariae were collected under a stereomicroscope, counted into Eppendorf tubes containing 96% ethyl alcohol and left at -20°C. The parasites were identified morphologically according to Bykhovskaya-Pavlouskaya (1964), Hoffman (1967), Kennedy (1974), Ekingen (1983), Williams & Jones (1994).

**DNA Isolation**

DNA isolation was performed according to the Qiagen DNaseasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA) protocol with fish samples containing mertaserkariae larvae in their eyes.

The concentration and purity analyzes of the total DNA obtained were performed on the spectrophotometer (NanoDrop 2000, Thermoscientific). The samples, whose concentrations were determined by spectrophotometric analysis in adjusting the DNA concentration, were diluted to 200 ng / µl in order to avoid contamination of the main stocks and to standardize the DNA concentrations of all samples.

**PCR Amplification and Sequencing**

The lyophilized primer set (COX-1F 5´CGTTTTRAATTATACGGATCC 3´ and COX-1R 5´AGCATAGTAAATMGCACGC 3´ (Moszczynska et al., 2009)) used in the PCR process was diluted with distilled water up to 10 times the mmol values given by the manufacturer, to obtain 100 pm/µl stock primer. The PCR reaction was carried out as a total volume of 25 µl. In this volume, 10X PCR buffer, 2.5 mM MgCl₂, 1.25 pmol forward and 1.25 pmol reverse primer, 50 µM dNTPs, 0.6 U Taq polymerase, 50 ng purified DNA sample and 15 µl ddH₂O were used.

PCR conditions were adjusted for 30 seconds at 94°C, 30 seconds at 50°C and 60 seconds at 72°C for 35 cycles, followed by 2-minute pre-denaturation with Thermal cycler, followed by 72°C for 10 minutes last elongation step was applied.

After PCR, products were added with ethidium bromide. Electrophoresis was performed in a 2% agarose gel, and after the electrophoresis agarose gel, it was determined that the PCR reaction was working correctly by obtaining the band approximately of 500 bp fragment of the COI gene under UV light.

**Data Analysis**

For each sample to be analyzed in the series, purification was carried out through service procurement. Performing the sequence analysis of the samples was carried out by using specific primers used in the PCR process, through bilateral (forward and reverse) service procurement (UN Laboratory Systems).

DNA sequencing was performed in the Macrogen (Netherlands) laboratory using the ABI 3730XL Sanger sequencing device (Applied Biosystems, Foster City, CA).
and the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and consequently raw DNA sequences was obtained.

In the study, the DNA sequences of the samples were aligned using ClustalW algorithm in BioEdit v. 7.1.11 and the haplotype number and diversity were determined by DnaSP Version 5.10 program (Thompson et al., 1994; Hall, 1999; Librado, and Rozas, 2009). Genetic distance and nucleotide composition were calculated for haplotypes. BLAST (Basic Local Alignment Search Tool) analysis were performed to compare sequences with data available on GenBank. *Tylodelphys* (FJ469596) was used as the outgroup for phylogenetic analyses.

Nucleotide composition was prepared with Bioedit program for all haplotypes obtained as a result of sequence analysis. The most suitable model has been determined in the trees to be created using the MEGA7 program (Kumar et al., 2015). Using these haplotypes, intra-genetic distance value was calculated with MEGA7 program based on Kimura-2-Parameter (K2P) and base change model.

Maximum Parsimony (MP), Maximum Likelihood (ML), and Neigbor Joining (NJ) analyzes were performed using Mega 7 program. The Maximum Likelihood (ML) tree was given as a figure because the analysis results were found to be similar to each other. Bayesian Inference (Bi) method was done in Mr. Bayes program.

In the study MrBayes v. The program 3.2.1 was used (Ronquist, 2012). Gene sequences to be analyzed in MrBayes program were prepared by EMBOSS in Nexus format in Clustal-Omega. Haplotype in prepared nexus format; lset was analyzed using nst = 6 rates = invgamma and mcmc ngen = 500000 samplefreq = 100 printfreq = 100 diagnfreq = 1000 model settings. As a result of 500,000 iterations, a total of 7,502 trees were created to evaluate probabilities. One consensus tree was created from these trees. The revision and display process in the consensus tree was prepared in FigTree v1.4.3 program.

**Results**

In the eyes of a total of 150 fish samples collected *Diplostomum* sp. parasites were screened and fish containing at least 20 metacercariae were included in the study. DNA isolation was obtained by centrifuging at least 20 metacercariae together. DNA isolation was obtained from metacercariae isolated from a total of 53 fish samples, and the gene regions of 49 samples were reproduced by PCR procedures. Purification and bidirectional sequence analyzes from PCR samples were performed with service procurement.

As a result of haplotype analyses, eight haplotypes were determined (Table 1) and, haplotype diversity (Hd) was obtained as 0.6692. Each haplotype sequence determined in the study was compared with GenBank BLAST analysis and it was determined that the samples were compatible with *Diplostomum spathaceum* (Rudolphi, 1819) with 100% ident ratio. While Haplotype 1 and Haplotype 2 differed by 0.2%; haplotype 1; differs from Haplotype 3, Haplotype 5, Haplotype 6, and Haplotype 8 by 0.4%. Again, Haplotype 1; differs by 0.8% from Haplotype 4 and Haplotype 7. The highest genetic distance was found between Haplotype 7 and Haplotype 4 with a value of 0.01264 (Table 2). Accordingly, it was determined that the T base ratio was highest in all haplotypes analyzed. The ratios of A and G bases were very close to each other. C has the least proportion. Thus, A + T ratio was found above 68% in all haplotypes (Table 3).

### Table 1. mtDNA COI gene region haplotype information.

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>DNA sequence identified samples</th>
<th>Species name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplotype 3</td>
<td>[Z45]</td>
<td><em>Diplostomum spathaceum</em></td>
</tr>
<tr>
<td>Haplotype 6</td>
<td>[Z40]</td>
<td><em>Diplostomum spathaceum</em></td>
</tr>
<tr>
<td>Haplotype 7</td>
<td>[Z47, Z53]</td>
<td><em>Diplostomum spathaceum</em></td>
</tr>
</tbody>
</table>

### Table 2. Gene region pairwise distance relationship (Haplotype 1, Haplotype 2, Haplotype 3, Haplotype 4, Haplotype 5, Haplotype 6, Haplotype 7, Haplotype 8).

<table>
<thead>
<tr>
<th>Haplotype 1</th>
<th>Haplotype 2</th>
<th>Haplotype 3</th>
<th>Haplotype 4</th>
<th>Haplotype 5</th>
<th>Haplotype 6</th>
<th>Haplotype 7</th>
<th>Haplotype 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplotype 2</td>
<td>0,00209</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 3</td>
<td>0,00419</td>
<td>0,00209</td>
<td>0,00419</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 4</td>
<td>0,00842</td>
<td>0,00630</td>
<td>0,00419</td>
<td>0,00419</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 5</td>
<td>0,00419</td>
<td>0,00209</td>
<td>0,00419</td>
<td>0,00841</td>
<td>0,00419</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 6</td>
<td>0,00419</td>
<td>0,00209</td>
<td>0,00419</td>
<td>0,00841</td>
<td>0,00419</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 7</td>
<td>0,00840</td>
<td>0,00629</td>
<td>0,00840</td>
<td>0,01264</td>
<td>0,00840</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Haplotype 8</td>
<td>0,00419</td>
<td>0,00209</td>
<td>0,00419</td>
<td>0,00842</td>
<td>0,00419</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
In the study, phylogenetic trees were created using Maximum Likelihood (ML), Neighbor Joining (NJ), Minimum Evolution (ME) and Bayesian methods to show intra-species molecular phylogenetic relationship according to 8 haplotypes.

Accordingly, the Maximum Likelihood method based on the Tamura-Nei (TN93) model was created (Figure 1). The tree with the highest probability is shown with (-739.9283). The first tree was obtained automatically by applying Neighbour Join and BioNJ algorithms to a matrix of estimated distances using the Maximum Likelihood approach and then selecting the topology with superior probability value.

Accordingly to the Neighbour Joining method, the most suitable tree is shown with total branch length = 0.20291874. In addition, the percentage of replication (10000) in which associated taxa are combined in the bootstrap test is indicated next to the branches. Distances are calculated according to Tamura-Nei model and numerical expressions of base substitutions are given.

According to the Minimum Evolution method, the most suitable tree was shown with total branch length = 0.20291874, and the percentage of replication (10000) in which the associated taxa were combined in the bootstrap test was indicated next to the branches. Distances are calculated according to Tamura-Nei model and numerical expressions of base substitutions are given. The Minimum Evolution tree was searched using the Close-Neighbor-Interchange (CNI) algorithm.

mtDNA sequences made from the eye lens of Capoeta umbla type fish obtained from Keban Dam Lake, Karakaya Dam Lake, and Hazar Lake regions were combined with their morphologies and the Mr. Bayes method was used. According to Bayesian Inference (BI) analysis, 8 haplotypes were formed and it was determined that these haplotypes were ~ 99-100% similar to each other. The phylogenetic tree obtained as a result of Bayesian Inference analysis is given in Figure 2.

Discussion

In this study, genetic characterization study of Diplostomum sp., which is located in the eyes of Capoeta umbla, that naturally spreads in Keban Dam Lake, Karakaya Dam Lake, and Hazar Lake, was carried out. Fish samples used in the study obtained from 5 regions from Keban Dam Lake, 6 regions from Keban Dam Lake and 4 regions from Hazar Lake.

Moszczyńska et al. (2009) determined primers for the barcode region of the COI gene for Platylhelmints in their study. There is sufficient evidence that the MT-COI gene can be used for species identification over a wide taxonomic range. There are studies showing that COI sequences give more accurate results than other DNA barcode regions in studies of digenea separation. The main advantage of barcoding in principle is that it provides comparability between the use of standard markers and sequence data (Moszczyńska et al 2009). In this study, COI sequences were used and this study supports the achievement of clear results.

Locke et al. (2010) in their study to identify the species of Diplostomum metacercaria obtained from fish collected in the Lawrence River of Canada; They reported that species could not be identified due to the morphological aspect of this parasite that parasitizes freshwater fish worldwide and they detected 47 species using COI barcode region sequences. They also demonstrated that host and tissue specificities of most species supported by COI sequences are supported by the ITS sequences of rDNA. In this study, it was not necessary to use a different gene region since the COI sequences gave clear results and the species was proved to be Diplostomum spathaceum.

Cavaleiro et al. (2012) explained the two detected morphotypes, 18S and 5.8S rDNA, were found to have 100% genetic compatibility in the ITS1 gene region. Since the genetic data of this metacercaria are different from the Diplostomum species found in GenBank, a data entry is provided for a new genotype (accession number GQ370809). When compared with molecular phylogenetic analysis of Diplostomum metacercaria obtained from flounder, cluster analysis based on main components and morphometric data, it was found that Diplostomum spathaceum differs with Diplostomum mergi, Diplostomum pseudospathaceum and Diplostomum paracaudum. The second of these species was found most similar to the existing material. In this study, the COI gene region was used and Cavaleiro et al. (2012) and the genetic sequence of Diplostomum spathaceum was found to be the same.

Georgieva et al. (2013), in the first study on molecular research for the diversity of Diplostomum species in Europe, the binding sequence data in COI gene region and ITS1, 5.8S rDNA, ITS2 barcode regions are obtained from larval and mature parasites of snails,
fish and waterfowl in Europe. They found that ITS1, 5.8S rDNA and ITS2 sequences support the results determined by COI by revealing the presence of six genetically different Diplostomum strains examined in snail and fish populations with COI. As can be understood from here, the results of the COI gene region and other regions are the same and showed high reliability in the COI gene region especially in phylogenetic studies. It is seen that it is supported by this study.

Chibwana et al. (2013) explained the lack of these studies, especially in tropical regions, by explaining that diplostomide trematodes consist of a large and diverse digenea group, which may cause serious parasitic pathogen effects in natural and cultured fish. In this study, COI barcode region and ITS1, 5.8S rDNA, ITS2 gene cluster sequences of parasites obtained from catfish (Clarias gariepinus) in Africa were used. It was similar to this study, which was carried out using the COI barcode region.

In their study, Perez et al. (2014) stated that there are several molecular studies that reveal the Diplostomum diversity in Central and Northern Europe, and reported that studies for arachnids (Lymnaea stagnalis and Radix spp.) are limited, and morphological and molecular evidence using this study. They stated that they aimed to fill. In addition, COI barcode region and complete sequences of 5.8S rDNA, ITS1, ITS2 gene cluster were used in the study. However, with this study, the same results were obtained in Diplostomum spathaceum although only the COI gene region was used.
Locke et al. (2015) used ITS1, 5.8S rDNA, ITS2, and COI gene regions for the identification of Diplostomum species in their study and it was reported that the COI gene region is much more effective than other gene regions. In this study, the COI gene region was preferred for the identification of Diplostomum samples. According to the data obtained in our study and the results of the phylogenetic trees, it has been revealed that the COI gene region is quite effective in species identification. Therefore, our study was performed by Locke et al. (2015) are similar.

Choudhary and Agrawal (2016) evaluated three nuclear DNA (18S, ITS1 and ITS2) between the isolates of Opisthorchis (O. pedicellata sp.1 and O. pedicellata sp.2) from two silhouetteroid fish hosts. Molecular analysis showed low similarities, showing 1.9% for ITS1 and 2.6% for ITS2 and 1.6% for 18S. The secondary structure of the ITS2 region showed a remarkable fourhrellex model with relatively short spirals 3 and 4, the longest helix (3) with U-U mismatch in helix 2 and 2 UGGG motifs. Phylogenetic analysis using MEGA 6’s adjacent fusion (NJ) and maximum separation (MP) methods showed that both isolates formed a close group with significant support (100%). In this study, it is similar that it is close (100%).

In our study, molecular characterization of Diplostomum species, that cause infection in the eyes of Capoeta umbila fish located in Karakaya Dam Lake, Kebar Dam Lake and Caspian Lake, located within the borders of Elazig province, was performed. Molecular analysis was evaluated according to COI gene sequence data and all samples were found to be Diplostomum spathaceum. Phylogenetic analysis indicates that there is a closeness between Haplotype 3 and Haplotype 4. The examples that make up Haplotype 3 and Haplotype 4 include examples in all three study regions. Since Hazar Lake forms a closed area, it has no connection with the other two regions. However, the similarity of the Diplostomum samples obtained from all three regions resulted in the idea that they may be transmitted by the main hosts of the parasite between the regions. It may have mediated the transmission of parasites, especially between bird species and all three regions. However, the Diplostomum spathaceum of all samples obtained from the COI sequence indicates that this parasite has a 100% prevalence in the research areas.

Ethical Statement

There is no need for the Ethics Committee (Fırat University Animal Experiments Local Ethics Committee Presidency) (Protocol No: 2016/19; Decision No: 69).

Funding Information

Financial support by TUBITAK (Project Number: 116YS03) is gratefully acknowledged.

Author Contribution

All authors designed the research; SDB collected parasites, analyzed the data. All authors have reviewed and approved the final version of the article.

Conflict of Interest

All authors declare that there is no conflict of interest.

Acknowledgements

This research was supported by The Scientific and Technological Research Council of Turkey with Project no. 116YS03. It has been summarized from Sibel Barata’s doctoral dissertation.

References


gariepinus (Siluriformes: Clariidae), Infection, Genetics and Evolution, 17, 62–70.


