Recep Tayyip Erdoğan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi, 3(1) 34-39 2022
Recep Tayyip Erdogan University Journal of Science and Engineering, 3(1) 34-39 2022
ISSN: 2687-2315 / E-ISSN: 2757-7686

# Population Structure and Molecular Phylogeny of Mesopotamian Spiny Eel (*Mastacembelus mastacembelus*) (Teleostei: Synbranchiformes) in Turkey

**Gökhan KALAYCI[1]*, Recep DURMAZ[1]**

[1] Recep Tayyip Erdoğan University, Faculty of Fisheries, 53100, Rize, Türkiye

*Sorumlu Yazar/Corresponding Author
E-mail: kalaycigokhan@gmail.com

*Araştırma Makalesi/Research Article*
*Geliş Tarihi/Received: 02.06.2022*
*Kabul Tarihi/Accepted: 24.06.2022*

**ABSTRACT**

The Mesopotamian spiny eel Mastacembelus mastacembelus is an endemic freshwater fish confined to the Euphrates and Tigris drainages of Mesopotamia. In this study, it was aimed to reveal the genetic variation of Mesopotamian spiny eel in Turkey with samples from the distribution regions in the Tigris and Euphrates river basins. For this purpose, the partial mitochondrial COI gene region (656 bp) was sequenced from 15 samples from seven populations (Aksu, Merzimen, Karasu, Sinnep, Ambar, Zarova streams and main body of Tigris River). Eight variable nucleotide positions were identified and three of which were polymorphic. Overall haplotype and nucleotide diversity are Hd= 0.642±0.081 and $\pi$=0.00339±0.00069. Four haplotypes were identified, one from the Tigris River basin, and three from the Euphrates River basin. Haplotype network analysis contains four unique haplotypes with at least one mutational step and no haplotype shared between main river drainages. The most common haplotype was H1 and three Tigris populations shared it. The phylogenetic inferences suggest that *Mastacembelus* populations are divided into two main clades. The first clade consists of haplotypes of the Tigris River basin while the second clade contains Euphrates River basin populations.

**Anahtar Kelimeler:** Population genetics, COI, Genetic diversity, Mesopotamian spiny eel

# Türkiye'deki Mezopotamya Dikenli Yılan Balığının (*Mastacembelus mastacembelus*) (Teleostei: Synbranchiformes) Populasyon Yapısı ve Moleküler Filogenisi

**ÖZ**

Mezopotamya dikenli yılan balığı *Mastacembelus mastacembelus*, Mezopotamya'nın Fırat ve Dicle nehri drenajlarıyla sınırlı endemik bir tatlı su balığıdır. Bu çalışmada, Mezopotamya dikenli yılan balığının Türkiye'deki genetik varyasyonunun Dicle ve Fırat nehir havzalarındaki dağılım bölgelerinden örneklerle ortaya çıkarılması amaçlanmıştır. Bu amaçla, yedi popülasyondan (Aksu, Merzimen, Karasu, Sinnep, Ambar ve Zarova dereleri ve Dicle Nehri ana gövde) 15 örneğin kısmi mitokondiyal COI gen bölgesi (656 bç) sekanslanmıştır. Sekiz değişken mükleotit pozisyonu tespit edilmiş olup ve bunların üç tanesi polimorfik nükleotit olarak belirlenmiştir. Genel haplotip ve nükleotid çeşitliliği Hd= 0,642±0,081 ve $\pi$=0,00339±0,00069 olarak hesaplanmıştır. Biri Dicle Nehri Havzasından, üçü Fırat Nehri Havzasından olmak üzere dört haplotip tespit edilmiştir. Haplotip ağ analizinde en az bir mutasyon farkı olan ve ana nehir drenajları arasında haplotip paylaşımı olmayan dört haplotip belirlenmiştir. En yaygın haplotip H1 haplotipi olup ve üç Dicle popülasyonu tarafından paylaşılmaktadır. Filogenetik analizler sonucunda, *Mastacembelus* popülasyonlarının iki ana haplogruba ayrıldığı ortaya çıkarılmıştır. Birinci haplogrubu, Dicle Nehri havzası populasyonlarının, ikinci haplogrubu ise, Fırat Nehri havzası popülasyonlarının ve haplotiplerinin oluşturduğu belirlenmiştir.

**Keywords:** Population genetiği, COI, Genetik çeşitlilik, Mesopotamya dikenli yılan balığı

**Cite as;**
Kalaycı, G., Durmaz, R. (2022). Population Structure and Molecular Phylogeny of Mesopotamian Spiny Eel (*Mastacembelus mastacembelus*) (Teleostei: Synbranchiformes) in Turkey, *Recep Tayyip Erdogan University Journal of Science and Engineering*, 3(1), 34-39. Doi: 10.53501/rteufemud.1124666
Orcid ID: G.Kalaycı, 0000-0003-1255-496X; R.Durmaz, 0000-0002-8686-7152

## 1. Introduction

The anguilliform Mastacembelidae (Teleostei: Synbranchiformes) family currently contains three genera (*Mastacembelus*, *Macrognathus*, and *Sinobdella*) and 86 valid species. Sixty-one of these valid species are members of the genus *Mastacembelus* (Froese and Pauly, 2022). *Mastacembelus* species have intercontinental distribution throughout the Afrotropics, southeast Asia, whereas in the Middle East, just a single species (*Mastacembelus mastacembelus*) was distributed.

Previous molecular studies have primarily focused on regions with high levels of sympatric diversity, biogeographical and phylogeography, and diversification patterns (Alter et al., 2015; Brown et al., 2010; Day et al., 2017) however, population genetic parameters within the group need to be investigated. Recently, Gholamhosseini et al. (2022) examined the morphological and molecular characteristics of *Mastacembelus mastacembelus* in some Tigris, Zohreh, and Persis river populations in Iran. The molecular and morphological differences were investigated in some Mesopotamian spiny eel populations from Turkish inland waters, Tohma Stream, Karakaya Reservoir, and Tigris river (Çakmak and Alp, 2010; Tutar, 2015). In the first population genetics study in Turkey, Tutar (2015) compared nine samples from these three populations based on the concatenated 16S rRNA and 12S rRNA mitochondrial DNA sequence data and found no differences among all three populations. However, molecular variations and the population genetic structure of the species have not been investigated in Turkey's Euphrates River basin populations. Also, Turkey's Tigris River basin populations have not been well studied. Hence, this study aims to reflect the population genetic structure and molecular phylogeny of *Mastacembelus mastacembulus* in Turkey's Tigris and Euphrates river basin populations.

## 2. Materials and Methods

### 2.1. Sample Collection, DNA Extraction and Sequencing

*Mastacembelus mastacembelus* samples were collected and stored Recep Tayyip Erdogan University Zoology Museum of the Faculty of Fisheries, Rize (FFR). We used 15 specimens from seven sampling sites from Turkey drainages of the Tigris and Euphrates river basin (Table 1, Figure 1).

**Table 1.** Locality no, locality and coordinates of *Mastacembelus mastacembelus* species in Turkey

| Locality no | Locality | Coordinates | |
|---|---|---|---|
| 1 | Aksu Stream, Euphrates River Basin | 37°42'33.0"N | 37°56'11.3"E |
| 2 | Merzimen Stream, Euphrates River Basin | 37°17'31.0"N | 37°34'19.7"E |
| 3 | Karasu Deresi Stream, Euphrates River Basin | 37°28'02.1"N | 37°36'58.1"E |
| 4 | Sinnep Stream, Euphrates River Basin | 36°46'21.0"N | 37°16'21.4"E |
| 5 | Ambar Stream, Tigris River Basin | 38°15'52.7"N | 40°27'36.5"E |
| 6 | Main body, Tigris River Basin | 37°50'28.1"N | 40°39'32.9"E |
| 7 | Zarova Stream, Tigris River Basin | 37°49'30.0"N | 41°52'51.0"E |

Genomic DNA was isolated from fin tissue using the GeneAll Genomic DNA purification kit. DNA quantity and quality were checked on 1 % agarose gel electrophoresis and the NanoDrop 2000/c spectrophotometer (Thermo Scientific, Rockford, IL, USA). The partial COI gene (656 bp) was amplified by the FishF1 (5'-TCAACCAACCACAAAGACATTGGCAC-3') and FishR2 (5'-CACTTCAGGGTGACCGAAGAATCAG -3') (Ward et al., 2005) primer pair. PCR reaction was carried out in a 50 µl reaction volume containing 5µl 10x PCR buffer, 3 mM $MgCl_2$, 0.5 mM of each primer, 0.5 mM dNTPs mix, 100 ng template DNA, and 1 u Taq DNA polymerase (New England Biolabs). The PCR was performed under the following conditions: first denaturation at 95 °C for 30 seconds,

denaturation at 95 °C for 30 seconds, annealing at 51 °C for 45 seconds, extension at 68 °C for 45 seconds through 35 cycles, and a final extension at 68 °C for 5 minutes using Biorad T100 (Bio–Rad, Hercules, CA, USA) thermal cycler. The PCR products were run 1 % agarose gel electrophoresis and visualized under UV Quantum–Capt ST4 system (Vilber Lourmat, France). Macrogen Europa Inc. (Amsterdam, Netherlands) purified and sequenced PCR products.


**Figure 1.** Map showing the sampling locations. (1-Aksu s., 2-Merzimen s., 3-Karasu s., 4-Sinnep s., 5-Ambar s. 6-Zarova s. 7- Main body, Tigris River). (Red colored circles represent the Euphrates River and the green ones represent the Tigris River populations.)

### 2.2. Genetic Structure and Phylogenetic Analysis

The nucleotide sequences of COI genes were aligned using the Clustal W algorithm (Thompson et al., 1994) implemented in Bioedit v7.2.5 (Hall, 1999) software. DnaSP version 6.12.03 (Rozas et al., 2017) program was used to determine haplotype number (H), variable and/or polymorphic nucleotide position, haplotype diversity (Hd), and nucleotide diversity ($\pi$) for each population. Nucleotide frequencies and transition/transversion rate were calculated using MEGA X (Kumar et al., 2018). To calculate genetic variation among and within the populations molecular variance analysis, AMOVA was performed using the Arlequin v3.5.1.2 (Excoffier and Lischer, 2010) software. Sequences were submitted to the NCBI GenBank with accession numbers ON827284 - ON827298. To determine the best-fit nucleotide substitution model according to the Akaike information criterion (AIC) and Bayesian information criterion (BIC) jModeltest v. 0.0.1 (Posada, 2008) was used and the HKY model (-ln=947.2231) was selected for phylogenetic analysis. Median-joining (MJ) algorithm (Bandelt et al., 1999) implemented in Network 5.0.0.1 software (www.fluxus-engineering.com) was used for the haplotype network inferences. Phylogenetic relationships among samples were estimated using the maximum likelihood (ML) approach using MEGA X (Kumar et al. 2018) and the Bayesian inference (BI) analysis using the MrBayes 3.1.2 software (Ronquist and Huelsenbeck, 2003). For all phylogenetic analyses *Mastacembelus liberiensis* (LT576957), was selected as outgroup taxa. Pairwise genetic distance estimation among the haplotypes, localities and populations were calculated by MEGA X (Kumar et al., 2018) software using the uncorrected p-distance model.

### 3. Result and Discussion

Nucleotide sequences of the COI gene region (656 bp) were examined in 16 specimens belonging to seven populations of *Mastacembelus mastacembelus* in Southeastern Turkey (Table 2). Fifteen specimens were

sequenced from this study also, a sequence in Ambar stream (LT577004) from Genbank was added to the analysis. The average nucleotide frequencies were determined as 26.70% A, 27.40% T, 29.20% C, and 16.80% G. The transition/transversion (Ti/Tv) rate k1 = 1000 (purines), k2 = 486.074 (pyrimidines) and the overall Ti/Tv bias were calculated as R=340.119. Four haplotypes were identified; one from the Tigris River basin (N=8), and three from the Euphrates River basin (N=8) (Table 2).

**Table 2.** Frequency, distribution, variable nucleotide (nt) positions and Genbank number of COI haplotypes of *Mastacembelus mastacembelus* in Turkey (the locality number is in parentheses)

| Haplotypes | Genbank Accession Number | Variable nt positions | | | | | | | | Euphrates | | | | | Tigris | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 64 | 262 | 328 | 518 | 538 | 559 | 629 | 631 | Aksu s. (1) | Merzimen s. (2) | Karasu s. (3) | Sinnep s. (4) | Ambar s. (5) | Main body, Tigris R. (6) | Zarova S. (7) |
| **H1** | ON827284 | A | T | C | A | T | T | C | A | | | | | 2 | 1 | 5 |
| **H2** | ON827291 | . | . | . | G | C | . | . | G | 2 | 3 | | 1 | | | |
| **H3** | ON827297 | G | . | . | G | C | . | . | G | | | 1 | | | | |
| **H4** | ON827298 | . | C | T | G | C | C | T | G | | | | 1 | | | |

Eight variable sites were identified, and three of which were polymorphic. Overall haplotype and nucleotide diversity are Hd= 0.642±0.081 and $\pi$=0.00339±0.00069. No polymorphisms in the population's haplotype and nucleotide diversity were determined except for the Sinnep population. Considering main river drainages, there are no polymorphisms for the Tigris basin whereas Hd= 0.464±0.200 and $\pi$= 0,00191±0,00110 for the Euphrates River basin. The AMOVA suggested that almost all observed genetic variations occur among groups determined as 82.76% for two main river basins. Haplotype network analysis contains four unique haplotypes with at least one mutational step. No haplotype is shared between main river drainages, which differ by at least three mutational steps. The most common haplotype was H1, which was shared by three Tigris populations, Ambar, Zarova Streams, and the main body of the Tigris River. All haplotypes were distributed in either one, two, or three populations (Figure 2). Pairwise genetic distances between haplotypes were calculated between 0.15% (H2 and H3) and 1.07% (H1 and H4). The genetic distance between Tigris and Euphrates populations was calculated as 0.55 %.

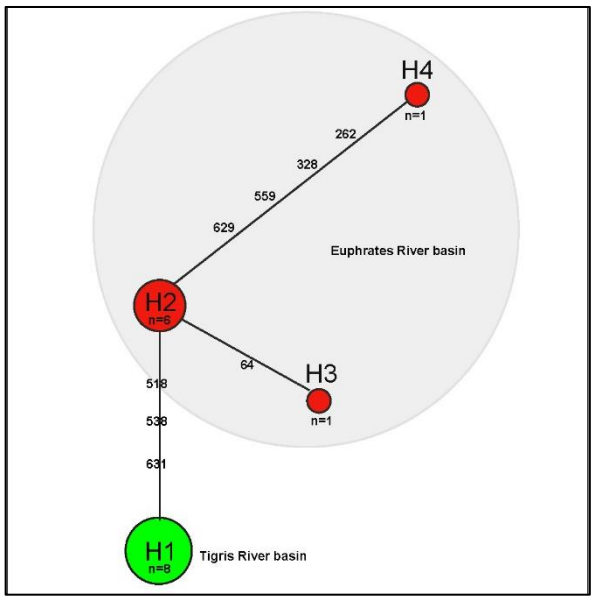


**Figure 2.** Median-joining network of the COI haplotypes. Circle size corresponds to sample size; the number in the lines indicates variable nucleotide position. (n=sample size)

The phylogenetic inferences suggest that *Mastacembelus* populations are divided into two main clades. The first clade consists of haplotypes of the Tigris River basin, while the second clade contains Euphrates River basin populations (Figure 3).

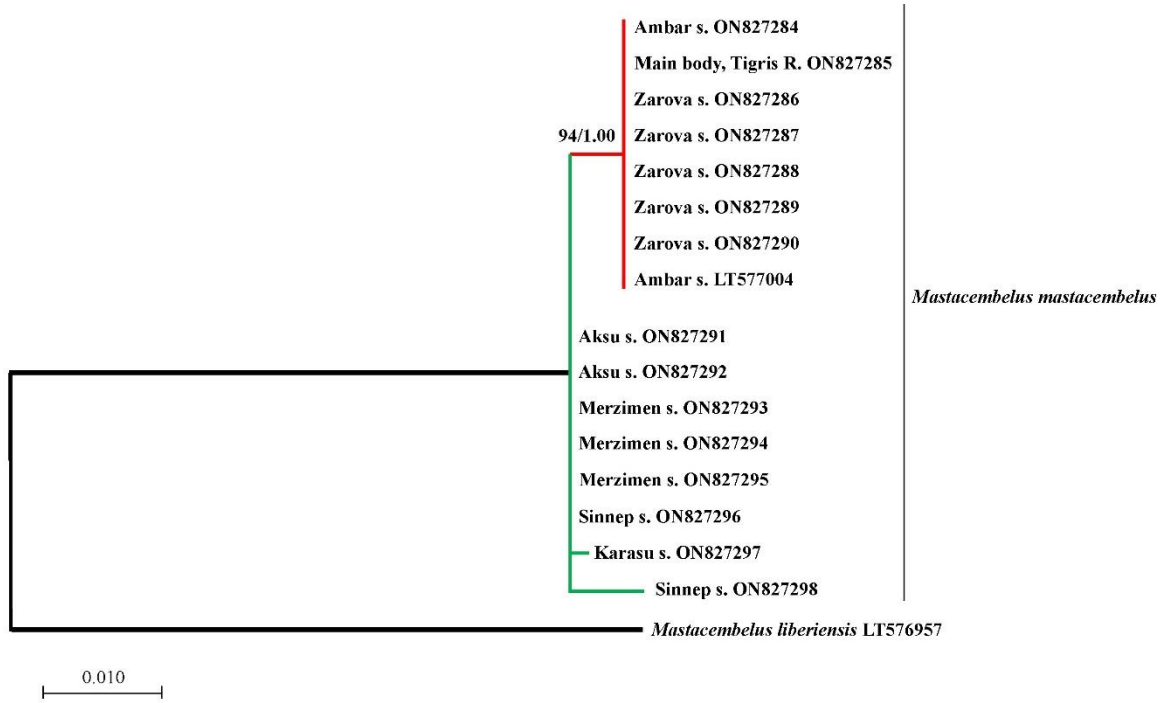


**Figure 3.** Maximum likelihood tree based on COI sequences of *Mastacembelus* populations.

Maximum likelihood and Bayesian inference analyses resulted in congruent trees. Bootstrap and posterior probability values are shown above nodes on a tree if 50% or higher.

Maximum likelihood and Bayesian inference analyses of mitochondrial COI trees are congruent and supported by high bootstrap values (>75) for the distinction of populations of Tigris and Euphrates river basins. Although the sample size is small, this inference is essential because it is the first comparison of Euphrates and Tigris river basin *Mastacembelus* populations. Gholamhosseini et al. (2022) showed that the samples from the Persis and Zohreh rivers, and the part of the Tigris River located on the Iranian borders river systems were not divided into three different groups in the phylogenetic analysis and showed relatively low genetic distances among Iranian tributaries. One of the main reasons may be that the localities belong to the Persian Gulf basin. Also, Tutar (2015) compared three Tigris populations of based on the combined 12S rRNA and 16S rRNA mitochondrial DNA sequence data and found no differences among all three populations. Similarly, in this study, there is quite a low genetic difference between the samples belonging to the same river basin. The samples share the same haplogroup in the phylogenetic trees. Contrast, samples from the Tigris River basin and the Euphrates River were divided into two haplogroups.

## 4. Conclusion

In this study, the genetic structure and comparison of Mesopotamian spiny eel *Mastacembelus mastacembelus* Tigris and Euphrates populations in Turkey were first studied by using mitochondrial COI marker. However, COI molecular marker primarily used for species and higher categorical levels can be informative for populations. This is the first study to identify Tigris and Euphrates populations of Mesopotamian spiny eel based on mitochondrial COI sequences. There is no genetic information or sequences about *M. mastacembelus* Euphrates populations in Genbank or other databases. These sequences are

valuable for further molecular, barcoding and phylogenetic studies.

**Acknowledgements**

I am pleased to thank Dr. Davut Turan, Dr. Cüneyt Kaya, Dr. Esra Bayçelebi and Dr. Münevver Oral for providing to samples.

**Ethics Approval**

This study was conducted with the museum samples stored in Recep Tayyip Erdogan University Zoology Museum of the Faculty of Fisheries, Rize (FFR). Therefore, there is no need for ethics approval.

**References**

Alter, S.E., Brown, B., Stiassny, M.L.J. (2015). Molecular phylogenetics reveals convergent evolution in lower Congo River spiny eels. *BMC Evolutionary Biology*, 15(1): 224. https://doi.org/10.1186/s12862-015-0507-x

Bandelt, H.J., Forster, P., Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Brown, K.J., Rüber, L., Bills, R., Day, J.J. (2010). Mastacembelid eels support lake tanganyika as an evolutionary hotspot of diversification. *BMC Evolutionary Biology*, 10(1): 188. https://doi.org/10.1186/1471-2148-10-188

Çakmak, E., Alp, A. (2010). Morphological differences among the Mesopotamian spiny eel, M*astacembelus mastacembelus* (Banks & Solander 1794), populations. *Turkish Journal of Fisheries and Aquatic Sciences*, 10(1), 87-92. DOI: 10.4194/trjfas.2010.0113

Day, J.J., Fages, A., Brown, K.J., Vreven, E.J., Stiassny, M.L.J., Bills, R., Friel, J.P., Rüber, L. (2017). Multiple independent colonizations into the Congo Basin during the continental radiation of African Mastacembelus spiny eels. *Journal of Biogeography*, 44(10), 2308–2318. https://doi.org/10.1111/jbi.13037

Excoffier, L., Lischer, H.E.L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Froese, R., Pauly, D. (2022). FishBase. 2022. *World Wide Web Electronic Publication. Available at: http://www.Fishbase.org (Accessed on 28 April 2022)*.

Gholamhosseini, A., Razbanian, M., Esmaeili, H.R., Eagderi, S. (2022). Molecular systematics and morphological variation in the Mesopotamian spiny eel *Mastacembelus mastacembelus* (Teleostei: Mastacembelidae). *The European Zoological Journal*, 89(1), 546–555. https://doi.org/10.1080/24750263.2022.2057604

Hall, T.A. (1999). BioEdit a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Posada, D. (2008). jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, 25(7), 1253–1256. https://doi.org/10.1093/molbev/msn083

Ronquist, F., Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572–1574. https://doi.org/10.1093/bioinformatics/btg180

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34(12), 3299–3302. https://doi.org/10.1093/molbev/msx248

Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22, 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Tutar, E. (2015). Combined mitochondrial DNA analysis of the Mesopotamian spiny eel, *Mastacembelus mastacembelus* (Banks & Solander 1794), and its phylogenetic position. *International Journal of Aquatic Biology*, 3(5), 314–322.

Ward, R.D., Zemlak, T.S., Innes, B.H., Last, P.R. Hebert, P.D.N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B*, 360, 1847–1857.