PENSOFT.

MUSEUM FÜR NATURKUNDE BERLIN

DNA barcoding of the genus *Alburnoides* Jeitteles, 1861 (Actinopterygii, Cyprinidae) from Anatolia, Turkey

Halim Canoglu[1], Ismail Aksu[2], Davut Turan[2], Yusuf Bektas[3]

1 *Department of Biology, Faculty of Arts and Sciences, Recep Tayyip Erdogan University, 53100, Rize, Turkiye*
2 *Department of Basic Sciences, Faculty of Fisheries, Recep Tayyip Erdogan University, 53100, Rize, Turkiye*
3 *Deparment of Molecular Biology and Genetic, Faculty of Sciences, Karadeniz Technical University, 61080, Trabzon, Turkiye*

https://zoobank.org/6FBE3E0A-9CE9-4189-AC0C-736A58762B9B

Corresponding author: Yusuf Bektas (ybektas@ktu.edu.tr)

Abstract

The present study investigated the ability of DNA barcoding to reliably identify the endemic freshwater species in Turkey, known as biodiversity hotspots. The barcode region (652 bp) of the mitochondrial cytochrome c oxidase subunit I (COI) was used to barcode 153 individuals from 13 morphologically identified species of the genus *Alburnoides*. Based on the Kimura two-parameter (K2P) evolution model, the average interspecific distance (0.0595) was 31-fold higher than the average intraspecific distance (0.0019). There was a clear-cut barcode gap (0.0158–0.0187) between maximum intraspecific distance (*A. tzanevi* and *A. velioglui*) and minimum nearest-neighbour distance (*A. freyhofi* and *A. kurui*) for Anatolian *Alburnoides* species and a common genetic threshold of 0.0158 sequence divergence was defined for species delimitation. The multiple species delimitation methods (ABGD, ASAP, GMYC and bPTP) revealed a total of 11 molecular operational taxonomic units (MOTUs) for 13 morphospecies. Neighbour-joining (NJ), Maximum Likelihood (ML) and Bayesian Inference (BI) tree analysis indicated that all haplotypes were clustered into two major clades, which corresponded to eleven *Alburnoides* species clusters, with strong bootstrap support. Furthermore, all the specimens clustered in concurrence with the morpho-taxonomic status of the species, except for two species (*A. coskuncelebii* and *A. emineae*) that were morphologically differentiated, but showed overlap in variation for COI-based DNA barcode data with other species. Overall, present results identified that COI-based DNA barcoding is effective for species identification and cataloguing of genus *Alburnoides* in Turkey.

Key Words

Introduction

Turkey is located at the intersection of three (Caucasus, the Irano-Anatolian and the Mediterranean) of the world's 35 biodiversity hotspots (Mittermeier et al. 2004; Fricke et al. 2007). Turkish freshwater wetlands contain numerous lakes, rivers and reservoirs, which host over 180 Cyprinid species (Freyhof et al. 2014; Kuru et al. 2014; Çiçek et al. 2020). However, recent reports have shown a decline in the number of fish caught from most Turkish inland waters due to habitat degradation, pollution, climate change, inadequate management of fisheries, highly-regulated

large river drainages, such as dams and hydropower plants and the introduction of non-native fishes (Fricke et al. 2007; Çiçek et al. 2015; Tarkan et al. 2015; Freyhof et al. 2020). Despite being faced with serious and increasing human activities-induced threats to fish biodiversity, especially in recent years, Turkey is not yet at the level it should be in terms of biodiversity and habitat conservation activities (Şekercioğlu et al. 2011). Phylogeny and phylogeography of *Alburnoides* species in Turkey were revealed by analysing the mitochondrial cytochrome b (Cyt b) gene sequences (Bektas et al. 2019a). However, COI-based DNA barcoding has not been performed for

the detection of intraspecific diversity and species delimitations in the genus *Alburnoides*, most of which are endemic to Anatolia. Hebert et al. (2003) proposed DNA barcoding technology using mitochondrial cytochrome c oxidase subunit I (COI) gene sequence, for which the intraspecific diversity is significantly lower than the interspecific diversity, as a barcode for species identification and classification. For over a decade, the COI gene has been widely used for classifying and identifying Cyprinid fishes (Hubert et al. 2008, 2015; Keskin and Atar 2013; Chakraborty and Ghosh 2014; Geiger et al. 2014; Karim et al. 2016; Lakra et al. 2016; Bayçelebi et al. 2018; Aksu and Bektas 2019; Bektas et al. 2019b; Rahman et al. 2019; Bahadır et al. 2020; Bektas et al. 2022). So far, the COI-based DNA barcode has been used to identify a limited number of cyprinid fish genera or species in Turkey (Keskin et al. 2012; Keskin and Atar 2013; Geiger et al. 2014; Bektas et al. 2019b) in Turkey. However, *Alburnoides* populations in Iran (Jouladeh-Roudbar et al. 2016; Eagderi et al. 2019) and the Caucasus (Levin et al. 2018), which are adjacent regions to Turkey, were barcoded, based on the COI sequences, but DNA barcodes of the *Alburnoides* species in Anatolia, an important centre of endemism, have not yet been generated.

This paper aims to determine intraspecific and interspecific genetic distances and to identify DNA barcodes of *Alburnoides* species from Anatolia.

Materials and methods

Sample collection and morphological identification

One hundred and fifty-three freshwater fish samples belonging to 13 *Alburnoides* species were collected from 47 different sites (some from type localities) in Turkish inland waters between 2016 and 2020 (Fig. 1, Suppl. material 1). Morphological identification was performed in situ by following the counts and measurements of Turan et al. (2016, 2017). After each specimen was morphologically identified, tail fin tissue was taken, preserved in 95% ethanol and stored at –20 °C.

### DNA extraction, PCR amplification and DNA sequencing

Genomic DNA extraction was performed using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) in an automated DNA extraction robot (QIAcube). DNA yield and purity were quantified using a NanoDrop 2000C spectrophotometer (Thermo Scientific). Extracted DNA were stored at -20 °C. Universal PCR primers FishF1 and FishR1 (Ward et al. 2005) were used for COI amplification. Polymerase chain reaction (**PCR**) was performed on BIORAD T100TM Gradient thermocycler (Biorad, Hercules, CA, USA) following the protocol of Ward et al (2005). Thermal cycling conditions were as follows:

3 min initial denaturation at 94 °C, followed by 35 cycles of 1 min at 94 °C, annealing for 40 s at 54 °C, elongation for 1 min at 72 °C and final extention for 7 min at 72 °C. PCR products were electrophoresed on a 1.4% agarose gel containing ethidium bromide and visualised under UV using a gel documentation system. PCR products were purified using the QIAquick kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions and purified PCR products were directly sequenced in both directions using an ABI PRISM 3730x1 automated DNA sequencer (Applied Biosystem, Foster City, USA) following the manufacturer's instruction.

Genetic distance, haplotyping and phylogenetic analyses

Sequences were edited using DNASTAR software (EditSeq V.7.1.0 (Lasergene, DNASTAR) and aligned using the MegAlign V.7.1.0 (Lasergene, DNASTAR), based on the Clustal W (Thompson et al. 1994). The mitochondrial COI sequences were aligned and adjusted to have an equal length of 652 bp. Pairwise genetic distance amongst and within species, polymorphic sites and nucleotide composition were computed using MEGA X (Kumar et al. 2018) with the Kimura 2-parameter (**K2P**) model (Kimura 1980). The haplotype data were generated using the DnaSP v.5 (Librado and Rozas 2009) software. All haplotype sequences were submitted to GenBank under the accession MZ539436–MZ539471 (Suppl. material 1).

Based on the COI barcode region, phylogenetic relationships amongst the Spirlin haplotypes were reconstructed using three methods, Neighbour-joining (**NJ**), Bayesian Inference (**BI**), and Maximum Likelihood (**ML**). The programme jModelTest 0.1.1 (Posada 2008) was used to obtain the appropriate model of evolution using the Akaike Information Criterion (**AIC**; Akaike (1973)) and the Bayesian Information Criterion (**BIC**; Schwarz (1978)). The NJ tree was generated using MEGA X (Kumar et al. 2018) with the Kimura-2 parameter (**K2P**). The node support was evaluated with 1000 bootstrap pseudo-replicates. Bayesian Inference analysis was conducted by using BEAST 1.8.2 (Drummond et al. 2012), setting a coalescent tree prior and an uncorrelated lognormal relaxed clock. Each Bayesian tree was run for 10 million generations with a sampling frequency of 1,000 and were combined with a burn-in set to 25% using LogCombiner 1.8.2 (Drummond et al. 2012). A majority rule consensus tree was constructed using TreeAnnotator 1.8.2 (Rambaut and Drummond 2015). In BEAST, the divergence date of *Alburnoides* and *Rutilus* (approximately 27 Mya; Levy et al. (2009); Perea et al. (2010)) was used as a calibration point for estimating the divergence times of *Alburnoides* species. The ML method was performed with RAxML v.8.1.21 (Stamatakis 2014). The support of each node was estimated using a rapid bootstrap analysis with 1,000 replicates. *Rutilus rutilus* (GenBank Accession No. MW473258) was used as an outgroup. The obtained trees were visualised using FigTree v.1.4.3 (Rambaut 2012) and modified using Adobe Photoshop CS6 software.
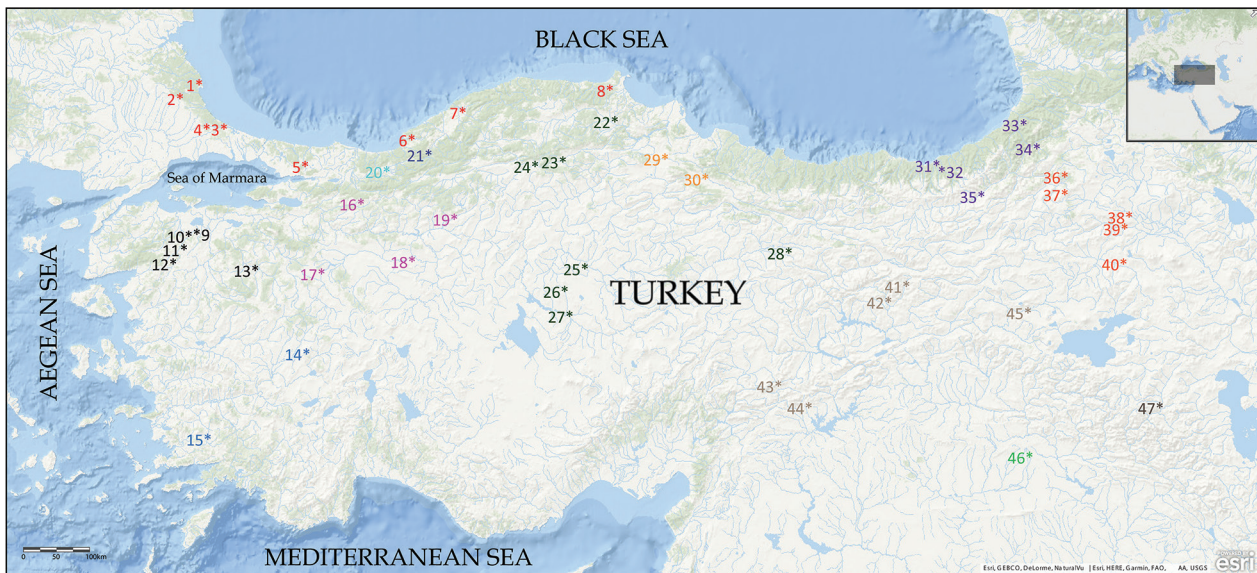
Figure 1. Map showing the range of distribution for each species and sampling points within each species.

Species delimitation methods

To estimate Molecular Operational Taxonomic Units (**MOTUs**), two distance-based methods, Automatic Barcode Gap Discovery (ABGD; Puillandre et al. (2012)) and Assemble Species by Automatic Partitioning (ASAP; Puillandre et al. (2020)) and two tree-based approaches, the General Mixed Yule-coalescent (GMYC; Pons et al. (2006)) and Bayesian poisson tree process (bPTP; Zhang et al. (2013)), were applied to the COI dataset.

For the species delimitation, we used Automatic Barcode Gap Discovery (**ABGD**) (Puillandre et al. 2012) and Assemble Species by Automatic Partitioning (**ASAP**) (Puillandre et al. 2020) (both distance-based methods) on the alignment of CO1 sequences. The ABGD species delimitation was performed via the ABGD web server with default settings (P_{min} = 0.001, P_{max} = 0.1, Steps = 10, X relative gap width = 1.5, Nb bins = 20) and testing of two substitution models (JC69 and K2P) for calculation of the Kimura 2-parameter distances (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html, accessed on 3 March 2022). ASAP species delimitation was performed via the ASAP web server with testing of all three substitution models for calculation of the distances (https://bioinfo.mnhn.fr/abi/public/asap/asapweb.html, accessed on 19 December 2022) using the distance matrix, generated through MEGA X (Kumar et al. 2018).

The single-locus and tree-based delimitation approaches, the General Mixed Yule-coalescent (**GMYC**; Pons et al. (2006)) and Bayesian poisson tree process (bPTP; Zhang et al. (2013)) models, distinguish specimens in species level using coalescence theory. The ultrametric tree required for the GMYC method was obtained using BEAST 1.8.2 (Drummond et al. 2012) on COI dataset. The MCMC chain was run for 10 million generations with the best substitution model calculated in jModelTest 0.1.1 (Posada 2008) and the Yule speciation model. The MCMC log-on prior and posterior values was examined

in Tracer 1.6 (Rambaut et al. 2014) and a burn-in of 1 million generations was used to avoid suboptimal trees in the final consensus tree. The single-threshold GMYC method was run using ultrametric haplotype BEAST trees on the GMYC web server (species.h-its.org/gmyc, accessed on 06 May 2022). For PTP-based MOTU estimation, the needed rooted phylogenetic input-tree was constructed with RAxML v.8.1.21 (Stamatakis 2014) with the best substitution model. The PTP model was implemented following the default parameters and 500,000 MCMC generations on the bPTP web server (http://species.h-its.org/ptp/).

We based species delimitation on a majority consensus of the results inferred from all four methods: two species were considered a single species if at least three out of four methods harmoniously merged two species as one.

Distance-based species assignment

Distance-based methods of species assignments in conjunction with computer simulations can determine the statistical significance of species identification success rates (Huang et al. 2013). To evaluate the potential of the COI dataset for correct species identification and to select the best threshold value, these were tested using the 'Best Close Match' (**BCM**), implemented in Species Identifier 1.8 programme of TaxonDNA software (Meier et al. 2006).

Results

Sequence variation of morphospecies

A total of 36 distinct haplotypes for COI barcode region were identified (Suppl. material 1). The average nucleotide frequencies were 29.71 (T), 25.98 (C), 19.20 (G) and

25.11 (A)%. The mean ratio is approximately 6:1. In this study, the average intraspecific K2P distance was 0.0019, compared with the 0.0595 intraspecific value. The mean interspecific distance was found to be 31-fold higher than the mean intraspecific distance.

Phylogenetic relationships within *Alburnoides*

Phylogenetic relationships amongst *Alburnoides* species were investigated using TIM3+I+G, which was selected as the best evolution model in BI and ML approaches for our dataset. Neighbour-joining, Maximum Likelihood and Bayesian Inference analyses generated similar tree topologies and two main clades were detected consisting of eleven deeply divergent haplogroups within *Alburnoides* (posterior probabilities ≥ 0.95 and 68% ≤ bootstrap values for ML and NJ, Fig. 3). The dataset was analysed using the strict molecular clock in BEAST. The timing of splitting-events of *Alburnoides* lineages are shown in Fig. 3.

Species delimitation

The morphological species hypothesis was not supported by our analyses using ABGD, ASAP, bPTP and GMYC, which delimited 13 species. Nevertheless, 11 putative species were recognised by all the methods that we applied to the COI dataset. We detected a distinct gap between intra- and interspecific genetic distances for 11 hypothetical species of *Alburnoides*, ranging from 0.0158 to 0.0187 for K2P distances (Fig. 2). Results from the four species delimitation analyses, ABGD, ASAP, bPTP and GMYC are summarised in Fig. 3. Relationships amongst the major clades, identified using the GMYC and bPTP analyses (Fig. 3), were identical to each other and to the concatenated phylogenies. Nodal support was high (> 0.95) for most nodes of the BEAST consensus tree (Fig. 3). Results from GMYC and bPTP tests corroborated the presence of eleven independent lineages in *Alburnoides* (Fig. 3). According to the four methods of MOTU delimitation, there was an overlap between *A. tzanevi-A. coskuncelebii* and

A. velioglui-A. emineae belonging to the genus *Alburnoides*. Therefore, the *Alburnoides* dataset was reduced to eleven groups, based on the results of four delimitation methods (Fig. 3).

Species identification with BCM

Due to the observed discrepancy between taxa delimited, based on the four MOTU delimitation methods described above and morphospecies, all *Alburnoides* species in Turkey were assigned into eleven of thirteen categories. Moreover, the DNA barcoding gap, which was the maximum intraspecific distance of each species against its minimum distance to the nearest neighbour, was calculated for all species. The maximum sequence divergence amongst individuals of some species (*A. tzanevi*: 0.01583 and *A. velioglui*: 0.01582, Table 1) exceeded some of the nearest-neighbour distance values (*A. velioglui* & *A. emineae*: 0.0142 and *A. tzanevi* & *A. coskuncelebii*: 0.0120). Therefore, the identification success of the BCM method was re-evaluated by treating *A. velioglui-A. emineae* and *A. tzanevi-A. coskuncelebii* as a species in the dataset using a final distance threshold of 0.0158 (Table 1) and the BCM approach returned 100% correct identifications.

Table 1. Mean and maximum intraspecific and nearest-neighbour (NN) distance (percentage) for all species.

Species	Mean intraspecific distance	Max. intraspecific distance	Nearest neighbour	Distance to NN
A. smyrnae	N/A	N/A	*A. tzanevi*	0.04801
A. tzanevi	0.00758	0.01583	*A. freyhofi*	0.02516
A. freyhofi	0.00145	0.00308	*A. kurui*	0.01875
A. turani	N/A	N/A	*A. eichwaldii*	0.03885
A. manyasensis	0.00028	0.00153	*A. turani*	0.04611
A. velioglui	0.00862	0.01582	*A. eichwaldii*	0.02848
A. kosswigi	0.00329	0.00776	*A. turani*	0.05350
A. diclensis	N/A	N/A	*A. eichwaldii*	0.03076
A. fasciatus	N/A	N/A	*A. eichwaldii*	0.04945
A. kurui	N/A	N/A	*A. freyhofi*	0.01875
A. eichwaldii	0.00168	0.00464	*A. velioglui*	0.02845

N/A denotes that species is a singleton.
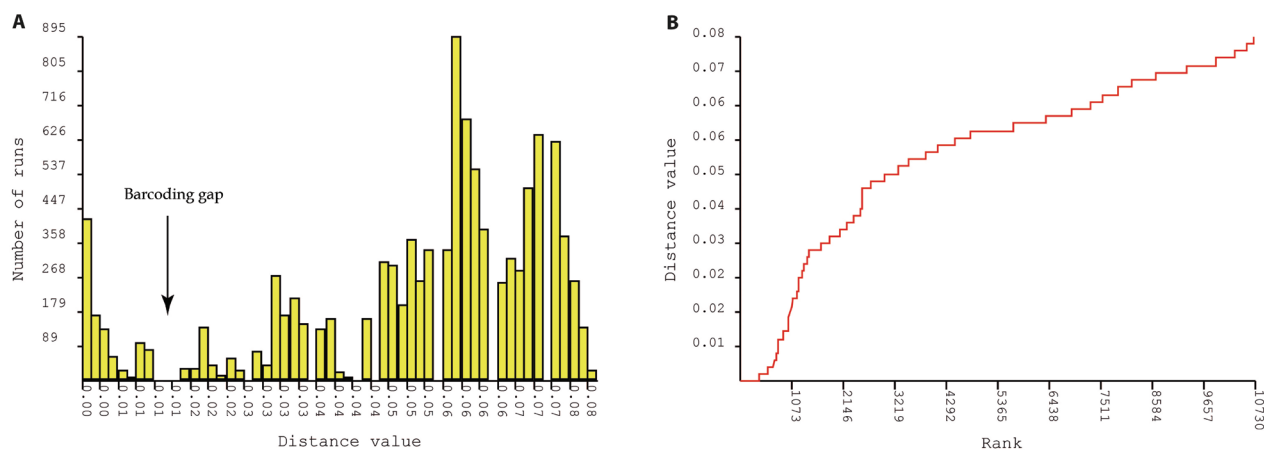
Figure 2. Histogram (**A**) and ranked pairwise (K2P) distances (**B**) amongst COI-based barcode sequences by the automated barcode gap discovery (ABGD) approach.
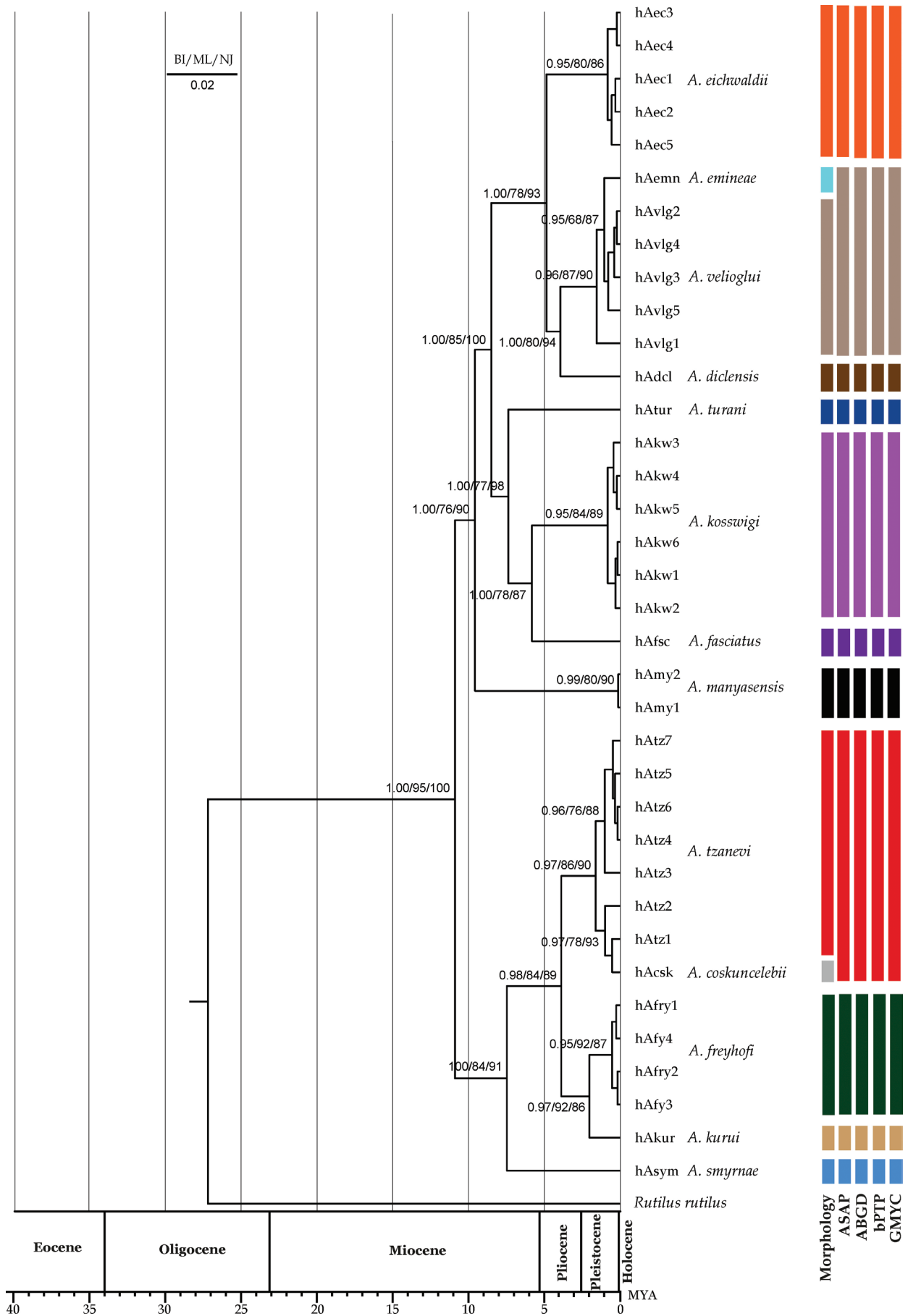
Figure 3. Species delimitation results visualised as bars on an ultrametric Bayesian phylogenetic tree of *Alburnoides* COI gene. The tree phylogroups were obtained according to ABGD, ASAP, GMYC, bPTP and morphology. As outgroup, *Rutilus rutilus* was used.

Discussion

Phylogenetic reconstructions

Morphological identification of Spirlin can be challenging due to their similar morphology and limited morphological information. As maternally-inherited mtDNA evolves about four times faster than that of the nuclear genome, mtDNA can be used to track differences in closely-related taxa and even within species (Stepien and Kocher 1997; Desalle et al. 2017). In this study, the phylogenetic relationships inferred using mitochondrial COI (Fig. 3) are consistent with previous studies using several mitochondrial markers (Jouladeh-Roudbar et al. 2016; Stierandová et al. 2016; Levin et al. 2018), but differ in some of the deepest nodes, with the currently accepted phylogenies of *Alburnoides* (Cyt b, Bektas et al. 2019a). As seen in the phylogenies of Bektas et al. (2019a), the phylogenetic reconstructions strongly support the topology of two clades (western and eastern), containing 13 *Alburnoides* lineages, each corresponding to the species, with moderate-high posterior probability and bootstrap values (> 0.95 pp, > 86 BT; Fig. 3). On the contrary, the analysis of the present COI dataset supports the existence of 11 instead of 13 major lineages in *Alburnoides*. The phylogenetic trees constructed (Fig. 3) showed that the most conspecific haplotypes cluster together, which substantially confirmed previous taxonomic assignments, based on morphology (Turan et al. 2013, 2014, 2016, 2017; Kaya 2020). Nevertheless, our phylogenetic tree topologies are completely not congruent with the morphological data: *A. emineae* is located inside *A. velioglui* and *A. coskuncelebii* is located inside *A. tzanevi*.

As most closely-related species (*A. cockuncelebii* / *A. tzanevi* and *A. emineae* / *A. velioglui*) do not share haplotypes and both species pairs are living allopatrically, the low genetic distances between them may be a result of recent speciation, thus representing the most recently diverged species in the study area. *Alburnoides tzanevi* is found in some rivers that flow into the Black Sea from the Bulgarian-Turkish border to Sinop, the northernmost point of Turkey, while *A. coskuncelebii* has only been reported from the Büyükmelen River, which is also the type locality. As can be seen in Fig. 1, the geographic distributions of *A. tzanevi* and *A. coskuncelebii* overlap. Corrected divergence estimates between the species (Fig. 3) indicate that this phylogeographical structure was initiated by historical events of the Early to Middle Pleistocene. Current distributions of two species are restricted to the Thrace and south-western Anatolia coastal areas of the Black Sea (Fig. 1), where they, thus, most likely survived several later glacial cycles in separate allopatric refugia. The Tigris-Euphrates River system originated in the Late Miocene, developed from small and probably short-lived isolated channels in the Pliocene and became the main drainage system of the region at present (Stow et al. 2020). *Alburnoides velioglui* is found in the upper tributaries of the Euphrates, the Karasu and Murat and *A. emineae* is in

the Beyazsu Stream, which is the upper tributary of the Khabur River in the Euphrates drainage system, within the borders of Turkey. According to our divergence time estimates (Fig. 3), these species were probably isolated in the distant tributaries of the Euphrates after the Pliocene (Trifonov et al. 2018), when the drainage system of the region evolved into its present form.

Species delimitations

In our study, the combination of four species delimitation methods (ABGD, ASAP, bPTP and GMYC) (Fig. 3) implemented, suggested that our results generally confirm the high taxonomic utility of the COI-barcoding region in *Alburnoides* and most closely correspond to the current morphology-based classification of *Alburnoides*. Contrary to the findings of Turan et al. (2014, 2019), which reported a total of 13 *Alburnoides* species, a total of eleven putative partitions were identified amongst Spirlin samples, based on COI-based analysis in this study. There was a strong consensus between the species delimitation results and the phylogenetic clustering, revealing the following lineages: *A. smyrnae*, *A. tzanevi*, *A. eichwaldii*, *A. freyhofi*, *A. kurui*, *A. turani*, *A. manyasensis*, *A. velioglui*, *A. kosswigi*, *A. diclensis* and *A. fasciatus* (Fig. 3).

Algorithms like bPTP/GMYC help us to objectivise species or as an argument to drop poorly-described morphotaxa. The reason why we use different approaches like bPTP and GYMC is that they may give different results. When we test the number of morphotaxa using both approaches for our dataset, we found that the results of bPTP and GYMC analyses are harmonised, but underestimate the number of species. We need to discuss these morphospecies when the pPTP and GMYC do not match the tip labels (phenotypic) or give a lower number. bPTP and GMYC give 11 species standing against the 13 annotated in ML and BI trees. Looking at the topology and branch-lengths distribution in NJ and BI trees, it shows 11 clades regarding intraclade coherence and interclade distinction (Fig. 3). This result tells us that the number of morphologically distinguished species is high. Most of the major clades include only one species, with only two including genetically quite similar species. These species (*A. emineae* and *A. coskuncelebii*) are poorly separated from other ones, so reducing the number of species makes sense. *A. emineae* falls into the *A. velioglui* group, which has a higher intra-species diversity. Similarly, *A. coskuncelebii* is nearly identical to some tzanevi, but the *A. tzanevi*'s intra-species diversity is higher. Therefore, there is little reason molecular-wise to keep *A. emineae* and *A. coskuncelebii* as a different species. Following the bPTP/GMYC result, species epithets, *A. velioglui* and *A. tzanevi*, can be dropped as they have priority. Consequently, the species delimitation algoritm showed that *A. tzanevi* and *A. coskuncelebii* merged into a single putative species, while *A. velioglui* and *A. emineae* were also merged into a single putative species.

Our results revealed two cases of discordance between morphology-based classification and COI-barcoding-based species delimitation. These discordances resulted from the underestimation of possible species in which taxa, currently recognised as valid taxa, collapse in single barcode units. The first case when COI-barcoding failed to recognise morphologically distinct taxa as independent MOTUs includes the *A. tzanevi* and *A. coskuncelebii* from the Black Sea drainage in Thrace and western Anatolia. In our analyses, all methods recognised *A. tzanevi* and *A. coskuncelebii*, as a single unit (Fig. 3). *Alburnoides coskuncelebii* was recently recognised from the Büyük Melen, a catchment situated between the Sakarya and Kızılırmak Rivers, while *A. tzanevi* from small streams in Thrace and adjacent Anatolia (Turan et al. 2019). These results may be explained by allopatric speciation which led to initial differentiation in mtDNA and morphology, which was reported earlier for Spirlins (Bektas et al. 2019a). The second case is the allopatric *A. velioglui* and *A. emineae*, inhabiting Murat, Munzur, Sultansuyu and Beyazsu Rivers (Suppl. material 1) in the upper tributaries of the Euphrates River Basin. These two species are easily distinguished by several diagnostic morphological characters (Turan et al. 2014), yet all species delimitation methods applied herein failed to recognise them as independent groups (Fig. 3). Further studies, including examination of nuclear DNA markers, are required to clarify evolutionary relationships and taxonomic status of these species.

Barcoding gap and automated identification

Species identification, using DNA barcoding, is based on the principle that the genetic distance between two species is much greater than that within a species (Ge et al. 2021). A 10-fold sequence difference between the average interspecific and the average intraspecific differences has been suggested as the standard COI threshold for animal species identification (Hebert et al. 2004). In our present study, the mean interspecific distance was found to be 31-fold higher than the mean intraspecific distance, which was higher than the 23-fold difference observed in Caucasian *Alburnoides* species (Levin et al. 2018) and the 15-fold differences observed in Iranian *Alburnoides* species (Jouladeh-Roudbar et al. 2016). This difference is due to the fact that the average interspecies genetic distance calculated in this study (0.0595) is higher than those (0.0395 for 13 Caucasian species and 0.0386 for 11 Iranian species) in previous studies (Jouladeh-Roudbar et al. 2016; Levin et al. 2018). In other words, our results point to clear genetic differences between the Anatolian species of *Alburnoides*. The high species diversity and clear interspecies genetic differentiation detected for *Alburnoides* can be attributed to the presence of three biodiversity hotspots (Mediterranean, Caucasus and Irano-Anatolian) that coincide in Anatolia (Mittermeier et al. 2011) and the fact that Turkey served as the refugium for southern European taxa during the last glacial age (Reyjol

et al. 2007; Bilgin 2011). Anatolia, indeed, exhibits high species diversity and endemic species richness compared to its neighbouring regions (Levy et al. 2009; Giannetto and Innal 2021). The families Nemacheilidae, Gobiidae, Cobitidae and Salmonidae provide evidence that Anatolia is such a centre of diversity (Kuru et al. 2014).

Based on mitochondrial Cyt b gene sequences, Stierandová et al. (2016) reported that *Alburnoides* lineages in Albania show low mitochondrial distances between 0.83% and 1.43%, which are more likely to be evaluated at an intraspecific level. On the other hand, Levin et al. (2018) reported a COI-based threshold value of 1.3% for *Alburnoides* species in the Caucasian and Middle East regions. Our results indicate that there is a clear barcoding gap (0.0158–0.0187) separating the intraspecific and interspecific distances. Consequently, a common genetic distance threshold of 0.0158 was observed for rapid and reliable identification of *Alburnoides* species. We think that our threshold value, which is higher than that reported by Levin et al. (2018), is due to the complex topography and geomorphology of Anatolia (EEA 2002).

Conclusion

Our study demonstrates the usefulness of DNA barcoding for the identification of *Alburnoides* species in Turkey. This study contributes to the construction of DNA reference barcode data for Turkish fish fauna. We also confirm that DNA barcoding could assist in resolving issues of ambiguity in morphological identification of species. Finally, it has been concluded that COI-based DNA barcoding is a reliable and useful approach for the genetic identification of morphospecies belonging to the genus *Alburnoides*.

Acknowledgements

References

Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki BF (Eds) Second International Symposium on Information Theory. Academiai Kiado: Budapest, 267–281.

Aksu I, Bektas Y (2019) Mitochondrial phylogeny and biogeography of the genus *Gobio* (Teleostei: Cyprinidae) in Turkey. Zoology in the Middle East 65(2): 128–141. https://doi.org/10.1080/09397140.2019.1586126

Bahadır D, Aksu I, Bektaş Y (2020) Genetic analysis of *Carasobarbus* Karaman, 1971 (Actinopterygii Cyprinidae) in Anatolia. Journal of Anatolian Environmental and Animal Sciences 5(4): 520–526. https://doi.org/10.35229/jaes.796051

Bayçelebi E, Kaya C, Turan D, Ergüden SA, Freyhof J (2018) Redescription of *Garra turcica* from southern Anatolia (Teleostei: Cyprinidae). Zootaxa 4524(2): 227–236. https://doi.org/10.11646/zootaxa.4524.2.6

Bektas Y, Aksu I, Kaya C, Baycelebi E, Atasaral Ş, Ekmekçi FG, Turan D (2019a) Phylogeny and phylogeography of the genus *Alburnoides* (Teleostei, Cyprinidae) in Turkey based on mitochondrial DNA sequences. Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis 30(7): 794–805. https://doi.org/10.1080/24701394.2019.1664493

Bektas Y, Aksu I, Kaya C, Turan D (2019b) DNA barcoding of the Genus *Capoeta* (Actinopterygii: Cyprinidae) from Anatolia. Turkish Journal of Fisheries and Aquatic Sciences 19(9): 739–752. https://doi.org/10.4194/1303-2712-v19_9_03

Bektas Y, Aksu I, Kaya C, Bayçelebi E, Turan D (2022) DNA barcoding and species delimitation of the genus *Oxynoemacheilus* (Teleostei: Nemacheilidae) in Anatolia. Journal of Fish Biology 101(3): 505–514. https://doi.org/10.1111/jfb.15114

Bilgin R (2011) Back to the future: The distribution of intraspecific genetic diversity in and around Anatolia. International Journal of Molecular Sciences 12(6): 4080–4103. https://doi.org/10.3390/ijms12064080

Chakraborty M, Ghosh SK (2014) An assessment of the DNA barcodes of Indian freshwater fishes. Gene 537(1): 20–28. https://doi.org/10.1016/j.gene.2013.12.047

Çiçek E, Sungur-Birecikligil S, Fricke R (2015) Freshwater fishes of Turkey: A revised and updated annotated checklist. Biharean Biologist 9(2): 141–157.

Çiçek E, Sungur S, Fricke R (2020) Freshwater lampreys and fishes of Turkey; a revised and updated annotated checklist. Zootaxa 4809(2): 241–270. https://doi.org/10.11646/zootaxa.4809.2.2

Desalle R, Schierwater B, Hadrys H (2017) MtDNA: The small workhorse of evolutionary studies. Frontiers in Bioscience (Landmark Edition) 22(5): 873–887. https://doi.org/10.2741/4522

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29(8): 1969–1973. https://doi.org/10.1093/molbev/mss075

Eagderi S, Jouladeh-Roudbar A, Imani Harsini J, Rostami M (2019) Phylogeny of the members of the genus *Alburnoides* in Iran using COI gene. Iranian Scientific Fisheries Journal 28(3): 125–136. https://doi.org/10.22092/ISFJ.2019.119160

European Environmental Agency [EEA] (2002) Europe's biodiversity – biogeographical regions and seas. http://www.eea.europa.eu/publications/report_2002_0524_154909 [06/02/2023]

Freyhof J, Ekmekçi FG, Ali A, Khamees NR, Özuluğ M, Hamidan N, Küçük F, Smith KG (2014) Freshwater fishes. In: Smith KG, Barrios V, Darwall WRT, Numa C (Eds) The status and distribution of freshwater biodiversity in the Eastern Mediterranean, Chapter 3, Cambridge, UK, Malaga, Spain and Gland, Switzerland: IUCN, 19–42.

Freyhof J, Bergner L, Ford M (2020) Threatened Freshwater Fishes of the Mediterranean Basin Biodiversity Hotspot: Distribution, extinction risk and the impact of hydropower. EuroNatur and RiverWatch [i–viii +] 1–348.

Fricke R, Bilecenoğlu M, Sarı HM (2007) Annotated checklist of fish and lamprey species (*Gnathostomata* and *Petromyzontomorphi*) of Turkey, including a Red List of threatened and declining species. Stuttgarter Beitrage zur Naturkunde. Serie A, Biologie, 706 pp.

Ge Y, Xia C, Wang J, Zhang X, Ma X, Zhou Q (2021) The efficacy of DNA barcoding in the classification, genetic differentiation, and biodiversity assessment of benthic macroinvertebrates. Ecology and Evolution 11(10): 5669–5681. https://doi.org/10.1002/ece3.7470

Geiger MF, Herder F, Monaghan MT, Almada V, Barbieri R, Bariche M, Berrebi P, Bohlen J, Casal-Lopez M, Delmastro GB, Denys GPJ, Dettai A, Doadrio I, Kalogianni E, Karst H, Kottelat M, Kovacic M, Laporte M, Lorenzoni M, Marcic Z, Ozulug M, Perdices A, Perea S, Persat H, Porcelotti S, Puzzi C, Robalo J, Sanda R, Schneider M, Slechtova V, Stoumboudi M, Walter S, Freyhof J (2014) Spatial heterogeneity in the mediterranean biodiversity hotspot affects barcoding accuracy of its freshwater fishes. Molecular Ecology Resources 14(6): 1210–1221. https://doi.org/10.1111/1755-0998.12257

Giannetto D, Innal D (2021) Status of Endemic Freshwater Fish Fauna Inhabiting Major Lakes of Turkey under the Threats of Climate Change and Anthropogenic Disturbances: A Review. Water (Basel) 13(11): 1534. https://doi.org/10.3390/w13111534

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings. Biological Sciences 270(1512): 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004) Identification of Birds through DNA Barcodes. PLoS Biology 2(10): 1657–1663. https://doi.org/10.1371/journal.pbio.0020312

Huang J, Zhang A, Mao S, Huang Y (2013) DNA Barcoding and Species Boundary Delimitation of Selected Species of Chinese Acridoidea (Orthoptera: Caelifera) PLoS ONE 8(12): e82400. https://doi.org/10.1371/journal.pone.0082400

Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burridge M, Watkinson D, Dumont P, Curry A, Bentzen P, Zhang J, April J, Bernatchez L (2008) Identifying Canadian freshwater fishes through DNA barcodes. PLoS ONE 3(6): 2490–e2490. https://doi.org/10.1371/journal.pone.0002490

Hubert N, Wibowo A, Busson F, Caruso D, Sulandari S, Nafiqoh N, Pouyaud L, Rüber L, Avarre JC, Herder F, Hanner R, Keith P, Hadiaty RK (2015) DNA barcoding Indonesian freshwater fishes: Challenges and prospects. DNA Barcodes 3(1): 144–169. https://doi.org/10.1515/dna-2015-0018

Jouladeh-Roudbar A, Eagderi S, Esmaeili HR, Coad BW, Bogutskaya N (2016) A molecular approach to the genus *Alburnoides* using COI sequences data set and the description of a new species, *A. damghani*, from the Damghan River system (the Dasht-e Kavir Basin, Iran) (Actinopterygii, Cyprinidae). ZooKeys 579: 157–181. https://doi.org/10.3897/zookeys.579.7665

Karim A, Iqbal A, Akhtar R, Rizwan M, Amar A, Qamar U, Jahan S (2016) Barcoding of freshwater fishes from Pakistan. Mitochondrial DNA A DNA Mapping Seqequencing Analysis 27(4): 2685–2688. https://doi.org/10.3109/19401736.2015.1043544

Kaya C (2020) Spirlins of the southern Black Sea basin, with the description of a new species (Teleostei: Leuciscidae). Zootaxa 4763(3): 419–428. https://doi.org/10.11646/zootaxa.4763.3.6

Keskin E, Atar HH (2013) DNA barcoding commercially important fish species of Turkey. Molecular Ecology Resources 13(5): 788–797. https://doi.org/10.1111/1755-0998.12120

Keskin E, Ağdamar S, Tarkan AS (2012) DNA barcoding common non-native freshwater fish species in Turkey: Low Genetic Diversity but High Population Structuring. Mitochondrial DNA 24(3): 276–287. https://doi.org/10.3109/19401736.2012.748041

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16(2): 111–120. https://doi.org/10.1007/BF01731581

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

Kuru M, Yerli S, Mangıt F, Ünlü E, Alp A (2014) Fish biodiversity in inland waters of Turkey. Journal of Academic Documents for Fisheries and Aquaculture 1(3): 93–120.

Lakra WS, Singh M, Goswami M, Gopalakrishnan A, Lal KK, Mohindra V, Sarkar UK, Punia PP, Singh KV, Bhatt JP, Ayyappan S (2016) DNA barcoding Indian freshwater fishes. Mitochondrial DNA A DNA Mapping Seqequencing Analysis 27(6): 4510–4517. https://doi.org/10.3109/19401736.2015.1101540

Levin BA, Simonov E, Matveyev MP, Artaev ON, Mustafayev NJ, Pashkov AN, Roubenyan HR (2018) DNA barcoding of the fishes of the genus *Alburnoides* (Actinopterygii, Cyprinidae) from Caucasus. Mitochondrial DNA A DNA Mapping Seqequencing Analysis 29(1): 49–55. https://doi.org/10.1080/24701394.2016.1238900

Levy A, Doadrio I, Almada VC (2009) Historical biogeography of European leuciscins (Cyprinidae): Evaluating the Lago Mare dispersal hypothesis. Journal of Biogeography 36(1): 55–65. https://doi.org/10.1111/j.1365-2699.2008.01969.x

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25(11): 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in *Diptera*: A tale of high intraspecific variability and low identification success. Systematic Biology 55(5): 715–728. https://doi.org/10.1080/10635150600969864

Mittermeier RA, Gill PR, Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux J, da Fonseca GAB (2004) Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions. Cemex-Agrupacion Sierra Madre, Mexico City, Mexico, 392 pp.

Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C (2011) Global Biodiversity Conservation: The Critical Role of Hotspots. In: Zachos F, Habel J (Eds) Biodiversity Hotspots. Springer, Berlin, Heidelberg, 3–22. https://doi.org/10.1007/978-3-642-20992-5_1

Perea S, Böhme M, Zupančič P, Freyhof J, Sanda R, Özuluğ M, Abdoli A, Doadrio I (2010) Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. BMC Evolutionary Biology 10(265): 265. https://doi.org/10.1186/1471-2148-10-265

Pons J, Barraclough T, Gomez-Zurita J, Cardoso A, Duran D, Hazell S, Kamoun S, Sumlin W, Vogler A (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55(4): 595–610. https://doi.org/10.1080/10635150600852011

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

Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier MC, Cruaud C, Holford M, Samadi S (2012) Large-scale species delimitation method for hyperdiverse groups. Molecular Ecology 21(11): 2671–2691. https://doi.org/10.1111/j.1365-294X.2012.05559.x

Puillandre N, Brouille S, Achaz G (2020) ASAP: Assemble species by automatic partitioning. Molecular Ecology Resources 21(2): 609–620. https://doi.org/10.1111/1755-0998.13281

Rahman MM, Norén M, Mollah AR, Kullander SO (2019) Building a DNA barcode library for the freshwater fishes of Bangladesh. Scientific Reports 9(1): 9382. https://doi.org/10.1038/s41598-019-45379-6

Rambaut A (2012) Figtree v 1.4.0. http://tree.bio.ed.ac.uk/software/figtree/

Rambaut A, Drummond AJ (2015) TreeAnnotator v1.8.2: MCMC Output analysis. http://beast.bio.ed.ac.uk

Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer v1.6. http://beast.bio.ed.ac.uk/Tracer

Reyjol Y, Hugueny B, Pont D, Bianco PG, Beier U, Caiola N, Casals F, Cowx IG, Economou A, Ferreira MT, Haidvogl G, Noble R, De Sostoa A, Vigneron T, Virbickas T (2007) Patterns in species richness and endemism of European freshwater fish. Global Ecology and Biogeography 16(1): 65–75. https://doi.org/10.1111/j.1466-8238.2006.00264.x

Schwarz G (1978) Estimating the dimension of a model. Annals of Statistics 6(2): 461–464. https://doi.org/10.1214/aos/1176344136

Şekercioğlu CH, Anderson S, Akçay E, Bilgin R (2011) Turkey's rich natural heritage under assault. Science 334(6063): 1637–1639. https://doi.org/10.1126/science.334.6063.1637-b

Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9): 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stepien CA, Kocher TD (1997) Molecules and morphology in studies of fish evolution. In: Kocher TD, Stepien SD (Eds) Molecular Systematics of Fishes. Academic Press, San Diego, 1–11. https://doi.org/10.1016/B978-012417540-2/50002-6

Stierandová S, Vukic J, Vasil'eva ED, Zogaris S, Shumka S, Halačka K, Vetešník L, Švátora M, Nowak M, Stefanov T, Koščo J, Mendel J (2016) A multilocus as-sessment of nuclear and mitochondrial sequence data elu-cidates phylogenetic relationships among European spirlins (*Alburnoides*, Cyprinidae). Molecular Phylogenetics and Evolution 94: 479–491. https://doi.org/10.1016/j.ympev.2015.10.025

Stow D, Nicholson U, Kearsey S, Tatum D, Gardiner A, Ghabra A, Jaweesh M (2020) The Pliocene-Recent Euphrates river system: Sediment facies and architecture as an analogue for subsurface reservoirs. Energy Geoscience 1(3–4): 174–193. https://doi.org/10.1016/j.engeos.2020.07.005

Tarkan AS, Marr SM, Ekmekçi FG (2015) Non-native and translocated freshwater fish species in Turkey. Fishes in Mediterranean Environments 3: 1–28. https://doi.org/10.29094/FiSHMED.2015.003

Thompson JD, Higgins DG, Gibson TJ (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

Trifonov VG, Çelik H, Simakova AN, Bachmanov DM, Frolov PD, Trikhunkov YaI, Tesakov AS, Titov VV, Lebedev VA, Ozherelyev DV, Latyshev AV, Sychevskaya EK (2018) Pliocene –

early Pleistocene history of the Euphrates valley applied to late Cenozoic environment of the northern Arabian Plate and its surrounding, eastern Turkey. Quaternary International 493: 137–165. https://doi.org/10.1016/j.quaint.2018.06.009

Turan D, Ekmekci FG, Kaya C, Güçlü SS (2013) *Alburnoides manyasensis* (Actinopterygii, Cyprinidae), a new species of cyprinid fish from Manyas Lake basin, Turkey. ZooKeys 276: 85–102. https://doi.org/10.3897/zookeys.276.4107

Turan D, Kaya C, Ekmekçi FG, Bayçelebi E (2014) Three new species of *Alburnoides* (Teleostei: Cyprinidae) from Euphrates River, Eastern Anatolia, Turkey. Zootaxa 3754(2): 101–116. https://doi.org/10.11646/zootaxa.3754.2.1

Turan D, Bektaş Y, Kaya C, Bayçelebi E (2016) *Alburnoides diclensis* (Actinopterygii: Cyprinidae), a new species of cyprinid fish from the upper Tigris River, Turkey. Zootaxa 4067(1): 79–87. https://doi.org/10.11646/zootaxa.4067.1.6

Turan D, Kaya C, Bayçelebi E, Bektaş Y, Ekmekçi FG (2017) Three new species of *Alburnoides* from the southern Black Sea basin (Teleostei: Cyprinidae). Zootaxa 4242(3): 565–577. https://doi.org/10.11646/zootaxa.4242.3.8

Turan D, Kaya C, Aksu I, Bayçelebi E, Bektaş Y (2019) *Alburnoides coskuncelebii*, a new species from the stream Büyük Melen in northwestern Anatolia (Teleostei: Leuciscidae). Ichthyological Exploration of Freshwaters 29(3): 201–210. https://doi.org/10.23788/IEF-1107

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences 360(1462): 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Zhang J, Kapli P, Pavlidis P, Stamatakis A (2013) A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29(22): 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Supplementary material 1

List of species including voucher's specimen number, species name, locality information and Specimen voucher No/GenBank accession no.

Authors: Halim Canoglu, Ismail Aksu, Davut Turan, Yusuf Bektas

Data type: table (docx file)

Link: https://doi.org/10.3897/zse.99.94333.suppl1