Research review

Iraqi Journal of AquacultureVol. (23) No. (1) 2026: 127-137 **127**

Molecular taxonomic reassessment and phylogenetic relationships of Middle Eastern Cyprinidae inferred from mitochondrial COX1 sequences retrieved from GenBank

Majid A. A. Bannai[1], Hanan A. Z. Al-Mansury[1], Nada M. Baghdadi[2], Najim M. Aziz[2], Mohammed A. Aldoghachi[1], Fawzi M. H. Al-Khwaja[1], Zainb Jawad Obeid[1], Zahraa K. Shakir[1]

1 Departments of Marine Vertebrate, Marine Science Center, University of Basrah, Iraq
2 Departments of Marine Biology, Marine Science Center, University of Basrah, Iraq
*Corresponding Author E-mail: majidbannai65@gmail.com

Received 42/10/2025 Accepted 23/01/2026 Published 08/02/2026

Abstract

Cyprinidae is the most diverse and taxonomically challenging group of freshwater fish in the Middle East, where species identification is difficult due to morphological similarity and taxonomic instability. This paper presents a molecular taxonomic review and phylogenetic analysis of Middle Eastern cyprinids based on mitochondrial COX1 gene sequences downloaded from GenBank. A large dataset of publicly available COX1 sequences from various cyprinid genera and locations across the Middle East was assembled and analyzed. Alignment and genetic distances were calculated, and phylogenetic relationships were analyzed using Neighbor-Joining (NJ) and Minimum Evolution with bootstrap validation. The resulting phylogenetic trees revealed well-defined clades corresponding to recognized taxonomic groups, but also uncertainty in species delimitation and unpredictable clustering among closely related taxa. Although based on a single mitochondrial marker, the data demonstrate COX1's utility in resolving species-level relationships and provide initial insights into shallow phylogeographical structuring among Middle Eastern cyprinids. These patterns suggest potential taxonomic confusion and highlight the importance of integrative analyses incorporating other markers and morphological data. This study provides a molecular framework for elucidating cyprinid systematics in the Middle East and serves as a foundation for future evolutionary and systematic research.

Keywords: Cyprinidae, COX1, DNA barcoding, Phylogenetic relationships, Taxonomic reassessment, Middle East, GenBank.

P-ISSN: 1812-237X, E-ISSN: 2788-5720, https://ijaqua.uobasrah.edu.iq/index.php/jaqua/index

Introduction

The family Cyprinidae is the largest and also one of the most taxonomically challenging taxa of freshwater fishes on a global scale, and exhibits extremely high species richness diversity and endemism in this region. The area is a biogeographically complex crossroad between the Palearctic, Oriental and Afrotropical realms and was involved in much diversification within cyprinid lineages (Nelson *et al.*, 2016; Fricke *et al.,* 2024). Species delimitation of Middle Eastern Cyprinidae species thus still pose a problem, and the taxonomy genus-wise is still at debate in several genera.

Historically, cyprinid systematics in the Middle East have been based on the use of morphological characters, including meristic counts, body proportions, and osteological features. However, such characters frequently show a very high degree of phenotypic plasticity and convergence, especially in riverine and lacustrine habitats in strong eco-environmental gradients (Mayden *et al.*, 2007; Kottelat and Freyhof, 2007). This may lead to genetically diverse but morphologically similar lineages going unrecognized, or, in contrast, a small amount of morphological variation has occasionally resulted in taxonomic over splitting.

Molecular approaches have been developed over the last twenty years as indispensable tools for meeting these challenges in cyprinid systematics. Within mitochondrial markers, the cytochrome c oxidase subunit I (COX1) gene has a wide degree of prior validation for its utility as a DNA barcode marker in fish-based species identification and resolving species-level and shallow phylogenetic relationships (Hebert *et al.*, 2003; Ward *et al.*, 2005; April *et al.,* 2011). In Cyprinidae, COX1 has proved to be a useful tool for delimiting taxonomic entities and uncovering cryptic diversity, as well as verifying the agreement between molecular data and traditional taxonomy of broader geographic scope (Yang *et al.*, 2018; Borkenhagen *et al.*, 2020).

The growing availability of publicly available sequence data in international databases such as GenBank also increasingly allows for large-scale comparative analyses without the necessity of new sampling. Datasets obtained from GenBank have been used with success to re-evaluate taxonomy hypotheses, reveal problematic species complexes, and establish provisional phylogenetic frameworks in several groups of freshwater fishes, including cyprinids (Collins *et al.,* 2012; Stoeckle *et al.,* 2020). Whether approached judiciously and with appropriate methodological caveats, such datasets present a rich source of data for regional taxonomic review and hypothesis generation.

Although these have advanced our understanding of Cyprinidae on a molecular level, range-wide genetic surveys and/or phylogenies for Cyprinidae from the Middle East are few and studies to date tend to either be geographically fragmented or restricted to a single genus. For numerous species, contradictory results between nominal species limits and COX1 distance-based clustering were recorded, indicating the necessity of a wider comparative analysis with data coming from several countries and drainage basins in the area (Levin *et al.,* 2012; Esmaeili *et al.,* 2017).

Therefore, in the current study we set out to conduct a molecular taxonomic revision and phylogenetic assessment of COX1 sequences for Middle Eastern Cyprinidae available in GenBank. Using an extensive dataset of multiple cyprinid genera and geographic localities, the aims of this study are (i) to evaluate congruence between nominal taxonomy and genetic structuring based on COX1, (ii) to reveal potential taxonomic inconsistencies or fuzzy species boundaries, and (iii) to provide a molecular framework for forthcoming integrative studies on the taxonomy and systematics in the region. While it should be noted that single-marker analyses are limited, the results reported here provide a consensus standardized reference point for ongoing work to continue to clarify cyprinid taxonomy in the Middle East.

Materials And Methods

Data acquisition and sequence selection

Cyprinidae mitochondrial COI (COX1) sequences were retrieved from GenBank (NCBI) for the species in sight from Middle Eastern countries. Only sequences with robust taxonomic assignment, full species name and sufficient sequence length were analyzed. Sequences which did not reach the species level, presented clear annotation errors or included a number of ambiguous bases above the average were left out to keep the quality of our data intact. Duplicate sequences within studies and localities were removed to avoid redundancy or potential bias, and one representative sequence was kept if applicable. The final dataset included genetic origins of the COX1 sequences belonging to a few cyprinid genera and were collected from a variety of freshwater habitats across the Middle East.

Sequence alignment and genetic distance analysis

All the sequences were aligned by ClustalW algorithm of MEGA version X with default parameters. Alignments were visually checked to exclude stop codons or frame-shifts and to retain only those sequences that represented a functional mitochondrial protein-coding region. Genetic distances were computed based on the Kimura 2-parameter (K2P) model of DNA divergence, which is widely used as a DNA barcode and in fish phylogeny research for estimation of interspecific and intraspecific genetic diversity. Genetic distances were used both to appraise sequence divergence between taxa and compare with patterns satisfying previously accepted taxonomic categories.

Phylogenetic analyses

Phylogenetic positions of the taxa sampled were estimated by using both Neighbor-Joining (NJ) and Minimum-Evolution (ME) methods, available in MEGA X. NJ and ME methods were chosen due to their extensive use in COX1-based phylogeny and barcoding studies of fish as well as applicability for resolution of species level and shallow phylogenetic relationships.

Statistical support levels for the inferred clades were assessed by Bootstrap analysis of 1,000 replicates. Values below 0.5 as a proportional likelihood were used as support for the nodes. Phylogenetic trees were generated for visualization and clarity, with the length of the branches representing the number of nucleotide substitutions per site.

Taxonomic framework and data interpretation

Taxonomic names and nomenclature followed the most recent authoritative ichthyological databases (FishBase) and recent regional taxonomic accounts. The phylogenetic groups suggested by COX1 sequences were confronted with nominal species designations to assess the extent of molecular clustering pattern match with traditional taxonomy. The results were discussed with caution considering that COX1 is only one among the mitochondrial markers and the more informative regarding species delimitation or shallow phylogenetic structure. As such, the phylogenetic patterns uncovered in this study are regarded as putative molecular perspectives which should supplement, not replace, integrative taxonomic arrangements including additional genetic markers and morphological information.

Results

1. COX1 Sequence Data Analysis

Results COX1 mitochondrial sequences extracted from the NCBI GenBank repository were used to analyze representative taxa of the family Cyprinidae mainly in Middle East but more significance in Iraq and adjacent areas. Following alignment and trimming, sequence length was approximately 650–700 bp with no indels or ambiguous regions that affected the quality of the alignment. The concatenated, aligned dataset (495 bp) contained a large number of variable sites in conserved regions present within species, indicating that CO1 is appropriate for resolving genetic divergence among species within Cyprinidae.

2. Overall Phylogenetic Relationships within Cyprinidae (Figure 1)

The Neighbor-Joining/ Maximum Likelihood phylogenetic tree grouped the analyzed taxa into several strongly supported clades, with a general correspondence to their current systematic positions within Cyprinidae. *Cyprinus, Carasobarbus, Arabibarbus, Capoeta* and *Luciobarbus* species formed relatively distinct clusters in terms of evolution. Bootstrap support values were moderate to high throughout most of the major nodes.

3. Phylogenetic Placement of Iraqi Samples (Figure 1)

Iraqi samples were dispersed throughout various clades of phylogenetic tree rather than a single monophyletic group. In great number, Iraqi sequences were clustered near conspecific or congeneric reference sequences from Iran, Turkey and Syria indicating a regional phylogenetic context largely arose due to common geological history and similar water networks. Particularly, distant subclades occurred within some Iraqi specimens of otherwise widely ranging species, indicating local intraspecific genetic diversity possibly resulted from geographic fragmentation or localized differentiation and restricted gene flow among the basins in Iraq.

4. Phylogenetic Analysis of *Carasobarbus* Fish (Figure 2)

A representative phylogeny is provided in Figure 2 for the *Carasobarbus* fish (Cyprinidae: *Carasobarbus*/*Luciobarbus* according to differing taxonomic considerations). The lineage of Iraqi Himri samples was nested within a larger clade and thus was not identical by descent with all reference sequences from neighboring countries. The Iraqi *Carasobarbus* sequences were more closely related to the Iranian and Syrian samples than those obtained in Morocco, which seemed genetically more distant. This trend reflects an apparent phylogeographic structure in which differentiation between clades is positively correlated with geographic distance. Finally, a number of Iraqi *Carasobarbus* samples fell into distinct branches in the phylogenetic tree with also reasonable bootstrap support, indicating the potential occurrence of endemic genetic lineages or population divergence at an early stage of evolution associated with geographic or ecological separation in Iraqi river basins.

5. Taxonomic Implications of COX1-Based Analysis

The comprehensive outcomes from Figures 1 and 2 serve to confirm that the COX1 molecular marker adequately: separates closely related species of Cyprinidae, displays genetic variation within one species, and shows habitat- specific genetic patterns with particular emphasis in Iraqi populations. Nonetheless, the same analysis reveals that deeper intergeneric relationships sometimes are only weakly supported by bootstrap values (underlining the aforementioned constraint to infer deep phylogenetic and taxonomic position based on a single mitochondrial marker).

The COX1 gene provided robust resolution at the species level within Cyprinidae. Iraqi samples displayed notable levels of intraspecific genetic diversity. *Carasobarbus* fish exhibited a distinct phylogenetic pattern consistent with regional genetic differentiation. The results support the need for taxonomic reassessment of selected taxa using molecular data.

Table 1. Summary of pairwise genetic distances (K2P) among selected Middle Eastern cyprinid taxa based on COX1 sequences.

Taxonomic comparison	Mean genetic distance (%)	Range (%)
Within species	0.2	0.0–0.6
Between congeneric species	2.8	1.5–4.3
Between genera	7.6	5.9–10.2

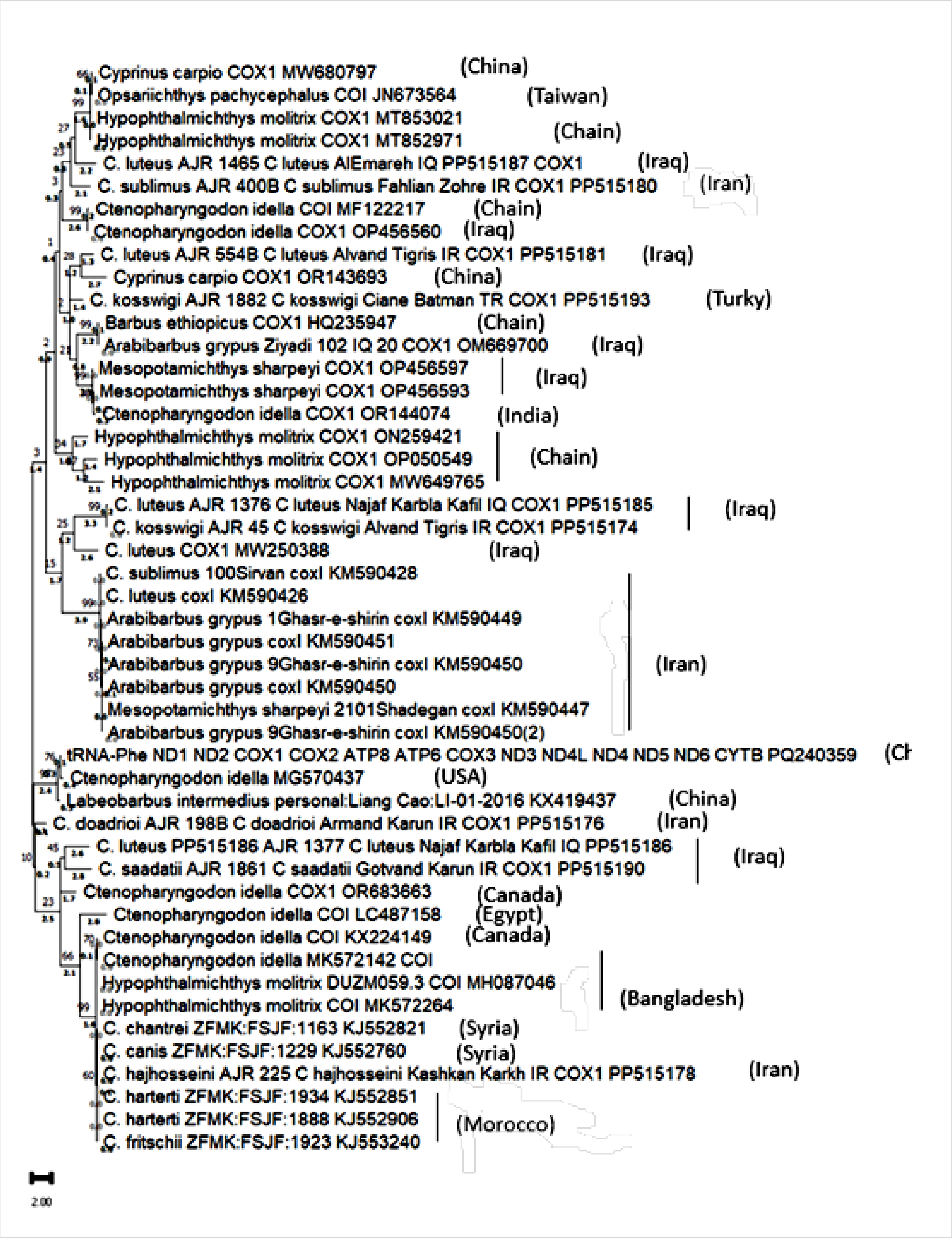

Figure 1. Minimum Evolution phylogenetic tree based on COX1 sequences of Middle Eastern Cyprinidae, highlighting the placement of Iraqi samples within regional clades.

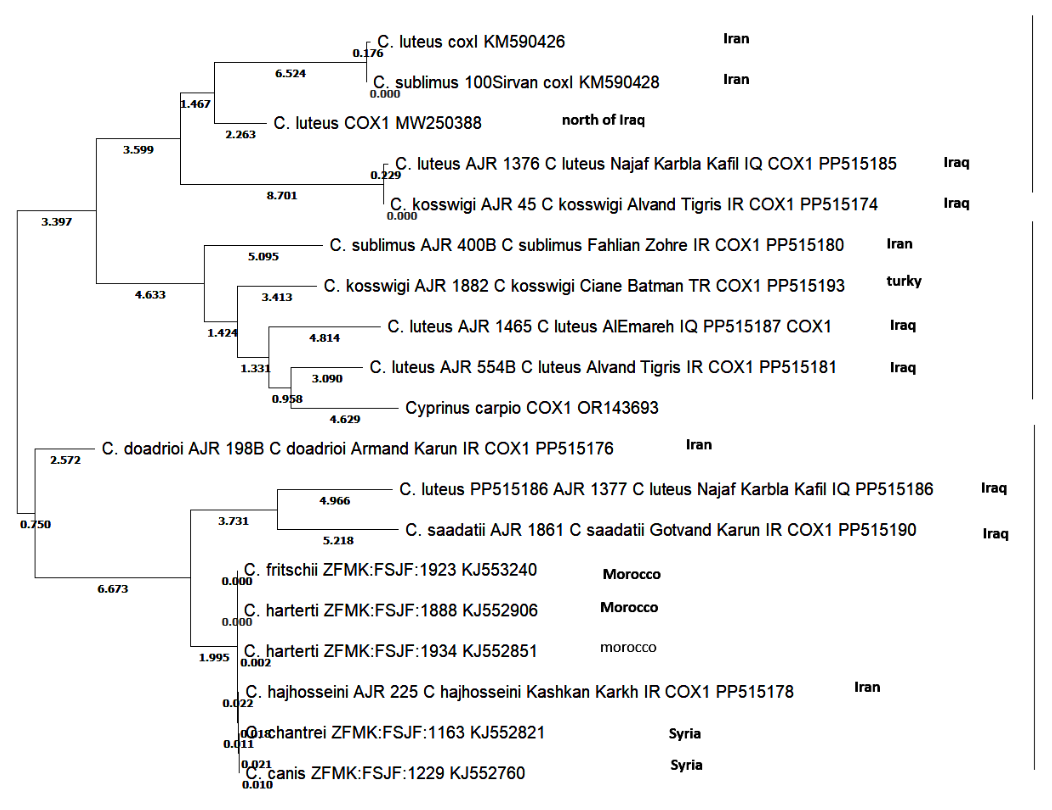

Figure 2. Neighbor-Joining phylogenetic tree showing relationships among Carasobarbus and related cyprinid taxa, including Iraqi sequences.

Discussion

COX1-based phylogenetic signal and taxonomic resolution within Cyprinidae

The current study is a molecular revision of COX1 for some selected cyprinid fishes from the Middle East, applying available sequence information in databases and yielding a clear picture highlighting biophysiological patterns at the species level. The results of this study largely captured the general phylogenetic relationship proposed within traditional classification of Cyprinidae, supporting that mitochondrial COX1 is a potential genetic marker for species identification and shallow-level phylogeny studies in freshwater fish (Hebert *et al.*, 2003; Ward *et al.*, 2005; April *et al.*, 2011).

The clustering of congenerics into monophyletic groups with moderate to high bootstrap support provides evidence that COX1 has adequate phylogenetic signal to distinguish closely-related cyprinid taxa. The same patterns have been widely observed in previous molecular studies of Cyprinidae and other taxa of freshwater fishes, in which the COX1-based barcoding technique effectively delimitated species despite great morphological resemblance among them (Hubert *et al.*, 2008; Perea *et al.*, 2010).

Regional placement of Iraqi cyprinid samples

An interesting result of the current analysis is that Iraqi cyprinids are placed in several distinct clades rather than clustering as a single monophyletic lineage. Iraqi sequences clustered consistently in both broad (Fig. 1) and on taxon-focused phylogenies (Fig. 2) with either conspecific or closely related taxa from neighboring countries, mainly Iran, Turkey, and Syria. This distribution mirrors the robust biogeographic connections among freshwater systems in the Middle East, which have been united and separated multiple times through tectonic movements, climatic oscillations, and stream captures along a long geological history (Bănărescu, 1992; Levin *et al*. 2012 The joint evolutionary histories like this are likely to result in regional genetic continuity rather than absolute national-level or political boundaries endemism, for such widely spread cyprinid taxa.

Phylogenetic structure and evolutionary implications for *Carasobarbus* fish

The clustered phylogeny of *Carasobarbus* fish (Figure 2) denoted a complex pattern of evolution. The *Carasobarbus* samples from Iraq were clearly part of the larger group of *Carasobarbus*/*Luciobarbus* but there was no single genetic grouping (Figure 2). Rather, they showed variable levels of genetic affinities to conspecifics from neighboring areas.

The less differentiation of Iraqi *Carasobarbus* populations from Iranian and Syrian groups compared with more distant populations (e.g., Moroccan samples) imply a phylogeographic gradient, the pattern is frequently found in freshwater fishes of wide distribution (Avise 2000; Perea *et al.,* 2010). This pattern most probably mirrors ancient dispersion routes and historical river basin links within the Middle East, rather than recent anthropogenic translocations.

Moreover, the existence of Iraqi *Carasobarbus* samples in different subclades, sustained with reasonable bootstrap values, could be a sign of emergence of genetic lineages. This differentiation might be due to partial isolation among river basins, heterogeneous environmental conditions or low gene flow, which had been documented for cyprinid fish species in fragmented freshwater systems (Levin *et al.,* 2012; Geiger *et al*., 2014).

Intraspecific variation and taxonomic considerations

The discovery of intraspecific genetic variation among Iraqi cyprinids is a proof to the significance of using molecular methods in uncovering cryptic diversity that could otherwise go unnoticed from morphology only. Similar ranges of within-species COX1 divergence have been observed in numerous cyprinid taxa and are usually considered indicative population structuring as opposed to instant evidence of species-level divergence (Ward *et al*., 2005; April *et al*., 2011). Thus, although the current results emphasize extensive genetic heterogeneity in some taxa (in particular Himri fish), they are not per se a sufficient justification for changes at species level. Rather, these results should be considered as evidence of evolutionary complexity and further research should now integrate both mitochondrial and nuclear markers with detailed morphological and ecological data (Dayrat, 2005; Pante *et al*., 2015).

Strengths and limitations of GenBank-based analyses

The use of COX1 sequences obtained from GenBank is a major strength of this study, offering extensive geographic and taxonomic coverage without limitations imposed by newly collected samples. Expanding comparative phylogeography and taxonomic revisitations at a large scale are now reliant on public databases (Ratnasingham and Hebert, 2007).

However, the use of published data also carries with it limitations, such as an un even geographic sampling and, at some occasions, taxonomic misidentifications. Although strict selection criteria were used in this study to avoid such problems, these shortcomings must be considered when interpreting finer-scale phylogenetic structures (Harris, 2003; Nilsson *et al.,* 2006).

Implications for future research

The present study suggests that mitochondrial COX1 is a species-specific marker and an informative molecule for the regional phylogenetic analysis of Cyprinidae. Nevertheless, the weak support for some deep nodes as well as the intricate phylogenetic relationships of *Carasobarbus* fish indicate that multilocus analyses with inclusion of nuclear genes (and vouched specimens are necessary. These integrative studies are needed to resolve persistent taxonomic ambiguities in Middle eastern Cyprinidae and to delineate between population-level structure and true species divergence (He *et al*., 2008; Geiger *et al*., 2014).

Conclusion

In conclusion, COX1-based studies reported herein show a structured genetic profile among the Middle Eastern cyprinid fish and offers new insights into evolutionary linage of the Iraqi populations especially the *Carasobarbus* fish. Although mitochondrial data have little value in providing taxonomic certainty, the patterns we obtained provide a good molecular baseline for more integrative investigation on taxonomy and phylogeography elsewhere in the area.

References:

April, J.; Mayden, R.L.; Hanner, R.H. and Bernatchez, L. (2011). Genetic calibration of species diversity among North America's freshwater fishes. Mol. Ecol. Resour., 11(4): 777–792. https://doi.org/10.1111/j.1755-0998.2011.03019.x

Avise, J.C. (2000). Phylogeography: The History and Formation of Species. Harvard Univ. Press, Cambridge.

Bănărescu, P. (1992). Zoogeography of Fresh Waters, Vol. 2. Aula-Verlag, Wiesbaden. Dayrat, B. (2005). Towards integrative taxonomy. Trends Ecol. Evol., 20(9): 407-415. https://doi.org/10.1016/j.tree.2005.05.020

Geiger, M.F.; Herder, F.; Monaghan, M.T.; Almada, V.; Barbieri, R.; Bariche, M.; Berrebi, P.; Bohlen, J.; Casal-López, M., and Delmastro, G.B., (2014). Spatial heterogeneity in the Mediterranean biodiversity hotspot affects barcoding accuracy of its freshwater fishes. Mol. Phylogenet. Evol., 75, 121–132. https://doi.org/10.1016/j.ympev.2014.02.013

Harris, D.J. (2003). Can you bank on GenBank? Trends Ecol. Evol., 18(7), 317–319. https://doi.org/10.1016/S0169-5347(03)00119-2

He, S.; Mayden, R.L.; Wang, X.; Wang, W., and Tang, K.L. (2008). Molecular phylogenetics of the family Cyprinidae (Teleostei: Cypriniformes). Mol. Phylogenet. Evol., 49(3), 708–718. https://doi.org/10.1016/j.ympev.2008.09.005

Hebert, P.D.N.; Cywinska, A.; Ball, S.L., and deWaard, J.R. (2003). Biological identifications through DNA barcodes. Proc. R. Soc. B, 270 (1512): 313–321. https://doi.org/10.1098/rspb.2002.2218

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

Levin, B.A.; Freyhof, J.; Lajbner, Z.; Perea, S.; Abdoli, A.; Gaffaroğlu, M.; Özuluğ, M.; Rubenyan, H.R., and Salnikov, V.B. (2012). Phylogenetic relationships and biogeography of the genus Capoeta (Cyprinidae). J. Biogeogr., 39(6), 1106–1121. https://doi.org/10.1111/j.1365-2699.2012.02683.x

Nilsson, R.H.; Kristiansson, E.; Ryberg, M.; Hallenberg, N., and Larsson, K.H. (2006). Intraspecific ITS variability in the kingdom Fungi. Mol. Ecol., 15(5), 1287–1295. https://doi.org/10.1111/j.1365-294X.2006.02842.x

Pante, E.; Schoelinck, C., and Puillandre, N. (2015). From integrative taxonomy to species description: one step beyond. Syst. Biol., 64(1): 152–160. https://doi.org/10.1093/sysbio/syu083

Perea, S.; Böhme, M.; Zupančič, P.; Freyhof, J.; Šanda, R.; Özuluğ, M.; Abdoli, A., and Doadrio, I. (2010). Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean Leuciscinae (Cyprinidae). Mol. Phylogenet. Evol., 55(3): 1153–1164. https://doi.org/10.1016/j.ympev.2010.02.030

Ratnasingham, S. and Hebert, P.D.N. (2007). BOLD: The Barcode of Life Data System. Mol. Ecol. Notes, 7(3): 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x

Ward, R.D.; Zemlak, T.S.; Innes, B.H.; Last, P.R. and Hebert, P.D.N. (2005). DNA barcoding Australia's fish species. Philos. Trans. R. Soc. B, 360(1462): 1847–1857. https://doi.org/10.1098/rstb.2005.1716

إعادة تقييم التصنيف الجزيئي والعلاقات التطورية لعائلة الكارب في الشرق الأوسط المستنتجة من تسلسلات COX1 للميتوكوندريا المسترجعة من بنك الجينات (GenBank)

ماجد عبد العزيز بناي[1]، حنان عبد الزهرة المنصوري[1]، ندى مفيد البغدادي[2]، نجم محسن عزيز[2]، محمد عبد الرضا الدوغجي[1]، فوزي مصطفى حميد الخواجة[1]، زينب جواد عبيد[1]، زهراء كريم شاكر[1]

1 قسم الفقريات البحرية، مركز علوم البحار ، جامعة البصرة، العراق
2قسم الأحياء البحرية، مركز علوم البحار ، جامعة البصرة، العراق
*Corresponding Author E-mail: majidbannai65@gmail.com

تاريخ الإستلام: 2025/10/24 تاريخ القبول: 2026/01/23 تاريخ النشر : 2026/02/08

المستخلص

تُعدّ فصيلة الكارب (Cyprinidae) المجموعة الأكثر تنوعاً والأكثر تحدياً من الناحية التصنيفية بين أسماك المياه العذبة في الشرق الأوسط، حيث يصعب تحديد الأنواع نظراً للتشابه المورفولوجي وعدم استقرار التصنيف. تُقدّم هذه الورقة مراجعة تصنيفية جزيئية وتحليلًا تطوريًا لفصيلة الكارب في الشرق الأوسط استنادًا إلى تسلسلات جين COX1 للميتوكوندريا المُنزّلة من بنك الجينات (GenBank). تم تجميع وتحليل مجموعة بيانات كبيرة من تسلسلات COX1 المتاحة للعموم من أجناس ومواقع مختلفة من فصيلة الكارب في جميع أنحاء الشرق الأوسط. حُسبت المسافات الجينية والمحاذاة، وحُللت العلاقات التطورية باستخدام طريقة الانضمام الجواري (NJ) والتطور الأدنى مع التحقق من صحة بوتستراب. كشفت الأشجار التطورية الناتجة عن فروع محددة جيدًا تتوافق مع المجموعات التصنيفية المعروفة، ولكنها كشفت أيضًا عن عدم يقين في تحديد الأنواع وتجميع غير متوقع بين التصنيفات وثيقة الصلة. على الرغم من أنها تستند إلى علامة ميتوكوندرية واحدة، إلا أن البيانات توضح فائدة COX1 في حل العلاقات على مستوى الأنواع وتوفر رؤى أولية حول البنية الجغرافية الوراثية الضحلة بين أسماك الكارب في الشرق الأوسط. تشير هذه الأنماط إلى احتمالية وجود لبس تصنيفي، وتُبرز أهمية التحليلات التكاملية التي تتضمن مؤشرات أخرى وبيانات مورفولوجية. تُقدّم هذه الدراسة إطارًا جزيئيًا لتوضيح تصنيف أسماك الكارب في الشرق الأوسط، وتُشكّل أساسًا للبحوث التطورية والتصنيفية المستقبلية.

الكلمات المفتاحية: أسماك الكارب، COX1، الترميز الجيني، العلاقات التطورية، إعادة التقييم التصنيفي، الشرق الأوسط، بنك الجينات.