Received: 1 March 2022 | Revised: 29 March 2022 | Accepted: 11 April 2022

DOI: 10.1111/mec.16540

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

# Large-scale DNA barcoding of the subfamily Culterinae (Cypriniformes: Xenocyprididae) in East Asia unveils a geographical scale effect, taxonomic warnings and cryptic diversity

Weitao Chen[1,2,3] | Nicolas Hubert[4] | Yuefei Li[1,2,3] | Denggao Xiang[1] | Xingwei Cai[5] | Shuli Zhu[1,2,3] | Jiping Yang[1,2,3] | Chuanjiang Zhou[6] | Xinhui Li[1,2,3] | Jie Li[1,2,3]

[1]Pearl River Fisheries Research Institute, Chinese Academy of Fishery Science, Guangzhou, China

[2]Key Laboratory of Aquatic Animal Immune Technology of Guangdong Province, Guangzhou, China

[3]Guangzhou Scientific Observing and Experimental Station of National Fisheries Resources and Environment, Guangzhou, China

[4]Institut de Recherche pour le Développement, UMR 226 ISEM (UM-CNRS-IRD), Montpellier, France

[5]Hainan Academy of Ocean and Fisheries Sciences, Haikou, China

[6]College of Fisheries, Henan Normal University, Xinxiang, China

**Correspondence**
Nicolas Hubert, Institut de Recherche pour le Développement, UMR 226 ISEM (UM-CNRS-IRD), Montpellier, France.
Email: nicolas.hubert@ird.fr

Jie Li, Pearl River Fisheries Research Institute, Chinese Academy of Fishery Science, Guangzhou, Guangdong, China.
Email: lijie1561@163.com

**Funding information**
National Key R&D Program of China, Grant/Award Number: 2018YFD0900902 and 2018YFD0900903; Project of innovation team of survey and assessment of the Pearl River fishery resources, Grant/Award Number: 2020TD-10 and 2020ZJTD-04.

**Handling Editor:** Carla Lopes

## Abstract

Geographical scale might be expected to impact significantly the efficiency of DNA barcoding as spatially comprehensive sampling provides opportunities to uncover intricate relationships among closely related species and to detect cryptic diversity for widespread taxa. Here, we present a DNA barcoding study on a Xencyprididae subfamily (Culterinae) involving the production of 998 newly generated DNA barcodes from East Asian drainages (80 localities). Together with 513 barcodes mined from BOLD and GenBank, a reference library consisting of 1511 DNA barcodes (116 localities) for 42 species was assembled, accounting for 66% of known Culterinae species. Intraspecific genetic distances are positively correlated to geographical scale, while a negative correlation is detected between interspecific genetic distances and geographical scale. The present study demonstrates that geographical scale influences the efficiency of DNA barcoding by narrowing the width of the barcoding gap. DNA-based species delimitation analyses delimited 44 molecular operational taxonomic units (MOTUs). Rampant cryptic diversity is detected within eight species with multiple MOTUs, whereas 25 species present mismatch between morphological and molecular delimitations. A total of 18 species are lumped into nine MOTUs due to low interspecific divergence and/or mixed lineages. Several MOTU divergences are hypothesized to relate to known biogeographical barriers and geological events during the Pliocene and Pleistocene. This study provides new insights into the taxonomy and phylogeography of the subfamily Culterinae.

**KEYWORDS**
fisheries management, phylogeography, species delimitation, species divergence, species identification

## 1 | INTRODUCTION

DNA barcoding, proposed as a solution to overcome the rarefaction of taxonomists worldwide (i.e., taxonomic impediment), provides a reliable framework for sustainable species identification and assessment of neglected or poorly investigated taxonomic groups (April et al., 2011; Chen et al., 2015; Galimberti et al., 2012; Galimberti et al., 2021; Hebert et al., 2003; Shen et al., 2019). Since this approach, and associated data quality standards, was developed, a large number of empirical DNA barcoding studies have been conducted in various taxa (April et al., 2011; Chen et al., 2015; Delrieu-Trottin et al., 2019; Galimberti et al., 2012; Galimberti et al., 2021; Hebert et al., 2003; Hebert et al., 2004; Hubert et al., 2008; Hubert et al., 2012; Pereira et al., 2013; Shen et al., 2019; Sholihah et al., 2020). Most of these studies yielded a broad match between morphological identifications and DNA barcode clusters (Ball et al., 2005; Breman et al., 2016; Delrieu-Trottin et al., 2020; Hebert et al., 2003; Hebert et al., 2004; Kim et al., 2020; Lara et al., 2010; McCusker et al., 2013; Panprommin et al., 2019; Ribeiro et al., 2012; Smith et al., 2005), resulting in the detection of a barcoding gap, which is the lack of overlap between the distributions of the maximum intraspecific and minimun interspecific genetic distances (Meyer & Paulay, 2005). However, most previous reports focused on particular areas, such as a nation or a biogeographical province (Ball et al., 2005; Breman et al., 2016; Hebert et al., 2003; Hebert et al., 2004; Kim et al., 2020; Lara et al., 2010; McCusker et al., 2013; Panprommin et al., 2019; Ribeiro et al., 2012; Smith et al., 2005). If only restricted areas are involved, only a subset of global diversity of resident lineages is considered, leading to two shortcomings, which artificially inflate the width of the barcoding gap: (i) in spatially structured populations, only a fraction of intraspecific genetic diversity is sampled, and maximum intraspecific genetic distances are underestimated; and (ii) if allopatric speciation mostly contributed to species pools, sister-species are under-represented and minimum genetic distances between species are over-estimated (Bergsten et al., 2012). As such, regional species assemblages yield a greater match between morphological delimitations and DNA barcode clusters than entire lineages, a trend calling for broad-scale assessments in the case of highly diversified lineages.

The Xenocyprididae subfamily Culterinae (Ostariophysi: Cypriniformes) is a group of Asian endemic fishes (Figure 1), mostly residing in East Asian drainages (Froese & Pauly, 2020; Luo & Chen, 1998). A total of 64 recognized species, belonging to 18 genera, have been documented in this group (Dai et al., 2005). Among recognized species, many are economically important in East Asia (e.g., *Chanodichthys* spp., *Culter* spp., *Hemiculter* spp., *Parabramis* spp. and *Megalobrama* spp.) and are targeted by inland fisheries due to their medium to large size (Froese & Pauly, 2020; Luo & Chen, 1998). Additionally, several species such as *Anabarilius* spp., *Pogobrama barbatula*, *Hainania serrata*, *Ancherythroculter wangi* and *Megalobrama pellegrini* are on the list of endangered species in the "Red List of China's Vertebrates" (Jiang et al., 2016). As such, this subfamily has attracted much research interest in recent decades because of their conservation status (Cao & Wang, 2010; Feng et al., 2009; Jiang et al., 2018; Li et al., 1993; Zheng, 1989).

In the subfamily Culterinae, many closely related and morphologically similar species have been described (Froese & Pauly, 2020; Luo & Chen, 1998), and several cases of discordant results between morphological and genetic species delimitation have been observed (Luo, 2016; Wang et al., 2019; Xie et al., 2012), casting doubt on whether accurate species identification can be performed using DNA barcodes for this subfamily. For example, several morphologically different species within the genera *Megalobrama* or *Culter* were found to be nested within one lineage and/or share mitochondrial and nuclear haplotypes (Luo, 2016; Wang et al., 2019; Xie et al., 2012). Moreover, crossing experiments among the four

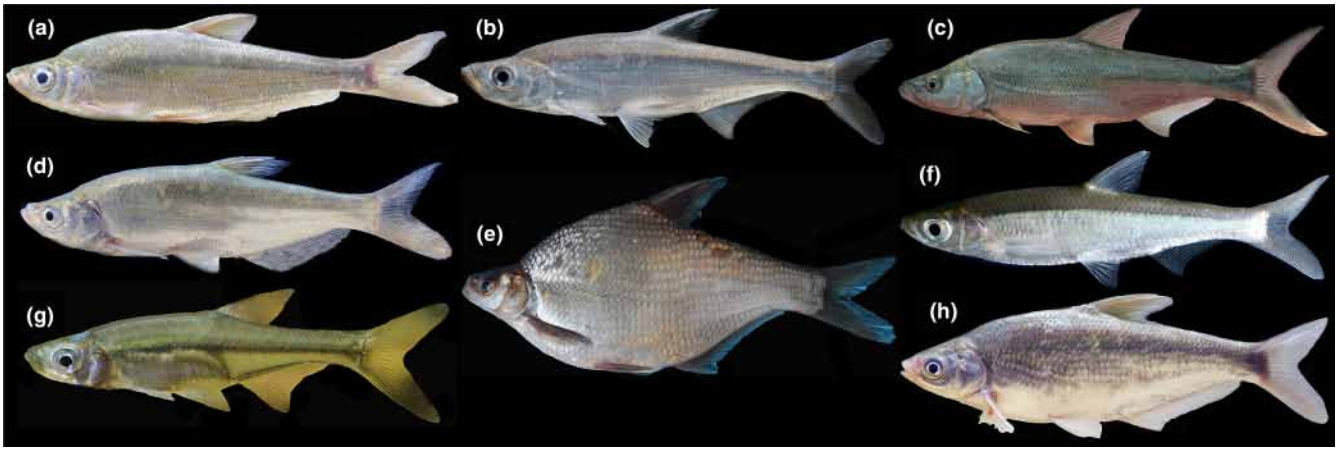


**FIGURE 1** Selected species of Culterinae, which illustrate the diversity of the subfamily. (a) *Ancherythroculter lini* Luo, 1994; maximum standard length (SL) of 19.1 cm. (b) *Culter recurviceps* (Richardson, 1846); maximum total length (TL) of 54.6 cm. (c) *Culter mongolicus* (Basilewsky, 1855); maximum SL of 100.0 cm. (d) *Chanodichthys erythropterus* (Basilewsky, 1855); maximum TL of 102.0 cm. (e) *Megalobrama pellegrini* (Tchang, 1930); maximum SL of 25.1 cm. (f) *Toxabramis houdemeri* Pellegrin, 1932; maximum SL of 14.8 cm. (g) *Metzia formosae* (Oshima, 1920); maximum SL of 7.3 cm. (h) *Parabramis pekinensis* (Basilewsky, 1855); maximum TL of 55.0 cm. Maximum SL and TL data from fishbase (Froese & Pauly, 2020)

*Megalobrama* species showed that they can hybridize (Zhang et al., 2014), introducing further difficulties in accurate molecular species identification.

Given some widespread Culterinae species are distributed across East Asia, a proportion of these species appear to form multiple hidden lineages under the influence of biogeographical barriers along with palaeogeological events. The most famous biogeographical barrier for freshwater fishes in China is the Nanling Mountains, which separates the Yangtze and the Pearl Rivers (Figure 2). This barrier has been demonstrated to have shaped distinct ichthyofauna in these two watersheds (Chen et al., 1986; Zheng, 1989) and to have led to a substantial number of cryptic lineages during the uplift of the Qinghai–Tibetan plateau and the evolution of the Asian monsoon (An et al., 2001; Chen et al., 2017; Perdices et al., 2005; Yang et al., 2009; Zhong et al., 2019). Other biogeographical barriers such as the Qiongzhou Strait and the Yunkai Mountains were demonstrated to have accelerated the divergence of local fish taxa in southern China (Chen et al., 2007; Chiang et al., 2013; Yang et al., 2012; Yang et al., 2016). Accordingly, if genetically structured populations are observed across those barriers, inflated intraspecific genetic divergences are expected for widespread taxa.

Here, we present the results of a large-scale DNA barcoding campaign for Culterinae in East Asia. A total of 1511 DNA barcodes from 116 locations, covering most important rivers in East Asia, were included. Furthermore, widely distributed species were sampled across different biogeographical regions, providing suitable candidates to probe biogeographical patterns. Thus, the goals of the present study were: (i) to assess the performance of DNA barcoding in capturing Culterinae species boundaries in large-scale sampling, (ii) to examine the impact of population genetic structure and geographical distance on the relative distribution of intraspecific and interspecific genetic distances, and (iii) to estimate the proportion of cryptic diversity and discuss its origin in the light of the biogeographical history of the region. Finally, this study aims to streamline the inventory of Culterinae species and promote further taxonomic studies.

## 2 | MATERIALS AND METHODS

### 2.1 | Sampling and identification

A total of 998 Culterinae specimens were collected between April 2010 and July 2020 across 80 sites that covered the Yangtze River, the Pearl River, the Yellow River, Heilongjiang River, the Huaihe River, and many independent and small coastal rivers (Figure 2; Table S1). Voucher specimens were morphologically identified using the monograph "Fauna Sinica Osteichthyes, Cypriniformes II, Ichthyography of the Pearl River and Freshwater Fishes of Guangxi, China" (Luo & Chen, 1998; Zheng, 1989; Zhou & Zhang, 2005). A fin clip or a muscle biopsy was taken from each specimen and fixed in 99% ethanol solution for further genetic analyses. Tissues and voucher specimens were deposited at the Pearl River Fisheries Research Institute, Chinese Academy of Fishery Science.

### 2.2 | DNA sequencing, data preprocessing and assembly

Genomic DNA was extracted using the Genomic DNA Isolation Kit (Axygen) according to the manufacturer's instructions. The standard 652-bp fragment from the 5′ end of the mitochondrial cytochrome *c* oxidase subunit I gene (COI) was amplified using the fish universal primers FishF1 and FishR1 (Ward et al., 2005). Amplifications included an initial denaturation step at 95°C for 5 min followed by 30 cycles of denaturation at 95°C for 1 min, annealing at 50°C for 45 s, and extension at 72°C for 45 s followed by a final extension at 72°C for 10 min. All amplified fragments were purified using 1.0% low-melting agarose gel electrophoresis and bidirectionally sequenced with the same primer pair using an ABI PRISM 3700 (Applied Biosystems) automatic DNA sequencer. COI contigs were checked and assembled using the SEQMEN procedure in the LASERGENE package (DNASTAR). Sequences were aligned and trimmed to the same length using MEGA version 6.0 (Tamura et al., 2013) and further translated into amino acids to check for the potential occurrence of stop codons. Aligned sequences, primer pairs, trace files, taxonomic information and collection data were deposited at the Barcode of Life Data system (BOLD; Ratnasingham & Hebert, 2013) in the project CUBAR "Barcodes of Culterinae species" (dx.doi.org/10.5883/DS-CUBAR). Alongside newly generated DNA barcodes, DNA barcode records belonging to Culterinae species were mined from GenBank and BOLD, where sample sources were available (Table S1) and sequences exceeded 600 bp.

### 2.3 | Genetic distance analyses

Using MEGA 6, we calculated Kimura 2-parameter (K2P) (Kimura, 1980) pairwise genetic distances within and among species. To check for the presence of a barcoding gap, species-level comparisons between maximum intraspecific genetic distances and minimum distances to the nearest neighbour (Meyer & Paulay, 2005) were performed in MEGA 6. To explore the effect of geographical scale on genetic distance estimates, individuals from multiple geographical locations were included. Linear regression and ANOVAs were performed using SPSS 18.0 to assess the relationship between species sample size and maximum intraspecific divergence, and between maximum spatial distance and maximum intraspecific divergence. The effect of geographical scale on the genetic distance to the closest species was investigated by creating five geographical distance categories, <1, <50, <200, <1000 and <5000 km. For each category, all interspecific genetic distances were calculated, and the minimum was recorded for each species. Finally, a neighbour-joining (NJ) tree was reconstructed for a visual inspection of the data, using K2P genetic distances with the "pairwise deletion" option in MEGA 6. We used

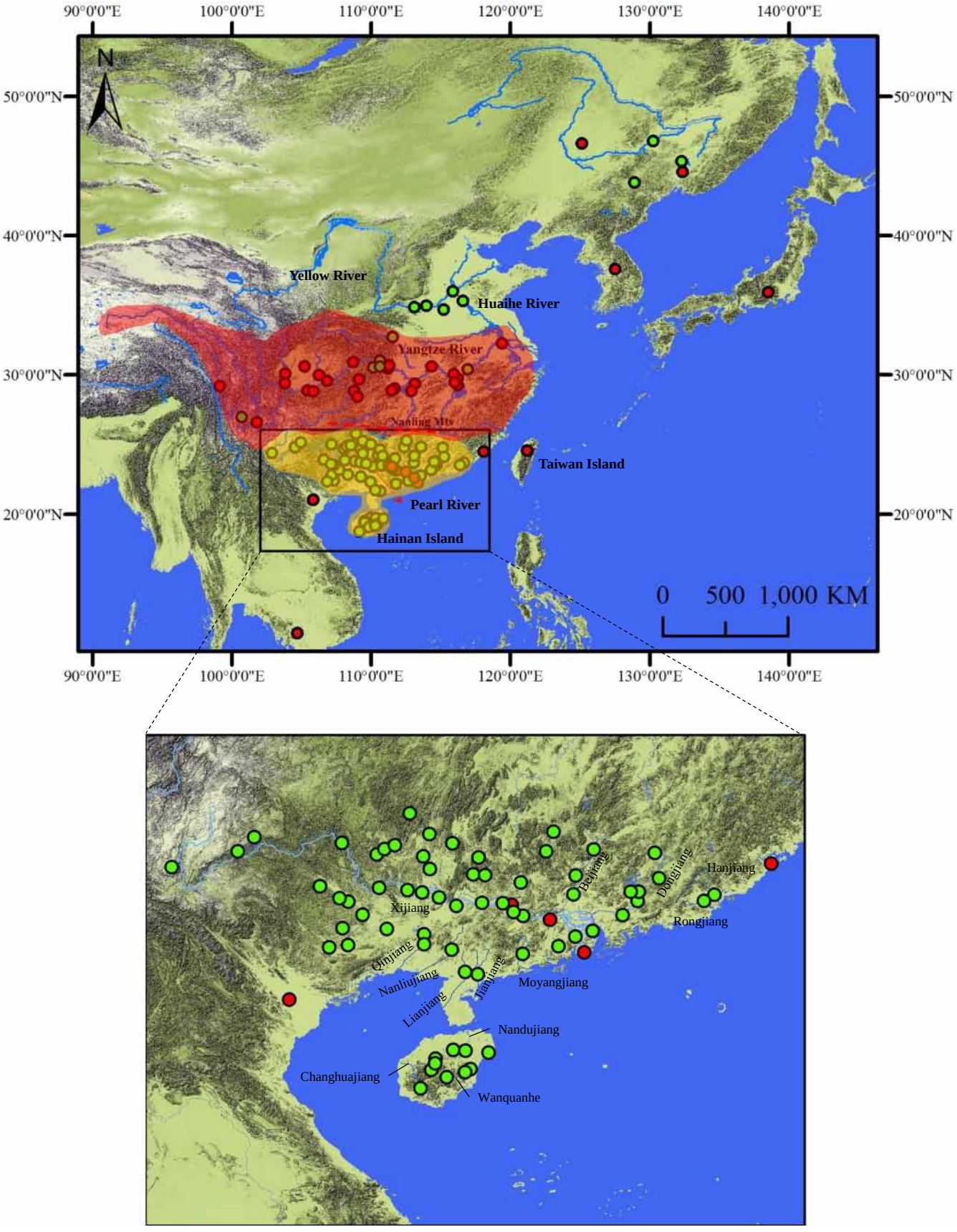


**FIGURE 2** Sample sites for the Culterinae species included in the present study. Details of the 116 sites and collection of DNA barcoded specimens are provided in Appendix S1. Green and red circles represent sampling sites for novel obtained barcodes and downloaded barcodes, respectively. Red and yellow shadows indicate the Yangtze and the Pearl River drainage, respectively. This map was created in ARCGIS version 10.2 (http://www.esri.com/arcgis/about-arcgis)

1000 bootstrap replicates to assess the branch support of the NJ tree.

## 2.4 | Sequence-based species delimitation

To group specimens into molecular operational taxonomic units (MOTUs), defined as diagnosable molecular lineages (Avise, 2012; Moritz, 1994), four sequence-based species delimitation methods were used, including Refined Single Linkage (RESL) (Ratnasingham & Hebert, 2013), Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012), Poisson Tree Process (PTP) (Kapli et al., 2017; Zhang et al., 2013), and the General Mixed Yule-coalescent (GMYC) (Fujisawa & Barraclough, 2013). A general delimitation scheme was established based a majority-rule consensus of the four algorithms used. For the sake of clarity, species identified by morphological features are referred to as species, and species delimited using DNA barcodes are referred to as MOTUs.

RESL analysis was used to produce Barcode Index Numbers (BINs) as implemented in BOLD. ABGD analysis was performed on the web interface (www.abi.snv.jussieu.fr/public/abgd/) using the default value for the relative gap width ($X = 1.5$) and K2P genetic distance. PTP in its multiple rates version was run using the web server (http://mptp.h-its.org). As the mPTP algorithm needs a phylogenetic tree as an input file, a maximum-likelihood (ML) tree was built using RAXML-VI-HPC (Stamatakis, 2006). ML analyses were conducted using the GTR + I + Γ model, according to the model selection procedure implemented in MRMODELTEST (Nylander, 2004). Single- and multiple-threshold GMYC analyses were conducted using a web server (http://species.h-its.org/gmyc/). As GMYC requires a fully resolved, ultrametric gene tree as input for the analysis, we reconstructed a Bayesian tree using BEAST version 1.8.1 (Drummond & Rambaut, 2007). Bayesian reconstructions were conducted using a Yule pure-birth model and GTR + I + Γ substitution model. Duplicated sequences were collapsed into haplotypes using ALTER (Glez-Pena et al., 2010) prior to running BEAST analyses. An uncorrelated relaxed lognormal clock model was used with the rate estimated from the data and ucld-mean parameters with a uniform prior value of 0 as a lower boundary and 10 as an upper boundary. All other settings were kept as the default values. Markov chain Monte Carlo (MCMC) runs were performed using 50 million generations, sampled every 1000 trees. Stability around unimodal posteriors and effective sample sizes (ESS) for all parameters were assessed using TRACER version 1.5 (Rambaut & Drummond, 2007). After removing 50% of the trees as burn-in, trees were summarized in a maximum credibility tree using TREEANNOTATOR version 1.8.1 (Drummond & Rambaut, 2007).

## 2.5 | Cryptic diversity and phylogeography

For any species pairs or groups of species with shallow genetic divergence, a network approach based on the median-joining algorithm was used to explore the relationships among haplotypes, as implemented in NETWORK 4.6 (Bandelt et al., 1999).

With respect to species dissolving into multiple MOTUs, an NJ tree was built for visual inspection using K2P distances, with 1000 bootstrap replicates, and lineages were projected onto a map. In addition, divergence times among MOTUs were estimated using BEAST 1.8.1 following the procedure mentioned previously. Trees were reconstructed using a strict-clock model, a substitution rate of 1.2% per million years (Bermingham et al., 1997), a Yule diversification model and GTR + I + Γ substitution model. Analyses were conducted with 100 million MCMC generations, sampled every 1000 generations. Trees were summarized with TREEANNOTATOR version 1.8, and the first 50% of sampled trees were removed.

# 3 | RESULTS

Within 998 newly acquired samples between 2010 and 2020, 26 Culterinae species, belonging to 13 genera, were delimited and identified based on their external morphology. Among them, 998 DNA barcodes were successfully produced. All the sequences were above 600 bp in length without stop codons or insertions, indicating collected sequences represent functional coding regions. Alongside newly generated DNA barcodes, 513 DNA barcodes belonging to 35 Culterinae species were mined from GenBank and BOLD (Table S1). After aligning newly generated and mined sequences, the final alignment consisted of 602 bp for 1511 specimens from 116 locations, covering most drainages in East Asia (Figure 2), and corresponding to 42 species and 15 genera (Table S1). The number of sequences per species ranged between one and 261, with an average of 34, originating from two to 38 sampling sites (Table 1).

Mean genetic divergence was 0.93% (0%–6.88%; standard error [SE] = 0.0009) within species and 3.51% (0.67%–5.82%; SE = 0.0034) between species within genera. Maximum intraspecific distances ranged between 0% and 10.78%, and nearest neighbour distances ranged from 0% to 16.43% (Table 1). In addition, eight species displayed maximum intraspecific distances above 2%, including *Hemiculter leucisculus* (5.56%), *Hemiculterella sauvagei* (3.07%), *Sinibrama melrosei* (7.67%), *Chanodichthys erythropterus* (7.07%), *Culter mongolicus* (3.96%), *Pseudohemiculter dispar* (10.78%), *Pseudolaubuca sinensis* (3.96%) and *Metzia formosae* (6.88%) (Figure S1; Table 1). Plotting maximum intraspecific and nearest neighbour K2P genetic distances revealed a lack of barcoding gaps with maximum intraspecific genetic distances exceeding minimum distances to the nearest neighbour in multiple cases (Figure 3a). Linear regressions of maximum intraspecific genetic distances as a function of the number of sampled individuals were significant ($R = 0.565$, $p = .001$), but had low explanatory power (adjusted $R^2 = .058$). Maximum intraspecific genetic distances also depended on geographical scale of sampling (adjusted $R^2 = 0.143$, $p = .019$) (Figure 3c). Genetic distances to the nearest neighbour declined from an average of 5.10% to 2.81% as the geographical range

**TABLE 1** Summary statistics of the number of individuals per species ($N$), the number of sampling sites per species ($N_P$), maximum intraspecific genetic distance ($D_M$), genetic distance to the closest neighbour ($D_N$) and the geographical extent of sampling in kilometres (DIS)

| Species | $N$ | $N_P$ | $D_M$ | $D_N$ | DIS |
|---|---|---|---|---|---|
| *Anabarilius brevianalis* | 1 | 1 | NA | 0.0468 | NA |
| *Anabarilius grahami* | 3 | 1 | 0.0033 | 0.0698 | 0 |
| *Anabarilius liui chenghaiensis* | 2 | 1 | 0.000 | 0.0101 | 0 |
| *Anabarilius liui yalongensis* | 1 | 1 | NA | 0.0101 | NA |
| *Ancherythroculter kurematsui* | 2 | 1 | 0.0000 | 0.0017 | 0 |
| *Ancherythroculter lini* | 38 | 12 | 0.0050 | 0.0441 | 713 |
| *Ancherythroculter nigrocauda* | 15 | 3 | 0.0017 | 0.0017 | 536 |
| *Ancherythroculter wangi* | 1 | 1 | NA | 0.0017 | NA |
| *Chanodichthys compressocorpus* | 1 | 1 | NA | 0.0325 | NA |
| *Chanodichthys erythropterus* | 103 | 20 | 0.0707 | 0.0325 | 3526 |
| *Culter alburnus* | 133 | 28 | 0.0118 | 0.0113 | 1646 |
| *Culter dabryi* | 23 | 7 | 0.0033 | 0.0017 | 2829 |
| *Culter mongolicus* | 54 | 14 | 0.0396 | 0.0196 | 3560 |
| *Culter oxycephaloides* | 1 | 1 | NA | 0.0311 | 0 |
| *Culter recurviceps* | 261 | 29 | 0.0169 | 0.0113 | 877 |
| *Hemiculter bleekeri* | 70 | 9 | 0.0101 | 0.0036 | 2911 |
| *Hemiculter leucisculus* | 220 | 38 | 0.0556 | 0.0266 | 3657 |
| *Hemiculter lucidus* | 6 | 1 | 0.0050 | 0.0036 | 0 |
| *Hemiculter tchangi* | 20 | 3 | 0.0067 | 0.0332 | 136 |
| *Hemiculterella sauvagei* | 36 | 8 | 0.0307 | 0.0542 | 855 |
| *Hemiculterella wui* | 24 | 5 | 0.0017 | 0.0643 | 581 |
| *Ischikauia steenackeri* | 1 | 1 | NA | 0.0577 | NA |
| *Macrochirichthys macrochirus* | 1 | 1 | NA | 0.1643 | NA |
| *Megalobrama amblycephala* | 28 | 8 | 0.0118 | 0.0057 | 1076 |
| *Megalobrama pellegrini* | 11 | 2 | 0.0000 | 0 | 156 |
| *Megalobrama skolkovii* | 8 | 4 | 0.0000 | 0 | 301 |
| *Megalobrama terminalis* | 48 | 10 | 0.0033 | 0.0299 | 836 |
| *Metzia formosae* | 2 | 2 | 0.0688 | 0.0368 | 1624 |
| *Metzia lineata* | 11 | 5 | 0.0067 | 0.0242 | 477 |
| *Metzia longinasus* | 1 | 1 | NA | 0.0367 | NA |
| *Metzia mesembrinum* | 1 | 1 | NA | 0.0242 | NA |
| *Parabramis pekinensis* | 47 | 11 | 0.0050 | 0.0471 | 1656 |
| *Pseudohemiculter dispar* | 57 | 14 | 0.1078 | 0.0913 | 836 |
| *Pseudohemiculter hainanensis* | 8 | 3 | 0.0135 | 0.0542 | 688 |
| *Pseudolaubuca engraulis* | 25 | 11 | 0.010 | 0.0892 | 1259 |
| *Pseudolaubuca sinensis* | 103 | 15 | 0.0396 | 0.0984 | 1220 |
| *Sinibrama macrops* | 21 | 6 | 0.0017 | 0.0381 | 962 |
| *Sinibrama melrosei* | 35 | 11 | 0.0767 | 0.0652 | 748 |
| *Sinibrama taeniatus* | 10 | 1 | 0.0151 | 0.0346 | 0 |
| *Sinibrama wui* | 1 | 1 | NA | 0.0381 | NA |
| *Toxabramis houdemeri* | 74 | 15 | 0.0084 | 0.0646 | 771 |
| *Toxabramis swinhonis* | 3 | 1 | 0.0000 | 0.0266 | 0 |

of sampling was increased from <1 to <5000 km (Figure 3d), with a significant effect (ANOVA; df = 3.934, $p$ = .005).

A visual inspection of the NJ tree based on K2P genetic distances indicated that 18 of 42 species analysed displayed concordant clusters of DNA barcodes (Figure S1). The 24 remaining species encompass multiple lineages, either lumped or mixed due to haplotype sharing. MOTU delimitation analyses yielded varying numbers of MOTUs according to methods with 41, 45, 36 and 53 MOTUs

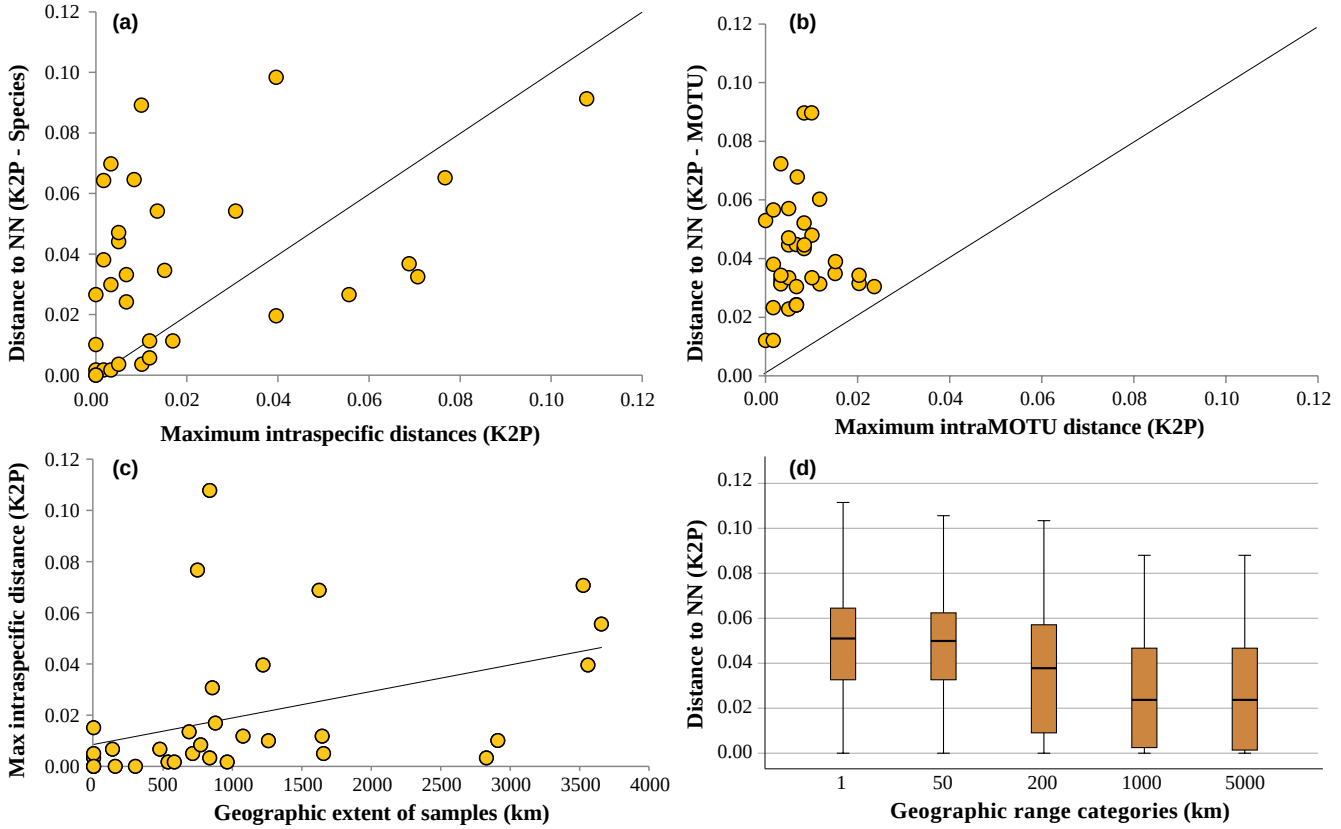


**FIGURE 3** Distribution of genetic distances within and between species and MOTUs, and relationships to geographical scales. (a) Relationship between maximum intraspecific and nearest-neighbour genetic distance among species. (b) Relationship between maximum intraspecific and nearest-neighbour genetic distance among MOTUs. Lines indicate the 1:1 ratio. (c) Maximum intraspecific variation against maximum geographical extent (km) of sampled individuals (adjusted $R^2 = 0.143$, $p = .019$). (d) the effect of geographical scale of sampling on the closest interspecific divergence. Minimum interspecific divergence across species in five distance categories. In each category, all interspecific distances between individuals with a pairwise geographical distance of less than the category value were calculated. Genetic distance was significantly smaller in the 5000-km category compared with 1- and 50-km categories (one-way ANOVA, Tukey's HSD, $p < .05$)

delimited by RESL, ABGD, PTP and GMYC, respectively (Figure 4). MOTU number was underestimated by RESL as BINs were not available for two sequences corresponding to two distinct species in BOLD (Table S1). The final consensus, established by a majority-rule consensus, consisted of 44 MOTUs. A total of 17 species were unambiguously delimited by the delimitation analyses, eight species displayed multiple MOTUs (Table 2) and 18 species displayed mixed genealogies with nine MOTUs shared by more than one species (Figure 4; Table 3). Distributions of both maximum intraspecific distances and distances to the nearest neighbour separated all observed MOTUs (Figure 3b).

Haplotype networks were reconstructed for 18 species displaying shallow genetic divergence and/or haplotype sharing (Figure 5). Five cases of haplotype sharing were observed between *Culter alburnus* and *Cu. mongolicus* (Figure 5a), *Ancherythroculter nigrocauda* and *Culter dabryi* (Figure 5b), *Chanodichthys erythropterus* and *Ch. compressocorpus* (Figure 5d), *Hemiculter bleekeri* and *H. lucidus* (Figure 5e), and *Megalobrama skolkovii* and *M. pellegrini* (Figure 5g). For each of these five cases, haplotype sharing involves ancestral haplotypes, as revealed by their central position in haplotype network reconstructions. Scattered haplotypes across haplotype networks were observed for *Ch. erythropterus* (Figure 5a), with haplotypes originating from Hainan Island being more closely related to haplotypes of *Culter recurviceps* than conspecific haplotypes sampled in the drainages of China's mainland.

NJ trees built for species displaying multiple, and deeply diverging, MOTUs revealed a diversity of patterns. Two lineages were detected in *Hemiculter leucisculus*, *Ch. erythropterus*, *Pseudolaubuca sinensis*, *Hemiculturella sauvagei*, *Pseudohemiculter hainanensis* and *Metzia formosae* and four lineages were observed in *Pseudohemiculter dispar* and *Sinibrama melrosei* (Figure 6). In most cases of two MOTUs, a south–north differentiation was observed, as exemplified by *Ch. erythropterus*, *Hemiculter leucisculus*, *Hemiculturella sauvagei* and *Pseudolaubuca sinensis*, with one lineage distributed in the southern drainages including Hainan Island, Pearl River, and some coastal rivers in southern China, and another lineage largely occurring in central and northern drainages such as the Yangtze River and/or its northern drainages (Figure 6). However, the co-occurrence of haplotypes from distinct MOTUs is observed in some tributaries of the Yangtze and the Pearl River (Figure 6a–d). Alternatively, allopatric distributions of conspecific MOTUs involved different patterns with Vietnam vs. Taiwan distribution of MOTUs for *M. formosae*,

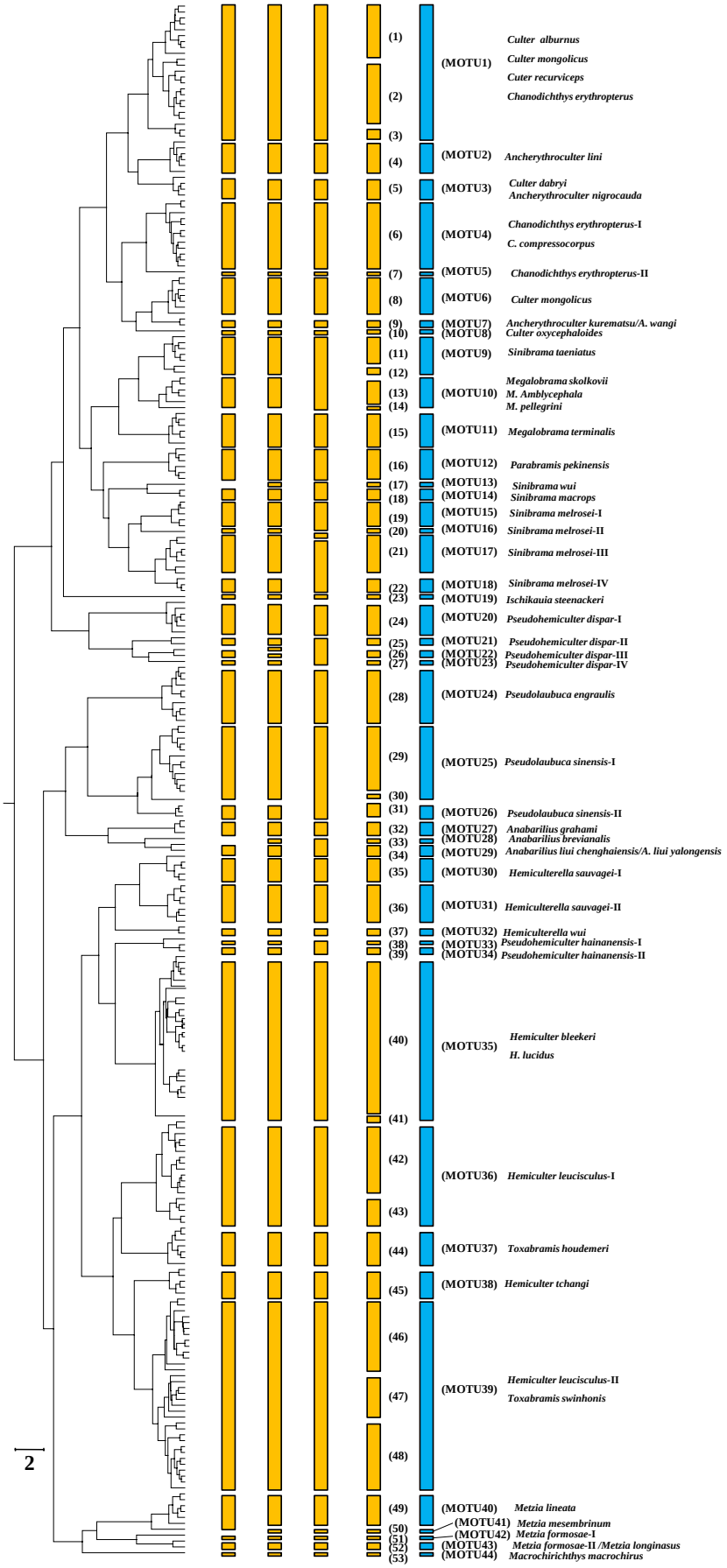


**FIGURE 4** DNA-based species delimitation within the 42 Culterinae species in this study. MOTU delimitation schemes collected from the RESL, ABGD, PTP and GMYC algorithms (yellow), including the consensus delimitation scheme (blue). A Bayesian tree produced in BEAST version 1.8.2 is used as a base to summarize delimitation results. Scale bar corresponds to Million years

**TABLE 2** List of species with multiple MOTUs including their barcode index number in BOLD, maximum genetic distances within an MOTU ($D_M$), and distance to their nearest MOTU ($D_N$). The nearest-neighbour distance for each MOTU corresponds to the distance to the closest MOTU in this study

| | $D_M$ | $D_N$ |
|---|---|---|
| ***Chanodichthys erythropterus*** | | |
| MOUT4 (BIN:ACU0863) | 0.0067 | 0.0448 |
| MOUT5 (BIN:ADW3458) | 0.0000 | 0.0530 |
| ***Hemiculter leucisculus*** | | |
| MOUT36 (BIN:ADM2843) | 0.0152 | 0.0390 |
| MOUT39 (BIN:ACB5189) | 0.0236 | 0.0305 |
| ***Hemiculterella sauvagei*** | | |
| MOUT30 (BIN:ADG7506) | 0.0067 | 0.0242 |
| MOUT31 (BIN:ADC8405) | 0.0067 | 0.0242 |
| ***Metzia formosae*** | | |
| MOUT42 (BIN:ADC6157) | NA | 0.0678 |
| MOUT43 (BIN:ADC6328) | 0.0069 | 0.0678 |
| ***Pseudohemiculter dispar*** | | |
| MOUT 20 (BIN:AEE2041) | 0.0084 | 0.0897 |
| MOUT 21 (BIN:AEH4884) | 0.0084 | 0.0521 |
| MOUT 22 (BIN:AEH7332) | 0.0084 | 0.0434 |
| MOUT 23 (BIN:AEH7333) | NA | 0.0434 |
| ***Pseudohemiculter hainanensis*** | | |
| MOUT33 (BIN:AEH4806) | 0.0000 | 0.0121 |
| MOUT34 (BIN:AEH4807) | 0.0017 | 0.0121 |
| ***Pseudolaubuca sinensis*** | | |
| MOUT25 (BIN:ADC6253) | 0.0203 | 0.0343 |
| MOUT26 (BIN:ACU0850) | 0.0033 | 0.0343 |
| ***Sinibrama melrosei*** | | |
| MOUT15(BIN:AEH7279) | 0.0050 | 0.0571 |
| MOUT16 (BIN:AEH3997) | NA | 0.0567 |
| MOUT17 (BIN:AEH3998) | 0.0101 | 0.0335 |
| MOUT18 (BIN:AEH8015) | 0.0050 | 0.0335 |

continental vs. Hainan Island MOTUs for *Pseudohemiculter hainanensis*, and erratic distribution in the case of *Pseudohemiculter dispar* and *S. melrosei* (Figure 6). Divergence time estimates suggest most MOTU divergence events originated during the Pleistocene, and several exceptions are detected within *Pseudohemiculter dispar*, *S. melrosei* and *M. formosae* with MOTU divergence dated back to the late Pliocene (Figure S2).

## 4 | DISCUSSION

The present study provides the first comprehensive DNA barcode reference library of the subfamily Culterinae, including 42 species (~66% of known Culterinae species) and covering most major East Asian watersheds (Luo & Chen, 1998). Notably, all *Anabarilius* species and several monotypic genera such as *Pogobrama* and *Hainania* are missing. These species have very restricted distribution ranges and little information is available regarding their abundance and ecology (Jiang et al., 2016; Luo & Chen, 1998). In addition, DNA sequences are scarce for these species in international repositories such as GenBank, EMBL and BOLD. By contrast, most widespread species were aptly covered, with multiple individual sequences originating from different East Asian drainages, which allows assessment of the impact of geographical scale on the performance of DNA barcoding in this subfamily.

### 4.1 | Shallow divergence, haplotype sharing and taxonomic confusion

DNA-based species delimitation analyses converged with morphological delimitation in only 17 species, indicating COI sequences fail at capturing species boundaries in multiple Culterinae. This relatively low success rate (~40%) was unexpected, considering the large number of empirical studies which yielded better matches between morphological and DNA-based delimitations (>90%) than here (April et al., 2011; Dahruddin et al., 2017; Delrieu-Trottin et al., 2019; Hou et al., 2018; Hubert et al., 2008; Keskin & Atar, 2013; Lara et al., 2010; Pereira et al., 2013; Ribeiro et al., 2012; Shen et al., 2016; Sholihah et al., 2020; Sonet et al., 2019; Ward et al., 2005). The lower success rate observed here is probably the result of the broad spatial scale involved and high coverage of closely related species in poorly known taxa. This is exemplified by the lower mean interspecific distance (3.51%) than observed in previous studies focused on fish families or subfamilies within narrow geographical areas (Adeoba et al., 2018; Alcantara & Yambot, 2016; Hubert et al., 2010; Jaafar et al., 2012; Tzeng & Chiu, 2012; Zhang et al., 2017). This trend is further supported by the statistically significant effect of geographical distance on both maximum intraspecific genetic distance (positively correlated) and genetic distance to the nearest neighbour (negatively correlated) (Figure 3). This result suggests the presence of a substantial proportion of recently diverged species among Culterinae species (Bellafronte et al., 2013), a trend further supported by the discovery of 18 cases of mixed interspecific genealogies (*Culter* spp., *Chanodichthys* spp., *Hemiculter* spp., *Megalobrama* spp., *Metzia* spp. and *Toxabramis swinhonis*). However, mismatches between morphological and DNA-based species delimitations may arise from introgressive hybridization and incomplete lineage sorting, but taxonomic confusion and erroneous identifications cannot be discarded (Dahruddin et al., 2017; Durand et al., 2017; Hubert & Hanner, 2015; Sholihah et al., 2020; Sonet et al., 2019).

In the genus *Megalobrama*, three species (i.e., *M. amblycephala*, *M. skolkovii*, and *M. pellegrini*) are molecularly indistinguishable, as they share the same mitochondrial lineage (MOTU10, Figure 3), and haplotype sharing was uncovered between *M. skolkovii* and *M. pellegrini* (Figure 5g). These close genetic relationships were previously mentioned (Bai et al., 2015; Xie et al., 2012), but no haplotype sharing was observed at the mitochondrial ND2 gene

| | $D_M$ | $D_N$ | Species |
|---|---|---|---|
| MOTU1 | 0.0203 | 0.0315 | *Culter alburnus*, *Culter mongolicus*, *Cuter recurviceps*, *Chanodichthys erythropterus* |
| MOTU3 | 0.0033 | 0.0328 | *Ancherythroculter nigrocauda*, *Culter dabryi*, |
| MOTU4 | 0.0067 | 0.0448 | *Chanodichthys compressocorpus*, *Chanodichthys erythropterus* |
| MOTU7 | 0.0017 | 0.0233 | *Ancherythroculter kurematsu*, *Ancherythroculter wangi* |
| MOTU10 | 0.0118 | 0.0314 | *Megalobrama Amblycephala*, *Megalobrama pellegrini*, *Megalobrama skolkovii* |
| MOTU29 | 0.0101 | 0.0480 | *Anabarilius liui chenghaiensis*, *Anabarilius liui yalongensi* |
| MOTU35 | 0.0117 | 0.0602 | *Hemiculter bleekeri*, *Hemiculter lucidus* |
| MOTU39 | 0.0236 | 0.0305 | *Hemiculter leucisculus*, *Toxabramis swinhonis* |
| MOTU43 | 0.0069 | 0.0678 | *Metzia formosae*, *Metzia longinasus* |

**TABLE 3** List of MOTUs shared by multiple species, including their maximum genetic distances within an MOTU ($D_M$), distance to their nearest MOTU ($D_N$), and a list of species detected for each MOTU. The nearest-neighbour distance for each MOTU corresponds to the distance to the closest MOTU without considering species boundaries

(Bai et al., 2015), suggesting this lack of resolution results from the lower substitution rate at the COI than ND2 gene. Interspecific hybridization experiments demonstrated that *Megalobrama* species can hybridize (Zhang et al., 2014), but *M. skolkovii* and *M. pellegrini* have allopatric distributions. *M. skolkovii* occurs in the middle and lower Yangtze River, and the Pearl River, and *M. pellegrini* occurs in the upper Yangtze River (Luo & Chen, 1998). A similar pattern is observed within the genera *Culter*, *Hemiculter*, *Ancherythroculter* and *Chanodichthys* with haplotype sharing occurring between species pairs (*Cu. mongolicus* and *Cu. recurviceps*, Figure 5a; *Ancherythroculter nigricauda* and *Culter dabryi*, Figure 5b; *Chanodichthys erythropterus* and *Ch. compressocorpus*, Figure 5d; *Hemiculter bleekeri* and *Hemuculter lucidus*, Figure 5e). In these particular cases, shared haplotypes have a central position in reconstructed haplotype networks, suggesting these are ancestral haplotypes. In addition, no haplotype sharing is detected among recently diverged haplotypes, suggesting that recent divergence of mitochondrial lineages is more probably responsible for this shortcoming than introgressive hybridization (Hubert & Hanner, 2015). Along the same line, taxonomic gaps or erroneous identifications are unlikely because of the presence of diagnostic morphological characters such as mouth shape, ventral ridge, chamber number and body type for most of the species pairs involved in haplotype sharing (Luo & Chen, 1998). This is particularly evident in *Ancherythroculter nigrocauda* and *Cu. dabryi*, which have been assigned to distinct genera due to different numbers of chamber in the swim bladder, *Ancherythroculter* spp. and *Culter* spp. having two and three chambers, respectively (Ding, 1994; Luo & Chen, 1998; Zheng, 1989). By contrast, *Hemiculter bleekeri* and *Hemuculter lucidus* are distinguished by quantitative characters including head–body length ratio (Luo & Chen, 1998). Moreover, *Hemuculter lucidus* was previously considered as a subspecies of *Hemiculter bleekeri* due to their morphological similarity (Luo & Chen, 1998; Wu & Yi, 1959). Considering that there is no overlap in the distribution of the two species (Luo & Chen, 1998), taxonomic confusion is more likely in this case and the two species should probably be regarded as a single species.

## 4.2 | Cryptic diversity and phylogeographical patterns

The mean level of intraspecific divergence observed here (0.93%) is higher than most of those previously reported in DNA barcoding studies focusing on fish families or subfamilies (0.75%) (Adeoba et al., 2018; Hubert et al., 2010; Jaafar et al., 2012; Laskar et al., 2013; Shen et al., 2016; Tzeng & Chiu, 2012; Zhang et al., 2017). The high intraspecific divergence observed in Culterinae is probably explained by the large geographical scale of the present study, widespread species being mostly spatially structured (Avise, 2000; Bergsten et al., 2012; Nekola & White, 1999; Wright, 1943), a trend previously observed in other Cypriniformes subfamilies such as Rasborinae in Southeast Asia (Sholihah et al., 2020). Several cases of multiple, highly divergent intraspecific MOTUs were detected in eight species, including *Chanodichthys erythropterus*, *Hemiculter leucisculus*, *Hemiculterella sauvagei*, *Metzia formosae*, *Pseudolaubuca sinensis*, *Pseudohemiculter* spp. and *Sinibrama melrosei* (Figure 4). Most intraspecific lineage divergences are dated to between the Pliocene and Pleistocene (Figure S2), suggesting the influence of historical geological events.

In *Ch. erythropterus*, *Hemiculter leucisculus*, *Hemiculturella sauvagei* and *Pseudolaubuca sinensis*, high divergence (>3.2% maximum intraspecific genetic distance) between MOTUs is observed. For *Ch. erythropterus* and *Hemiculter leucisculus*, MOTUs display alternative geographical distributions in the southern coastal watersheds, or the Pearl River and the Yangtze River. Similar distribution patterns were previously reported in phylogeographical studies of *Hemiculter leucisculus* and *Opsariichthys bidens* (Chen et al., 2017; Perdices et al., 2005). For *Hemiculturella sauvagei* and *Pseudolaubuca sinensis*,

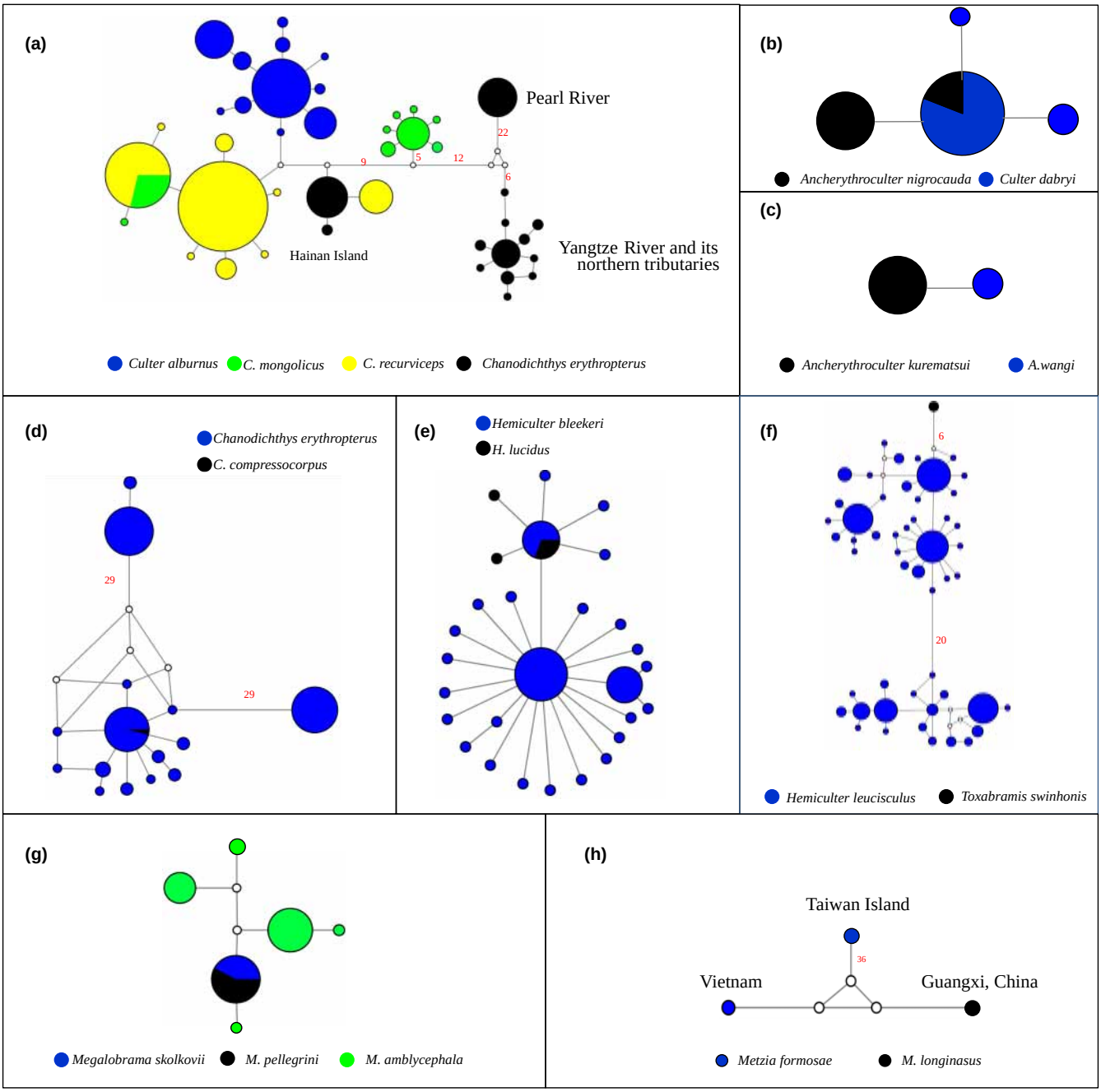


FIGURE 5 Median-joining networks of DNA barcode haplotypes for the species groups involved in mixed genealogies and/or haplotype sharing. (a) Three *Culter* species and *Chanodichthys erythropterus*. (b) *Ancherythroculter nigrocauda* and *Culter dabryi*. (c) *Ancherythroculter kurematsui* and *A. wangi*. (d) *Ch. Erythropterus* and *Ch. Compressocorpus*. (e) *Hemiculter bleekeri* and *Hemuculter lucidus*. (f) *Hemiculter leucisculus* and *Toxabramis swinhonis*. (g) Three *Megalobrama* species. (h) *Metzia formosae* and *M. longinasus*. Empty circles and numbers indicate unsampled haplotypes and mutation steps, respectively

two MOTUs were identified from the Pearl River and Yangtze River sampling sites, with one MOTU in the Yangtze and the other in both the Pearl River and southern tributaries of the Yangtze River (Figure 6). These distribution patterns are consistent with the influence of the Nanling Mountains, elevation of which created a biogeographical barrier separating the Yangtze River and the Pearl River, a scenario previously proposed by phylogeographical studies of multiple fish taxa, such as *Hemiculter leucisculus* (Chen et al., 2017), *O. bidens* (Li et al., 2009; Perdices et al., 2005), *Glyptothorax fokiensis* (Chen et al., 2007) and *Siniperca scherzeri* (Wang et al., 2006). Elevation of the Nanling Mountains occurred during uplift of the Qinghai–Tibetan plateau between 3.6 and 0.15 million years ago (Ma), a time frame matching our divergence time estimates among these MOTUs, dated at 1.36–2.32 Ma (Figure S2). Interactions between geology and palaeoclimatic fluctuations are also likely, as the Asian monsoon systems became established at 3.6–2.6 Ma. These climatic

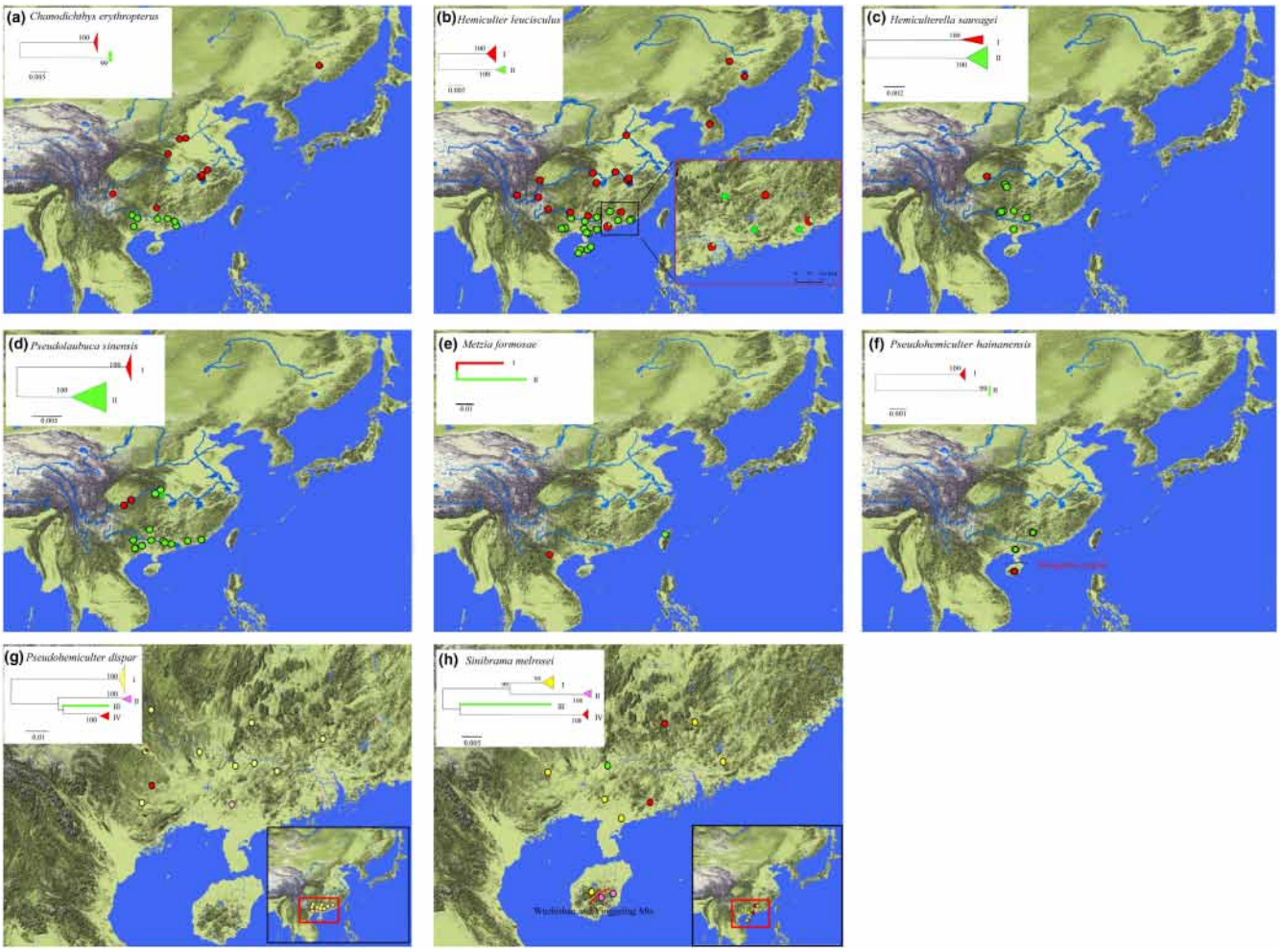


**FIGURE 6** Phylogeographical patterns of eight species with multiple lineages and/or MOTUs. (a) *Chanodichthys erythropterus*. (b) *Hemiculter leucisculus*. (c) *Hemiculterella sauvagei*. (d) *Pseudolaubuca sinensis*. (e) *Metzia formosae*. (f) *Pseudohemiculter hainanensis*. (g) *Pseudohemiculter dispar*. (h) *Sinibrama melrosei*. Trees at the top left of each map are neighbour-joining trees of the corresponding species K2P model. Different colour circles represent different lineages and/or MOTUs. Scale bars correspond to K2P genetic distances

and geological events probably drove landscape features and drainage systems of eastern Asia, promoting geographical isolation and subsequent divergence in many fish species (Chen et al., 2007; Chen et al., 2017; Guo et al., 2006; He et al., 2001; Perdices et al., 2005; Wang, 2005; Yang et al., 2009; Yu et al., 2014). In addition, several sequences of *Hemiculturella sauvagei* and *Pseudolaubuca sinensis* in some southern tributaries of the Yangtze River were grouped with the Pearl River lineage, and many sequences of *Hemiculter leucisculus* and *Ch. erythropterus* from the northern tributaries of the Pearl River were clustered with sequences of the Yangtze River and its northern tributaries. This is a common pattern observed in many fish taxa distributed in neighbouring tributaries of these two river basins (Chen et al., 2007; Chen et al., 2017; Perdices et al., 2005; Wang, 2005; Yang et al., 2009) due to intricate landscape changes involving watershed reconfiguration through headwater capture events, which may have thereby promoted fish dispersal by temporarily connecting upper branches of the Pearl and Yangtze Rivers (Ren et al., 1959).

With regard to *Pseudohemiculter dispar* and *Sinibrama melrosei*, four observed lineages were detected and did not show marked geographical distribution patterns (Figure 6g,h), suggesting these two species may have initially undergone allopatric divergence, followed by secondary contacts due to drainage rearrangements occurring in southern China. The above-mentioned climatic and geological events during the Pliocene and Pleistocene may have contributed to generate multiple MOTUs. In addition, previous studies demonstrated that geomorphological changes of drainage networks in response to the late Pleistocene glaciations had occurred in southern China (Liu et al., 2003; Wang et al., 1999), providing many opportunities for fish dispersal.

A repeated pattern of alternative distribution of MOTUs or haplotypes was also observed between Hainan Island and the mainland in *Pseudohemiculter hainanensis* (Figure 6f). This pattern suggests the formation of the Qiongzhou Strait triggered the divergence of populations within *Pseudohemiculter hainanensis* during the mid-Pleistocene (~0.53 Ma). The influence of this strait on population divergence has been previously suggested for several freshwater fish species such as *Garra orientalis* (Yang et al., 2016), *Micronemacheilus pulcher* (Qiu et al., 2008), *Megalobrama terminalis* (Chen et al., 2020) and *Glyptothorax hainanensis* (Chen et al., 2007).

### 4.3 | Implications for Culterinae taxonomy and future prospects

Our study provides the most comprehensive geographical, taxonomic and molecular sampling of this subfamily to date, with several findings contradicting current systematic and taxonomic knowledge. The most challenging cases were those involving shallow interspecific divergence between species, a trend shared by 18 species. Given that most of these cases may arise from recent divergence during the Pleistocene, additional molecular markers with higher substitution rates, such as the mitochondrial ND2 gene and control region, would certainly help in better distinguishing cases of recent divergence from actual haplotype sharing. Nevertheless, haplotype sharing was also observed between *Ancherythroculter nigrocauda* and *Culter dabryi*, for instance, calling for a broader assessment of the molecular species boundaries using biparentally inherited nuclear loci, and resulting in the revision of the generic status of the two species. With regard to *Hemiculter bleekeri* and *Hemuculter lucidus*, the relevance of the head–body length ratio in distinguishing each species should be reconsidered. However, we cannot rule out the occurrence of introgressive hybridization between these closely related species. Introgression is known to distort the congruence between morphological and DNA-based species delimitations (Petit & Excoffier, 2009), and constitutes a challenge for species identification based on DNA barcodes. Using additional nuclear loci in combination with COI in a spatially explicit context will probably shed light on species identity and ongoing evolutionary dynamics driving their genetic divergence (Oliveira et al., 2010; Raupach et al., 2010).

At the other end of the spectrum, the detection of multiple, and highly divergent, MOTUs in eight species suggests Culterinae diversity is currently underestimated and potentially new Culterinae species are awaiting a formal description. This pattern can be explained by overlooked, subtle morphological differences, but cryptic diversity and unrecognized speciation events might also be responsible for this result (Hebert et al., 2004; Hubert et al., 2012). The taxonomic status of these cryptic lineages delimited by mitochondrial sequences is unstable (Satler et al., 2013). These putative species constitute primary species hypotheses awaiting further assessments of their morphological divergence and phylogenetic placement using additional nuclear loci to produce secondary species hypotheses and robust species descriptions.

## 5 | CONCLUSION

The present large-scale genetic study in East Asia demonstrates the limit of DNA barcoding for automated identification of Culterinae species, as DNA barcodes and morphological characters are congruent for only 17 species. The 18 cases belonging to a complex of closely related species currently challenge the implementation of routine molecular species identification, but also question the evolutionary origin of a substantial number of Culterinae species. Moreover, newly detected mitochondrial lineages in nine species suggest Culterinae diversity might be underestimated, which calls for a substantial effort to better characterize the diversity of this particular group and has important implications in terms of conservation and management. This result further indicates East Asia provided important evolutionary reservoirs during the diversification of Culterinae. As such, this study warrants further research developments and provides guidelines for future taxonomic and genomic studies.

### AUTHORS CONTRIBUTIONS

W.C., N.H., X.L. and J.L. designed the study; W.C., D.X. and C.Z. performed molecular experiments; W.C., Y.L., X.C., S.Z., J.Y., X.W. and C.Z. collected specimens and coordinated morphological identifications. W.C. and N.H. performed data analyses and wrote an initial draft of the manuscript. All authors contributed to the writing of the final draft.

### ACKNOWLEDGEMENTS

The experiments were performed in accordance with the Ethics Committee of the Pearl River Fisheries Research Institute at the Chinese Academy of Fishery Science (CAFS). This research was supported by a project of the innovation team of survey and assessment of the Pearl River fishery resources (2020TD-10 and 2020ZJTD-04) and National Key R&D Program of China (2018YFD0900902 and 2018YFD0900903). We are grateful to Carla Lopes and the three anonymous reviewers for their insightful comments on an earlier version of the manuscript.

### CONFLICT OF INTEREST

The authors declare no competing interests.

### OPEN RESEARCH BADGES

This article has obtained an Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.org/10.5883/DS-CUBAR.

### DATA AVAILABILITY STATEMENT

Novel sequence data are available on the Barcode of Life Data system (BOLD) in the projects "General Project–CUBAR Barcodes of Culterinae species" (10.5883/DS-CUBAR; see Table S1).

### ORCID

*Weitao Chen* https://orcid.org/0000-0002-4324-1926
*Nicolas Hubert* https://orcid.org/0000-0001-9248-3377

### REFERENCES

Adeoba, M. I., Kabongo, R., Van der Bank, H., & Yessoufou, K. (2018). Re-evaluation of the discriminatory power of DNA barcoding on some specimens of African Cyprinidae (subfamilies Cyprininae

and Danioninae). *ZooKeys*, *746*, 105–121. https://doi.org/10.3897/zookeys.746.13502

Alcantara, S. G., & Yambot, A. V. (2016). DNA barcoding of commercially important grouper species (Perciformes, Serranidae) in The Philippines. *Mitochondrial DNA Part A*, *27*(6), 3837–3845. https://doi.org/10.3109/19401736.2014.958672

An, Z. S., Kutzbach, J. E., Prell, W. L., & Porter, S. C. (2001). Evolution of Asian monsoons and phased uplift of the Himalayan Tibetan plateau since late Miocene times. *Nature*, *411*(6833), 62–66. https://doi.org/10.1038/35075035

April, J., Mayden, R. L., Hanner, R. H., & Bernatchez, L. (2011). Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(26), 10602–10607. https://doi.org/10.1073/pnas.1016437108

Avise, J. C. (2000). *Phylogeography: The history and formation of species*. Harvard University Press.

Avise, J. C. (2012). *Molecular markers, natural history and evolution*. Springer Science & Business Media.

Bai, X., Guo, X., Zhang, X., Song, W., Li, Y., Luo, W., Cao, X., & Wang, W. (2015). Species identification and evolutionary inference of the genera Megalobrama and Parabramis (Cyprinidae: Cultrinae) in China. *Mitochondrial DNA*, *26*(3), 357–366. https://doi.org/10.3109/19401736.2013.823166

Ball, S. L., Hebert, P. D., Burian, S. K., & Webb, J. M. (2005). Biological identifications of mayflies (Ephemeroptera) using DNA barcodes. *Journal of the North American Benthological Society*, *24*(3), 508–524. https://doi.org/10.1899/04-142.1

Bandelt, H. J., Forster, P., & Rohl, 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

Bellafronte, E., Mariguela, T. C., Pereira, L. H. G., Oliveira, C., & Moreira-Filho, O. (2013). DNA barcode of Parodontidae species from the La Plata river basin-applying new data to clarify taxonomic problems. *Neotropical Ichthyology*, *11*(3), 497–506. https://doi.org/10.1590/S1679-62252013000300003

Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G. N., Ribera, I., Nilsson, A. N., Barraclough, T. G., & Vogler, A. P. (2012). The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology*, *61*(5), 851–869. https://doi.org/10.1093/sysbio/sys037

Bermingham, E., McCafferty, S., Martin, A. P., & Kocher, T. D. (1997). Fish biogeography and molecular clocks: Perspectives from the Panamanian isthmus. In *Molecular systematics of fishes* (pp. 113–128). Academic Press.

Breman, F. C., Loix, S., Jordaens, K., Snoeks, J., & Van Steenberge, M. (2016). Testing the potential of DNA barcoding in vertebrate radiations: The case of the littoral cichlids (Pisces, Perciformes, Cichlidae) from Lake Tanganyika. *Molecular Ecology Resources*, *16*(6), 1455–1464. https://doi.org/10.1111/1755-0998.12523

Cao, X., & Wang, W. (2010). Haematological and biochemical characteristics of two aquacultured carnivorous cyprinids, topmouth culter *Culter alburnus* (Basilewsky) and yellowcheek carp *Elopichthys bambusa* (Richardson). *Aquaculture Research*, *41*(9), 1331–1338. https://doi.org/10.1111/j.1365-2109.2009.02421.x

Chen, Y., Cao, W., & Zheng, C. (1986). Ichthyofauna of the Zhujiang River with a discussion on zoogeographical divisions for freshwater fishes. *Acta Hydrobiologica Sinica*, *10*, 228–236.

Chen, X. L., Chiang, T. Y., Lin, H. D., Zheng, H. S., Shao, K. T., Zhang, Q., & Hsu, K. C. (2007). Mitochondrial DNA phylogeography of *Glyptothorax fokiensis* and *Glyptothorax hainanensis* in Asia. *Journal of Fish Biology*, *70*, 75–93. https://doi.org/10.1111/j.1095-8649.2007.01370.x

Chen, W., Li, C., Chen, F., Li, Y., Yang, J., Li, J., & Li, X. (2020). Phylogeographic analyses of a migratory freshwater fish (*Megalobrama terminalis*) reveal a shallow genetic structure and pronounced effects of sea-level changes. *Gene*, *737*, 1–10. https://doi.org/10.1016/j.gene.2020.144478

Chen, W., Ma, X., Shen, Y., Mao, Y., & He, S. (2015). The fish diversity in the upper reaches of the Salween River, Nujiang River, revealed by DNA barcoding. *Scientific Reports*, *5*, 17437. https://doi.org/10.1038/srep17437

Chen, W., Zhong, Z., Dai, W., Fan, Q., & He, S. (2017). Phylogeographic structure, cryptic speciation and demographic history of the sharpbelly (*Hemiculter leucisculus*), a freshwater habitat generalist from southern China. *BMC Evolutionary Biology*, *17*(1), 216. https://doi.org/10.1186/s12862-017-1058-0

Chiang, T. Y., Lin, H. D., Zhao, J., Kuo, P. H., Lee, T. W., & Hsu, K. C. (2013). Diverse processes shape deep phylogeographical divergence in *Cobitis sinensis* (Teleostei: Cobitidae) in East Asia. *Journal of Zoological Systematics and Evolutionary Research*, *51*(4), 316–326. https://doi.org/10.1111/jzs.12030

Dahruddin, H., Hutama, A., Busson, F., Sauri, S., Hanner, R., Keith, P., Hadiaty, R., & Hubert, N. (2017). Revisiting the ichthyodiversity of Java and Bali through DNA barcodes: Taxonomic coverage, identification accuracy, cryptic diversity and identification of exotic species. *Molecular Ecology Resources*, *17*(2), 288–299. https://doi.org/10.1111/1755-0998.12528

Dai, Y., Yang, J., & Chen, Y. (2005). Phylogeny and zoogeography of the subfamily Cultrinae (Cyprinidae). *Acta Zootaxonomica Sinica*, *30*(2), 213–233.

Delrieu-Trottin, E., Durand, J., Limmon, G., Sukmono, T., Sugeha, H. Y., Chen, W., Busson, F., Borsa, P., Dahruddin, H., & Sauri, S. (2020). Biodiversity inventory of the grey mullets (Actinopterygii: Mugilidae) of the Indo-Australian archipelago through the iterative use of DNA-based species delimitation and specimen assignment methods. *Evolutionary Applications*, *13*(6), 1451–1467. https://doi.org/10.1111/eva.12926

Delrieu-Trottin, E., Williams, J. T., Pitassy, D., Driskell, A., Hubert, N., Viviani, J., Cribb, T. H., Espiau, B., Galzin, R., Kulbicki, M., Lison de Loma, T., Meyer, C., Mourier, J., Mou-Tham, G., Parravicini, V., Plantard, P., Sasal, P., Siu, G., Tolou, N., & Planes, S. (2019). A DNA barcode reference library of French Polynesian shore fishes. *Scientific Data*, *6*(1), 1–8. https://doi.org/10.1038/s41597-019-0123-5

Ding, R. (1994). *The fishes of Sichuan*. Sichuan Science and Technology Press.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*, 214. https://doi.org/10.1186/1471-2148-7-214

Durand, J.-D., Hubert, N., Shen, K.-N., & Borsa, P. (2017). DNA barcoding grey mullets. *Reviews in Fish Biology and Fisheries*, *27*(1), 233–243. https://doi.org/10.1007/s11160-016-9457-7

Feng, X., Xie, N., Feng, J., Xu, B., & Zhao, J. (2009). Phylogenetic relationships of the Culterinae fishes based on cytochrome *b* sequences analysis. *Freshwater Fish*, *39*, 23–27.

Froese, R., & Pauly, D. (2020). *Fishbase*. Worldwide web electronic publication, www.fishbase.org, version (01/2020). http://www.fishbase.org

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, *62*(5), 707–724. https://doi.org/10.1093/sysbio/syt033

Galimberti, A., Assandri, G., Maggioni, D., Ramazzotti, F., Baroni, D., Bazzi, G., … Casiraghi, M. (2021). Italian odonates in the Pandora's box: A comprehensive DNA barcoding inventory shows taxonomic warnings at the Holarctic scale. *Molecular Ecology Resources*, *21*(1), 183–200. https://doi.org/10.1111/1755-0998.13235

Galimberti, A., Spada, M., Russo, D., Mucedda, M., Agnelli, P., Crottini, A., … Casiraghi, M. (2012). Integrated operational taxonomic units (IOTUs) in echolocating bats: A bridge between molecular

and traditional taxonomy. *PLoS One*, *7*(6), e40122. https://doi.org/10.1371/journal.pone.0040122

Glez-Pena, D., Gomez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., & Posada, D. (2010). ALTER: Program-oriented conversion of DNA and protein alignments. *Nucleic Acids Research*, *38*, 14–18. https://doi.org/10.1093/nar/gkq321

Guo, X. G., He, S. P., & Zhang, Y. G. (2006). Phylogeny and biogeography of Chinese sisorid catfishes re-examined using mitochondrial cytochrome *b* and 16S rRNA gene sequences. *Molecular Phylogenetics and Evolution*, *38*(1), 291. https://doi.org/10.1016/j.ympev.2005.09.013

He, S. P., Cao, W. X., & Chen, Y. Y. (2001). The uplift of Qinghai–Xizang (Tibet) Plateau and the vicariance speciation of glyptosternoid fishes (Siluriformes: Sisoridae). *Science in China Series C: Life Sciences*, *44*, 644–651. https://doi.org/10.1007/BF02879359

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S., & Francis, C. M. (2004). Identification of birds through DNA barcodes. *PLoS Biology*, *2*(10), e312. https://doi.org/10.1371/journal.pbio.0020312

Hou, G., Chen, W. T., Lu, H. S., Cheng, F., & Xie, S. G. (2018). Developing a DNA barcode library for perciform fishes in the South China Sea: Species identification, accuracy and cryptic diversity. *Molecular Ecology Resources*, *18*(1), 137–146. https://doi.org/10.1111/1755-0998.12718

Hubert, N., Delrieu-Trottin, E., Irisson, J.-O., Meyer, C., & Planes, S. (2010). Identifying coral reef fish larvae through DNA barcoding: A test case with the families Acanthuridae and Holocentridae. *Molecular Phylogenetics and Evolution*, *55*(3), 1195–1203. https://doi.org/10.1016/j.ympev.2010.02.023

Hubert, N., & Hanner, R. (2015). DNA barcoding, species delineation and taxonomy: A historical perspective. *DNA Barcodes*, *3*, 44–58. https://doi.org/10.1515/dna-2015-0006

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., 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), e2490. https://doi.org/10.1371/journal.pone.0002490

Hubert, N., Meyer, C. P., Bruggemann, H. J., Guérin, F., Komeno, R. J. L., Espiau, B., Causse, R., Williams, J. T., & Planes, S. (2012). Cryptic diversity in indo-pacific coral-reef fishes revealed by DNA-barcoding provides new support to the Centre-of-overlap hypothesis. *PLoS One*, *7*(3), e28987. https://doi.org/10.1371/journal.pone.0028987

Jaafar, T. N., Taylor, M. I., Nor, S. A. M., De Bruyn, M., & Carvalho, G. R. (2012). DNA barcoding reveals cryptic diversity within commercially exploited Indo-Malay Carangidae (Teleosteii: Perciformes). *PLoS One*, *7*(11), e49623. https://doi.org/10.1371/journal.pone.0049623

Jiang, Z., Jiang, J., Wang, Y., Zhang, E., Zhang, Y., Li, L., … Zheng, G. (2016). Red list of China's vertebrates. *Biodiversity Science*, *24*(5), 500–551. https://doi.org/10.17520/biods.2016076

Jiang, W., Qiu, Y., Pan, X., Zhang, Y., Wang, X., Lv, Y., … Chen, J. (2018). Genome assembly for a Yunnan-Guizhou plateau "3E" fish, *anabarilius grahami* (Regan), and its evolutionary and genetic applications. *Frontiers in Genetics*, *9*, 614. https://doi.org/10.3389/fgene.2018.00614

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, *33*(11), 1630–1638. https://doi.org/10.1093/bioinformatics/btx025

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

Kim, S., Lee, Y., Mutanen, M., Seung, J., & Lee, S. (2020). High functionality of DnA barcodes and revealed cases of cryptic diversity in Korean curved-horn moths (lepidoptera: Gelechioidea). *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-63385-x

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

Lara, A., Ponce de Leon, J. L., Rodriguez, R., Casane, D., Cote, G., Bernatchez, L., & Garcia-Machado, E. (2010). DNA barcoding of Cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources*, *10*(3), 421–430. https://doi.org/10.1111/j.1755-0998.2009.02785.x

Laskar, B. A., Bhattacharjee, M. J., Dhar, B., Mahadani, P., Kundu, S., & Ghosh, S. K. (2013). The species dilemma of northeast Indian mahseer (Actinopterygii: Cyprinidae): DNA barcoding in clarifying the riddle. *PLoS One*, *8*(1), e53704. https://doi.org/10.1371/journal.pone.0053704

Li, S., Cai, W., & Zhou, B. (1993). Variation in morphology and biochemical genetic markers among populations of blunt snout bream (*Megalobrama amblycephala*). In *Genetics in aquaculture* (pp. 117–127). Elsevier.

Li, G. Y., Wang, X. Z., Zhao, Y. H., Zhang, J., Zhang, C. G., & He, S. P. (2009). Speciation and phylogeography of *Opsariichthys bidens* (Pisces: Cypriniformes: Cyprinidae) in China: Analysis of the cytochrome *b* gene of mtDNA from diverse populations. *Zoological Studies*, *48*(4), 569–583.

Liu, Z., Trentesaux, A., Clemens, S. C., Colin, C., Wang, P., Huang, B., & Boulay, S. (2003). Clay mineral assemblages in the northern South China Sea: Implications for east Asian monsoon evolution over the past 2 million years. *Marine Geology*, *201*(1), 133–146. https://doi.org/10.1016/S0025-3227(03)00213-5

Luo, W. (2016). *The phylogenetic relationships and hybridization introgression of Chinese bream fishes in Cultrinae (Teleostei: Cypriniformes: Cyprinidae)* Master Dissertation. Southwest University.

Luo, Y., & Chen, Y. (1998). *Culterinae. Fauna Sinica, Osteichthyes, Cypriniformes II*. Science Press.

McCusker, M. R., Denti, D., Van Guelpen, L., Kenchington, E., & Bentzen, P. (2013). Barcoding Atlantic Canada's commonly encountered marine fishes. *Molecular Ecology Resources*, *13*(2), 177–188. https://doi.org/10.1111/1755-0998.12043

Meyer, C. P., & Paulay, G. (2005). DNA barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, *3*(12), e422. https://doi.org/10.1371/journal.pbio.0030422

Moritz, C. (1994). Defining 'Evolutionarily significant Units' for conservation. *Trends in Ecology and Evolution*, *9*(10), 373–375. https://doi.org/10.1016/0169-5347(94)90057-4

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26*(4), 867–878. https://doi.org/10.1046/j.1365-2699.1999.00305.x

Nylander, J. (2004). *MrModeltest v2. Program distributed by the author.* Evolutionary Biology Centre, Uppsala University.

Oliveira, R., Castro, D., Godinho, R., Luikart, G., & Alves, P. (2010). Species identification using a small nuclear gene fragment: Application to sympatric wild carnivores from South-Western Europe. *Conservation Genetics*, *11*(3), 1023–1032. https://doi.org/10.1007/s10592-009-9947-4

Panprommin, D., Soontornprasit, K., Tuncharoen, S., Pithakpol, S., & Keereelang, J. (2019). DNA barcodes for the identification of species diversity in fish from Kwan Phayao, Thailand. *Journal of Asia-Pacific Biodiversity*, *12*(3), 382–389. https://doi.org/10.1016/j.japb.2019.05.003

Perdices, A., Sayanda, D., & Coelho, M. M. (2005). Mitochondrial diversity of *Opsariichthys bidens* (Teleostei, Cyprinidae) in three Chinese drainages. *Molecular Phylogenetics and Evolution*, *37*(3), 920–927. https://doi.org/10.1016/j.ympev.2005.04.020

Pereira, L. H. G., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse neotropical freshwater fish fauna? *BMC Genetics*, *14*, 20. https://doi.org/10.1186/1471-2156-14-20

Petit, R. J., & Excoffier, L. (2009). Gene flow and species delimitation. *Trends in Ecology & Evolution*, *24*(7), 386–393. https://doi.org/10.1016/j.tree.2009.02.011

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, *21*(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Qiu, C., Lin, Y., Qing, N., Zhao, J., & Chen, X. (2008). Genetic variation and phylogeography of *Micronoemacheilus pulcher* populations among drainage systems between western South China and Hainan Island. *Acta Entomologica Sinica*, *51*(11), 1099–1128.

Rambaut, A., & Drummond, A. (2007). Trace v1.4. Retrieved from http://beast.bio.ed.ac.uk/Tracer

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS One*, *8*(7). https://doi.org/10.1371/journal.pone.0066213

Raupach, M. J., Astrin, J. J., Hannig, K., Peters, M. K., Stoeckle, M. Y., & Wägele, J. W. (2010). Molecular species identification of central European ground beetles (coleoptera: Carabidae) using nuclear rDNA expansion segments and DNA barcodes. *Frontiers in Zoology*, *7*(1), 1–15. https://doi.org/10.1186/1742-9994-7-26

Ren, M., Bao, H., & Han, T. (1959). The Jinsha River valley landforms and river-capture in northwestern Yunnan. *Acta Geographica Sinica*, *25*(2), 135–155.

Ribeiro, A. D. O., Caires, R. A., Mariguela, T. C., Pereira, L. H. G., Hanner, R., & Oliveira, C. (2012). DNA barcodes identify marine fishes of Sao Paulo state, Brazil. *Molecular Ecology Resources*, *12*(6), 1012–1020. https://doi.org/10.1111/1755-0998.12007

Satler, J. D., Carstens, B. C., & Hedin, M. (2013). Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, Aliatypus). *Systematic Biology*, *62*(6), 805–823. https://doi.org/10.1093/sysbio/syt041

Shen, Y., Guan, L., Wang, D., & Gan, X. (2016). DNA barcoding and evaluation of genetic diversity in Cyprinidae fish in the midstream of the Yangtze River. *Ecology and Evolution*, *6*(9), 2702–2713. https://doi.org/10.1002/ece3.2060

Shen, Y., Hubert, N., Huang, Y., Wang, X., Gan, X., Peng, Z., & He, S. (2019). DNA barcoding the ichthyofauna of the Yangtze River: Insights from the molecular inventory of a mega-diverse temperate fauna. *Molecular Ecology Resources*, *19*(5), 1278–1291. https://doi.org/10.1111/1755-0998.12961

Sholihah, A., Delrieu-Trottin, E., Sukmono, T., Dahruddin, H., Risdawati, R., Elvyra, R., Wibowo, A., Kustiati, K., Busson, F., Sauri, S., Nurhaman, U., Dounias, E., Zein, M. S. A., Fitriana, Y., Utama, I. V., & Hubert, N. (2020). Disentangling the taxonomy of the subfamily Rasborinae (Cypriniformes, Danionidae) in Sundaland using DNA barcodes. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-59544-9

Smith, M. A., Fisher, B. L., & Hebert, P. D. (2005). DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1825–1834. https://doi.org/10.1098/rstb.2005.1714

Sonet, G., Snoeks, J., Nagy, Z. T., Vreven, E., Boden, G., Breman, F. C., Decru, E., Hanssens, M., Ibala Zamba, A., & Jordaens, K. (2019). DNA barcoding fishes from The Congo and the lower Guinean provinces: Assembling a reference library for poorly inventoried fauna. *Molecular Ecology Resources*, *19*(3), 728–743. https://doi.org/10.1111/1755-0998.12983

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, *22*(21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. https://doi.org/10.1093/molbev/mst197

Tzeng, C. H., & Chiu, T. S. (2012). DNA barcode-based identification of commercially caught cutlassfishes (family: Trichiuridae) with a phylogenetic assessment. *Fisheries Research*, *127*, 176–181. https://doi.org/10.1016/j.fishres.2012.01.022

Wang, D. (2005). *Species diversity and zoogeographical analysis of freshwater fishes in Guangxi Zhuang autonomous region*. Master's Thesis. Institute of Zoology, Chinese Academy of Sciences.

Wang, L., Luo, X., Wang, D., Li, W., Li, Z., Zou, G., & Liang, H. (2019). Identification analysis of DNA barcoding based on mitochondrial COI and Cytb genes in *Culter* fishes. *Freshwater Fisheries*, *49*(4), 22–28.

Wang, L., Sarnthein, M., Erlenkeuser, H., Grimalt, J., Grootes, P., Heilig, S., … Pflaumann, U. (1999). East Asian monsoon climate during the late Pleistocene: High-resolution sediment records from the South China Sea. *Marine Geology*, *156*(1–4), 245–284. https://doi.org/10.1016/S0025-3227(98)00182-0

Wang, W., Zhao, J., & Li, S. (2006). Genetic variation of the mitochondrial DNA Cytb among six populations of *Siniperca scherzeri* in China. *Zoological Research*, *27*(6), 589–593.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1847–1857. https://doi.org/10.1098/rstb.2005.1716

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.

Wu, Q., & Yi, B. (1959). Preliminary ecological survey of *Hemiculter* and *Hemiculter* species in the Heilongjiang Basin. *Acta Hydrobiologica Sinica*, *2*, 157–168.

Xie, N., Liu, X., Feng, X., & Guo, S. (2012). Sequences analysis on mitochondrial cytochrome *b* gene fragment of *Megalobram* spp. *Modern Agriculture Science Technology*, *1*, 290–292.

Yang, J. Q., Hsu, K. C., Liu, Z. Z., Su, L. W., Kuo, P. H., Tang, W. Q., … Lin, H. D. (2016). The population history of *Garra orientalis* (Teleostei: Cyprinidae) using mitochondrial DNA and microsatellite data with approximate Bayesian computation. *BMC Evolutionary Biology*, *16*, 73. https://doi.org/10.1186/s12862-016-0645-9

Yang, L., Mayden, R. L., & He, S. (2009). Population genetic structure and geographical differentiation of the Chinese catfish *Hemibagrus macropterus* (Siluriformes, Bagridae): Evidence for altered drainage patterns. *Molecular Phylogenetics and Evolution*, *51*(2), 405–411. https://doi.org/10.1016/j.ympev.2009.01.004

Yang, J., Yang, J. X., & Chen, X. Y. (2012). A re-examination of the molecular phylogeny and biogeography of the genus *Schizothorax* (Teleostei: Cyprinidae) through enhanced sampling, with emphasis on the species in the Yunnan-Guizhou plateau, China. *Journal of Zoological Systematics and Evolutionary Research*, *50*(3), 184–191. https://doi.org/10.1111/j.1439-0469.2012.00661.x

Yu, D., Chen, M., Tang, Q. Y., Li, X. J., & Liu, H. Z. (2014). Geological events and Pliocene climate fluctuations explain the phylogeographical pattern of the cold water fish *Rhynchocypris oxycephalus* (Cypriniformes: Cyprinidae) in China. *BMC Evolutionary Biology*, *14*(1), 1–12. https://doi.org/10.1186/s12862-014-0225-9

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

Zhang, D., Nie, Z., Yang, Z., Luo, W., Yi, S., Gao, Z., … Chen, B. (2014). The interspecific hybridization in four freshwater bream *Megalobrama* sp. *Journal of Dalian Ocean University*, *29*(2), 121–125.

Zhang, Y. H., Qin, G., Zhang, H. X., Wang, X., & Lin, Q. (2017). DNA barcoding reflects the diversity and variety of brooding traits of fish species in the family Syngnathidae along China's coast. *Fisheries Research*, *185*, 137–144. https://doi.org/10.1016/j.fishres.2016.09.015

Zheng, C. (1989). *Ichthyography of the Pearl River*. Science Press.

Zhong, J., Yi, S., Ma, L., & Wang, W. (2019). Evolution and phylogeography analysis of diploid and polyploid *Misgurnus anguillicaudatus* populations across China. *Proceedings of the Royal Society B-Biological Sciences*, *286*(1901), 20190076. https://doi.org/10.1098/rspb.2019.0076

Zhou, J., & Zhang, C. (2005). *Freshwater fishes of Guangxi*, *China*. Guangxi People's Publishing House.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Chen, W., Hubert, N., Li, Y., Xiang, D., Cai, X., Zhu, S., Yang, J., Zhou, C., Li, X., & Li, J. (2022). Large-scale DNA barcoding of the subfamily Culterinae (Cypriniformes: Xenocyprididae) in East Asia unveils a geographical scale effect, taxonomic warnings and cryptic diversity. *Molecular Ecology*, *00*, 1–17. https://doi.org/10.1111/mec.16540