DNA Barcoding Reveals Cryptic Diversity in the Underestimated Genus *Triplophysa* (Cypriniformes: Cobitidae, Nemacheilinae) from the Northeastern Qinghai-Tibet Plateau

tai wang

Gansu Agricultural University

Yanping Zhang

Gansu fisheries research institute

Zhuoyu Yang

Gansu fisheries research institute

Zhe Liu

Gansu Agricultural University

Yanyan Du

Gansu fisheries research institute

Research article

Keywords: DNA barcode, Qinghai-Tibet, Tibetan loach, cryptic species

---

Abstract

Background: The northeastern part of the Qinghai-Tibet Plateau (QTP) is one of the areas where the number of species of plateau loach is the largest. As one of the three major groups of fishes distributed on the QTP, plateau loach have very important ecological value. However, their taxonomy and systematics are still controversial, and a large number of new species have been reported. The reason for this phenomenon is that the degree of morphological variation is low, the phylogenetic information provided by morphological and anatomical features used for species identification is relatively poor, and there are many cryptic species. Based on the high-density sampling points from the biodiversity hotspots surveyed, this study aims to evaluate the biodiversity of plateau loach in the northeastern of QTP and reveal the hidden diversity by comparing morphologic species with molecular operational taxonomic units (MOTUs).

Results: After careful identification and comparison of the morphology and DNA barcoding of 1,630 specimens, 22 species were identified, 20 of which were considered valid local species and two of which were new species that had not been described. Based on the combination of morphological and molecular methods, a total of 24 native species have been found, two of which are cryptic species: *Triplophysa robusta sp1* and *Triplophysa minxianensis sp1*. Fourteen of the 24 species form clusters of barcodes, which allow them to be reliably identified. The remaining cases involved 10 closely related species, some of which were rapidly differentiated, seemed to have experienced incomplete lineage sorting, or showed introgressions.

Conclusions: The results highlight the need to combine traditional taxonomies with molecular methods to correctly identify species, especially in closely related species such as the plateau loach. This study provides a basis for protecting the biodiversity of plateau loach.

Background

With problems such as global climate change, issues related to populations, the ecological environment, energy and food are becoming increasingly serious, and sustainable anthropogenic development and the ability to understand and meet the requirements of biodiversity is becoming urgent (Loreau et al, 2001; Isbell et al, 2011; Cardinale et al, 2012). There is a major global demand for accurate and rapid identification of species for the protection and sustainable use of biodiversity resources. Species identification and classification is a basic requirement for biological research. Based on morphological characteristics, classical taxonomy has made great contributions to species classification; however, due to morphological plasticity, traditional taxonomy cannot accurately distinguish all species, in particular, some forms of similar, related species (Robinson and Parsons, 2002; Pigliucci, 2005). Therefore, there is a need for a new way to support species identification with classical taxonomy methods. Tautz et al. (2002) first suggested using DNA sequencing, namely, DNA taxonomy, as the main platform for biological classification. Then, Professor Paul Hebert from the University of Guelph in Canada introduced the concept of DNA barcoding, highlighting its significance to the field of biological taxonomy and species

identification (Hebert et al., 2003; Remigio et al., 2003) and suggesting the use of the mitochondrial cytochrome C oxidase subunit I (COI) gene as the basis for animal DNA barcoding. The applicability of DNA barcoding to the identification of marine and freshwater fish species has been shown by using a short fragment of approximately 650 bp from the mitochondrial COI gene to identify species based on sequence differences (Ward et al., 2005; Zhang et al., 2012; Zhang et al., 2013; Valdez-Moreno et al., 2012; Lakra et al., 2011; Bhattacharjee et al., 2012; Hubert et al., 2008). A growing number of studies show that DNA barcodes are widely used in animal species identification, classification, cryptic species detection, phylogenetic research, etc. (Smith et al., 2008; Swartz et al., 2008; Rock et al., 2008; Almerón-Souza et al., 2018), and to construct barcode databases, such as the Barcode of Life Data Systems (BOLD) (http://www.boldsystems.org), in which approximately 96,425 fish specimens belonging to 10,267 species have been barcoded. DNA barcoding, as a compliment to traditional species identification, can be used to automate and standardize the process of specimen identification, reducing the dependence on the experience of taxonomists (Sales et al., 2018; Burrows et al., 2019).

The Qinghai-Tibet Plateau (QTP), known as "the roof of the world", is rich in biodiversity and is a relatively unique area with many endemic species (Khan, et al., 2005). The native fish living in the Qinghai-Tibet region belong to three orders: Salmoniformes, Siluriformes and Cypriniformes (Wu and Wu, 1992). *Triplophysa*, which belong to the family Nemacheilinae (Cypriniformes) are widely distributed on the QTP and in its adjacent regions (Wang et al., 2016). It is a special group adapted to the climatic characteristics of the QTP, such as cool temperatures and oxygen shortages (Zhu and Wu, 1981; Wu and Wu, 1992). In 1992, there were 33 *Triplophysa* species identified. However, over time, a large number of new species have been described; so far there are a total of 140 valid species (He et al., 2008; Li et al.,2017;). Although there may be some synonym species (He et al., 2008; Prokofiev et al., 2007), these studies show that a large amount of unknown biodiversity exists in the *Triplophysa*, and many species have not been recognized or described. The phenomenon of many new species being reported is mainly caused by the existence of cryptic species or the lack of careful classification review. The simple body structure and relatively conservative morphological evolution of the plateau loach fish, coupled with their weak migration ability due to the restrictions of the water system, have led to limited gene exchange between different populations. Over time, although morphologically imperceptible, the process of species differentiation, including genetic structural differentiation and reproductive isolation, may have occurred, and many hidden taxa may have been ignored. Therefore, the genus *Triplophysa* should be considered in the study of cryptic diversity.

Classical morphological classification has always played a dominant role in species identification, but it has limitations. In particular, for the fish of the genus *Triplophysa*, the phenotype is easily affected by biological factors and the external environment and there is morphological plasticity; therefore, morphological differences are not easily detected (He et al., 2008). Moreover, some species were named many years ago, and their morphological descriptions were relatively simple. All these factors have led to difficulties in the subsequent identification of species and taxonomic research. Due to the difficulty in obtaining detailed data for comparisons, it is possible that the distribution of some species is artificially expanded and mistakenly divided into different geographical populations (Ding et al., 1996; Prokofiev,

2007). This long confusing classification history is more obvious in the species classification of the *Triplophysa* fish. There is an urgent need to accelerate the discovery and recording of fish biodiversity, which is important for achieving effective species conservation and predicting the biodiversity response to climate change (McCloud, 2010; Peterson et al., 2002; Barley and Thomson, 2016). Although DNA barcode technology has certain limitation, such as software model or parameter improper leading to excessive division of species or gathered, DNA barcode itself introgression, pseudogenes, definition of molecule as a result (Shaw, 2002; Collins and Cruickshank, 2013; Carstens et al., 2013; Pecnikar and Buzan, 2014), the DNA barcode has been proposed as an important tool, to make up for the large gap in our current understanding of biological diversity (Lim and Meier, 2011; Barrett et al., 2005; Dinca et al., 2011; Dhar et al.⁠2017). Some relatively automated methods of species division using DNA barcoding data sets have been developed, which offer the hope of making species division more efficient and less subjective (Knowles and Carstens,2007; Ence and Carstens, 2011; Esselstyn et al., 2012).

To date, there are no studies on the identification or evaluation of cryptic biodiversity within the genus *Triplophysa* using DNA barcoding in the northeastern QTP. Herein, based on extensive sample collection in this area, DNA barcoding technology was used to evaluate the biodiversity of plateau loach in the northeastern of QTP, which is a hotspot of biodiversity. A DNA barcoding database of plateau loach in this area was obtained. The relationships between morphological species and molecular operational taxonomic units (MOTUs) were assessed and this hidden biodiversity in the genus was identified. This research will contribute to a more comprehensive understanding of its biodiversity and to the protection of this important fish population.

Results

A total of 1,630 native specimens were collected from the northeastern edge of the QTP (Table S1; Figure1; Figure2), and 22 morphospecies were identified including two undetermined species (*Triplophysa sp1* and *Triplophysa sp2*). Among the specimens, the endemic species *T. robusta* (n=413) had the largest number of individuals, followed by *T. minxianensis* (n=253). The undetermined species *T. sp1* (n=3) and *T. bleekeri* (n=5) had the lowest number of specimens, with 68 specimens per species on average (Table 1). A total of 1,630 COI sequences were obtained. The size of the sequences obtained was 606 bp after trimming to a consensus length. No stop codons were observed, and the mean nucleotide composition within the complete data set was 30.6% thymine (T), 26.7% cytosine (C), 24.3% adenine (A) and 18.4% guanine (G). There were 393 conserved sites, 213 variable sites, 178 parsimonious sites and 35 singleton sites. A total of 230 unique haplotypes were generated in the 1623 COI sequences. The haplotype number of *T. robusta* was the largest (Nh=46), followed by that of *T. obscura* (Nh=27) and *T. stoliczkai* (Nh=25). The haplotype numbers of *T. bleeker* and *T. orientalis* were the smallest (Nh=1). Correspondingly, the haplotype diversity of *T. robusta* was the highest (h=0.9360±0.006). The nucleotide diversity was the highest for *T. obscura* (π=0.00777±0.00145) (Table 1).

The phylogenetic tree was constructed by neighbor-joining method (NJ), maximum likelihood method (ML) and Bayesian inference method (BI).With *Homatula variegata* as the outgroup group (Gen Bank no. :

MF953219), the topological structure of the phylogenetic trees obtained by the three analysis methods was basically the same, and only the topological structure of NJ tree was retained here, and the values at the nodes respectively represent the support degree at the nodes of NJ/ML/BI tree. The Poisson tree processes model (PTP) analysis with a maximum likelihood partition and Bayesian implementation resulted in 17 molecular operational taxonomic units (MOTUs) (Figure 3). The general mixed Yule-coalescent model (GMYC) analysis resulted in the same 17 MOTUs as those obtained in the PTP analysis (likelihood ratio=76.41, *P*<0.0001), whereas the Automatic Barcode Gap Discovery (ABGD) and the Barcode of Life Data Systems (BOLD) detected 19 MOTUs for the 22 morphological species (Figure 3). *T. strauchii, T. orientalis, T. tenuis, T. wuweiensis, T. polyfasciata, T. bleekeri, T. sp1, T. sellaefer, T. minxianensis sp1, T. hsutschouensis* and *T. robusta* showed correspondence between the morphological species and MOTUs. The MOTUs of *T. minxianensis, T. pappenheimi, T. siluroides, T. pappenheimi* and *T. robusta sp1* cannot be distinguished by the PTP, GMYC, ABGD or BOLD analyses. The same phenomenon occurs between *T. stoliczkae* and *T. dalaica* and between *T. scleroptera* and *T. pseudoscleroptera. T. leptosome* and *T. papilloso-labiatus* cannot be distinguished by the PTP or GMYC analyses, but they can be distinguished by the ABGD and BOLD analyses. The same was true of *T. shiyangensis*.

The average Kimura 2-parameter (K2P) intraspecific distance ranged between 0 and 3.10% (Table 2). The maximum observed average K2P intraspecific distance was that of *T. robusta*. The maximum intraspecific K2P distance ranged from 0 to 7.90%. The largest K2P intraspecific distance was observed for *T. robusta*, followed by *T. minxianensis* with a value of 7.40%. The nearest neighbour distance ranged between 0 and 8.57%. For *T. robusta, T. minxanensis, T. siluroides* and *T. pappenheimi*, a nearest neighbour distance of 0% was observed. The nearest neighbour distance of 18 species was lower than the maximum K2P intraspecific distance. Only the nearest neighbour distance of *T. scleroptera* and *T. pseudoscleroptera* was less than 1%, at 0.40%. The distributions of the maximum K2P intraspecific distances and the nearest neighbour K2P genetic distances reflected the overlap; in addition, no barcode gap was found (Figure 4).

Most species form very good evolutionary clades in the neighbor-joining method (NJ) tree, and these main clades represent different taxonomic species. Monophyletic clades have also been observed for *T. stoliczkae* and *T. dalaica*, and *T. scleroptera* and *T. pseudoscleroptera*. Neither *T. minxianensis* nor *T. robusta* formed an independent monophyletic clade, but they formed two larger clades according to geographic distribution. Because of a shared haplotype between *T. minxianensis, T. pappenheimi, T. siluroides* and *T. robusta*, these four species form a larger clade. The trend of mixed genealogies was confirmed by the examination of the haplotype networks. Two species pairs (*T. stoliczkae* and *T. dalaica* (Figure 5 clade A-1) and *T. scleroptera* and *T. pseudoscleroptera* (Figure 5 clade A-4)) cannot be distinguished by the four algorithms used for MOTU delimitation, and there was no shared haplotype between them. Four haplotypes were shared among *T. minxianensis, T. pappenheimi, T. siluroides* and *T. robusta* (Figure 5 cladeB).

Discussion

In this study, a total of 24 species were reported, including two new species: a cryptic species in the *T. minxianensis* population and a cryptic species in the *T. robusta* population. The morphological and molecular data were consistent in 14 of the 22 species identified. The results show that there are two cryptic species that can be described in the biodiversity hotspot area, which reinforces the general view that there is still a large amount of unrecorded diversity in the plateau loach. There is only one haplotype in the clade of *T. bleekeri* and *T. orientalis*. It is necessary to collect more specimens and add sequences, but we do not rule out the possibility of identifying more cryptic species.

Different numbers of MOTUs were identified in the four DNA barcode analysis methods: 17 different MOTUs were identified using the PTP and GMYC models and 19 MOTUs were identified using the ABGD and BOLD methods. *T. shiyangensis* and *T. leptosoma* cannot be distinguished by the PTP or GMYC models, but the ABGD and BOLD methods allow different MOTUs to be assigned to each species (Figure 3). The inconsistent results of the four methods may be due to differences in the methods used to distinguish species, The ABGD and BOLD methods are based on the genetic distance between species and distinguish species by the difference of intraspecific and interspecific genetic distances. The BOLD method defaults to a genetic distance threshold of 2.2% and ABGD of 2.8%, resulting in the same number of MOUTs defined by the two methods. Although it has been pointed out that the RESL in the BOLD system has a stronger taxonomic performance than that in the ABGD system, showing better species identification and MOTU assignment results (Ratnasingham and Hebert, 2013), the two methods in this study achieved the same results, which may be related to the identified species. A key aspect implicit in DNA barcoding analysis is the genetic distance threshold values used to define the MOTUs. COI genetic distance values from 1% (Hubert et al., 2008) to 2% (Keskin et al., 2013) have been considered the threshold values for fish DNA barcoding analysis. However, these values are derived from comparative analyses of species diversity in different aquatic ecosystems. For example, 2% is used to represent the DNA barcodes for the community of fish in certain rivers (Pereira et al., 2013). However, when DNA barcoding analysis was used for a group of closely related species (e.g., the same genus), a lower genetic distance value has been reported (Carvalho et al., 2011; Pereira et al., 2011, 2013). In particular, a low threshold value of 0.92% is needed to distinguish MOTUs in the genus *Laemolyta* (Anostomidae) (Ramirez and Galetti, 2015). Although most of the values obtained in this paper are above 1.47% (14 out of 18 MOTUs, Table 2), the maximum threshold value of related species detected between the obtained MOTUs is 0.40%, and some species have shared haplotypes. The existence of haplotype sharing among different species of plateau fishes may be related to complex species differentiation mechanism or convergent evolution of local adaptation (Shen et al., 2019; Chen et al., 2020). A lower threshold of genetic distance may be obtained when the genetic relationships between different species within a genus are analysed. This approach based on genetic distance analysis is easy to run, but it lacks phylogenetic content, the boundaries used to distinguish species are artificial and lacks the objectivity of species evolution (Ortiz and Francke, 2016). GMYC and PTP are species defining methods based on evolutionary trees, GMYC is based on the ultrametric trees to define species (Fujisawa and Barraclough, 2013), and PTP is based on substitution calibrated trees to define species, which avoids the potential pitfalls of constructing time-calibrated species phylogeny. (Zhang et al., 2013). We believe that tree-based

techniques are effective in identifying individual species because identifying a particular branch representing a particular species requires a threshold to represent the clade length and/or the pair distance used to distinguish differences between individuals.(Gustafsson, et al., 2009; Vieites et al., 2009; Lim and Meier, 2011).Such thresholds are also required when DNA barcoding data is analysed using clustering methods and based on distance methods. A technical problem with clustering is that pair distances of three or more sequences need not be equal, so strict thresholds are usually impossible to apply (Meier et al. 2006). Both methods defined 17 MOUTs in this study. Clearly, the accuracy of DNA barcoding methods depends largely on the target species being analysed (Pentinsaari et al., 2017)

The difference in the number of MOTUs detected by the different analysis methods was mainly seen in two pairs of MOTUs: the genetic distance between *T. shiyangensis* and *T. stoliczkae* was relatively low (2.65%), as was the genetic distance between *T. leptosoma* and *T. papilloso-labiatus* (1.47%). These relatively low genetic distance values may be related to the late differentiation of these MOTUs. Notably, the MOTUs of relatively recent origin had less time than species of distant origin to accumulate genetic differences, which hindered their correct identification, even though the species differ greatly in their morphological characteristics. *T. papilloso-labiatus* has obvious swim bladder, while *T. leptosoma* does not (Zhao, 1984). The characteristics of the genetic diversity of these species are the same: there is a relatively high level of haplotype diversity (> 0.5) and relatively low levels of nucleotide diversity (< 0.5%) (Table 1). This indicates that after the differentiation of these species, influenced by the founder effect and environmental heterogeneity caused by water system changes, the population rapidly accumulated variation, resulting in a high haplotype diversity index. The accumulation time of the nucleotide diversity index was much longer than that of the haplotype diversity index. In terms of geographical distribution, these two species are mainly distributed in the Shulehe River and Heihe River. The possibility of sympatric speciation exists, but this needs to be confirmed by further analysis.

An example of incompletely separated species was also found. *T. minxianensis*, *T. robusta*, *T. pappenheimi* and *T. siluroides* are not sufficiently differentiated by COI gene differences, and there are also shared haplotypes among the four species (Figure 5). These phenomena can be explained as frequent Mitochondrial DNA introgression events before species differentiation (Feng et al., 2018) or phenotypic plasticity in fish (Robinson and Parsons, 2002; Thibert-Plante and Hendry, 2011). The morphological characteristics of *T. hsutschouensis,* which was identified as an independent species isolated from *T. robusta*, include bare and scaleless bodies and a relatively low ratio of body length to body height (Wang, 1991). *T. robusta* only has residual scales in specific parts of its body. The Jinghe River population of *T. robusta* has scales along the lateral line from the caudal fin to the front of the dorsal fin. Moreover, the Jinghe River population and other populations of *T. robusta* were clustered into two clades (Figure 3), and the genetic distance between the populations reached 7.9% (Table 2). These phenomena suggested the existence of cryptic species of *T. robusta*. There was no difference between *T. minxianensis* and *T. minxianensis sp1* in the degradation of the swim bladder, whether the end of the pelvic fin reached the anus, the starting point of the dorsal fin and the pelvic fin relative to each other or the morphological measurement data. But the scales of *T. minxianensis sp1* were only found in the caudal peduncle and this is quite different from *T. minxianensis*, in which all the body parts except the

head have obvious round scales. The genetic distance between the two populations was 7.4% (Table 2), which indicated that there were cryptic species in *T. minxianensis*. Similar to this example of incomplete species separation, Wang (1991) argued that the plateau loach groups without scales (*T. hsutschouensis*) come from scaly groups (*T. minxianensis*) following the degeneration of scales. The groups with remnant body scales (*T. robusta*) are the intermediate species between the two types. The presence or absence of scales marks a leap in the evolution of plateau loach populations. The cryptic species found in this study provide more evidence for this speculation.

The morphological characteristics and molecular characteristics were inconsistent in *T. pseudoscleroptera* and *T. scleroptera*. The two species have similar appearances but different internal anatomical structure. The anterior and posterior segments of the swim bladder of *T. pseudoscleroptera* were the same size, with a long pouch or oblong oval shape and no pyloric caecum. The posterior chamber of the swim bladder of *T. scleroptera* is developed, the anterior segment is thin and the posterior segment is enlarged into a long pouch (Zhu et al., 1981). Without the comparison of internal anatomical structure, these species are easy to misidentify and morphological identification may be incorrect (He et al., 2008). However, due to the low interspecific distance between the two species (0.40%), the two MOTUs cannot be correctly distinguished. This inconsistency was also found between *T. dalaica* and *T. stoliczkai*. The posterior chamber of *T. dalaica*'s swim bladder was oval, while the posterior chamber of *T. stoliczkai* 's swim bladder was degraded; this feature can be used to accurately distinguish the two species.

As shown by the two cases reported here, the DNA barcoding did not show enough difference to distinguish similar species because the lineages were not completely divided into different clades. The reason for this phenomenon is the process of incomplete lineage sorting. Due to the extremely short time of species differentiation, ancestral traits are randomly fixed in the differentiated species (Fontenot et al., 2011; Leavitt et al., 2017). Similar phenomena have been found in *Psorophora* (Chan-Chable et al., 2016), Syngnathidae (Zhang et al., 2017) and *Laemolyta* (Ramirez et al., 2015), and mixed lineage cases are particularly common in plateau fish (Shen et al., 2018). In this sense, to find evidence of reproductive isolation, it is important to combine nuclear genetic and ecological data for further research (Mardulyn et al., 2011; Versteirt et al., 2015; Beebe, 2018).

It is easy to identify species with morphological characteristics that are not significantly different as a single species. For example, *T. bleekeri* and *T. polyfasciata* have very similar morphological characteristics, there is no significant difference in the quantitative traits in different proportions of their bodies, and they have been identified as the same nominal species. Ding et al. (1996) believed that they should be divided into two different species based on molecular data and pointed out that the main distinguishing feature was that there were 10-12 wide, dark brown horizontal stripes on the side of the body. However, even among *T. bleekeri* individuals collected from the same site, the horizontal stripes on the side of its body can range from 0-10. Of the specimens collected from Wenchuanhe River in Sichuan Province, most had 5-7 horizontal stripes, and almost none had more than 10. It was concluded that the validity of *T. polyfasciata* was still questionable (He et al., 2008). In this study, the numbers of these two

species of plateau loach collected were relatively small, with 10 *T. bleekeri* and 5 *T. polyfasciata*, and 7-9 horizontal stripes were observed on the sides of the fish bodies. The division into two different species was also not supported by morphology, but the genetic distance between the two species reached 8.57%, far exceeding the threshold of genetic distance within the species of 2% (Pereira et al., 2013). Therefore, it is speculated that these two species have undergone genetic differentiation in terms of genetic material, but due to the small size of the individual (the length of the collected sample is 5-8 cm), the morphological difference is not obvious, so they have historically been regarded as one species. Obviously, the body colour or body markings of the plateau loach may not be an effective classification feature for the identification of species and cannot be used as the main basis for identification.

Herzenstein (1891) identified *T. papilloso-labiatus* as a subspecies of *T. strauchii*; this finding was also supported by Zugmeyer (1910). *T. strauchii* lack a developed mastoid process similar to that of *T. papilloso-labiatus*. Instead, they have only a strong, naked fold, while the mastoid process on the upper lip of the plateau loach living in the Hexi corridor is obviously a double line, and that on the lower lip is blurred double line. Characteristics such as the mastoid process and strong, naked crease are continuously transitive in a geographical distribution without obvious boundaries. However, the appearance of significant double lines on the mastoid marks discontinuity in the variation, and there are relatively stable differences in a series of other morphological traits. Thus, *T. papilloso-labiatus* should be regarded as an independent species (Li and Chang, 1974; Zhao, 1984). This is also supported in the phylogenetic tree constructed in this study (Figure 3). *T. strauchii* and *T. papilloso-labiatus* are clustered into two different clades and should be independent species.

There is little difference in the morphological characteristics between *T. wuweiensis* and *T. scleroptera*. Li and Chang (1974) regarded *T. wuweiensis* as an independent species based on 7 morphological traits. Zhu and Wu (1975, 1981) believed that there was a certain continuity in the identification characteristics of these two species. However, after collecting specimens of *T. scleroptera* distributed in the Datonghe River, only one mountain away from the *T. wuweiensis* specimens, Zhao (1984) believed that there were significant differences between the two species in the number of pectoral fin rays, intestinal shapes and gill rakers, supporting *T. wuweiensis* as an independent species. In this study, *T. wuweiensis* and *T. scleroptera* clustered in different clades, and the two species were greatly differentiated, which also supported the idea that *T. wuweiensis* is an independent species. The low genetic diversity of *T. wuweiensis* may be due to the short time since species differentiation and the low haplotype diversity and nucleotide diversity may be caused by the founder effect and the narrow distribution area (the species is only distributed in the east and west Shiyanghe River tributaries).

T. shiyangensis, *T. papilloso-labiatus* and *T. hsutschouensis* are distributed in three inland river systems in the Hexi corridor. The maximum intra-species genetic distance of these three species is more than 1%. This may be mainly due to the wide geographic distribution of the three species and the large population differentiation caused by the barriers created by the water systems. This phenomenon also appears in the sympatric distribution of *Gymnocypris chilianensis*, in which each geographic population is clustered into a single clade, with a large genetic differentiation (Zhao et al., 2011).

The different geographic populations of some widespread species are identified as different species or subspecies due to some more significant morphological differences. For example, *T. stoliczkae* was divided into 7 subspecies (Herzenstein, 1891) due to the differences in the number of gill rakers, the proportion of quantitative traits and the number of spiral loops of intestinal tubes with changes in altitude or water system. In this study, the samples were collected in three drainage systems (Yellow River, Jialing River and the inland rivers in the Hexi corridor). The maximum genetic distance within the species was greater than 1.2% (Table 2). However, the samples of different water systems have shared haplotypes. This indicates that different geographic populations of *T. stoliczkae* in the surveyed area are from a common ancestor.

The membranous swim bladder of *T. obscura* is very developed with a constriction in the middle, and its length accounts for approximately 2/3 of the abdominal cavity. Compared with *T. orientalis*, its body surface has obvious spines. It is regarded as an independent new species (Li, 2017). In this study, a relatively large number of samples (n=234) were collected in the distribution area. The phylogenetic tree showed that the samples from different water systems were clustered into different clades, the maximum genetic distance within the species was 2%, and the nucleotide diversity and haplotype diversity were relatively high (h=0.887, π = 0.00777). These findings indicate that there is a large differentiation between the two geographically separated populations of *T. obscura* and the possibility of allopatric speciation. *T. obscura* and *T. orientalis* are also divided into two different monophyletic lines in the phylogenetic tree, which is consistent with the results of the analysis of Wu (2017).

Although only 3 specimens of *T. sp1* were collected in the Liangdang section of the Jialing River, there are obvious differences in morphological characteristics from other species of plateau loach. It should be identified as a new species that has not been reported, but more specimens should be collected for further confirmation. *T. sp2* was collected in the Jialing River, and showed degeneration of the membranous swim bladder, leaving only a small chamber, an anus near the start of the anal fin, the end of the pelvic fin adjacent to the anus, a large spot on the back of the body, a spot on the side of the body and other morphological characteristics which were obviously different from those of the closely related species *T. obscura*. A detailed description of these newly discovered species is necessary to make it possible to record the relationship between morphology and molecular identification criteria (Versteirt et al., 2015).

Conclusions

This study is the first comprehensive assessment of plateau loach species in a biodiversity hotspot using standard DNA barcoding. A high-density sample collection was carried out in this area to collect all known nominal species of plateau loach in this region. Although 14 of the 24 taxonomic species can be easily identified by DNA barcoding and classical morphological classification, 10 species pose serious challenges to standardized and automated molecular identification through mitochondrial DNA. Newly discovered species and cryptic species identified through DNA barcoding technology revealed the need for a taxonomic revision of the genus. If combined with the MOTUs identified here, the study of

morphological features can be facilitated to support the delimitation of species and classification. At the same time, using just one standardized barcode gene is not a perfect way to identify species, because their lower nucleotide sequence differences may be ignored (Ratnasingham and Hebert, 2013). It is necessary to combine nuclear markers with ecological and biological data (Ajamma et al., 2016; Waldir et al., 2018; Durand et al., 2017; Wang et al., 2018) and expand the survey area and number of species to evaluate the species boundary of the plateau loach genus. This study provides a basis for protecting the biodiversity of plateau loach.

Methods

Sample collection

The samples were collected at 114 sampling sites in two exorheic rivers (Jialing River, which is the largest clade of the Yangtze River, and the upstream of the Yellow River) and three inland water bodies (Shiyanghe River, Heihe River and Shulehe River) located on the northeastern edge of the QTP from 2015 to 2018 (Figure 1, Figure 2). The specimens were caught using gill nets and cage nets. To accurately identify the fish based on taxonomic books, the fresh specimens were examined for specific morphological characters (Zhu and Wu, 1981; Wu and Wu, 1992; Wang, 1991). The muscle tissue of each specimen was preserved in 95% ethanol for DNA extraction, and the voucher specimens were stored in 10% formaldehyde solution for further examination of specific morphological characters (Table S1).

DNA extraction, amplification and sequencing

Total genomic DNA was extracted from the muscle tissue using the high-salt method, and a segment of 651 bp from the cytochrome oxidase I gene (COI) was amplified using the published primers FishF1 (5' TCAACCAACCACAAAGACATTGGCAC3') and FishR1 (5' TAGACTTCTGGGTGGCCAAAGAATCA3') (Ward et al., 2005). The PCR amplifications were performed in 30 μL, including 21.25 μL of molecular grade water, 3.0 μL of 10× PCR buffer, 1.5 μL of each primer (10 mM), 1.5 μL of dNTPs (10 mM), 0.375 μL of Taq polymerase, and 1 μL of template DNA. The PCR conditions consisted of initial denaturation at 94°C for 5 min followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 53°C for 30 s, extension at 72°C for 1 min, and a final extension step at 72°C for 10 min followed by a hold at 4°C. The PCR products were analysed in 1% agarose gels containing ethidium bromide stain and bidirectionally sequenced using sequencing primers. The purified PCR products were sequenced on an ABI 3730 XL DNA System.

Genetic distance analyses

The sequencing chromatogram were checked by Chromas 1.45 software, and the forward and reverse sequences were assembled and edited with the SeqMan program (DNASTAR Inc., WI, USA). The sequences of each specimen generated in this study were compared and aligned using the ClustalW program. Haplotype number, haplotype diversity and nucleotide diversity were calculated with DnaSP 5.0 (Librado and Rozas, 2009). A neighbour-joining (NJ) tree, intraspecific and interspecific genetic distances were constructed based on the Kimura 2-parameter (K2P) distance model using MEGA, version 5.0with

the bootstrap support values calculated with 1000 replicates (Tamura et al., 2011). Mr Bayes 3.2.5 software was used for Bayesian inference analysis (Ronquist et al., 2003). The posterior probability represents the credibility of each clade. In the BI method, the Bayesian analysis based on codon partition (1st, 2nd, 3rd) is used, the random tree is taken as the starting tree, and a GTR+G substitution model. 50 million MCMC generations with samples saved every 500 generations. The first 10% of trees were discarded as burn-in. The maximum likelihood method was analysed with PhyML 3.0 software (Guindon et al., 2010), the substitution model was defined as GTR+G, and the number of substitution rate categories was set as 6, 100 times of self-guided method to test the confidence of each clade. These values were used to calculate the maximum, minimum and mean intraspecific and interspecific molecular operational taxonomic unit (MOTU) distances (Ramirez et al., 2017). The average K2P intraspecific distance is the average within each nominal species and the maximum K2P intraspecific distance is the largest distance observed between sequence pairs within a nominal species, and the barcoding gap was checked by the intraspecific and interspecific genetic distances (Shen et al., 2018). Finally, we constructed haplotype networks for mixed genealogies species (or main clades B) using Network 4.6 software (Bandelt et al., 1999).

Species delimitation

The species identified based on morphological characters were referred to as valid species, and species delimited by DNA sequences were referred to as MOTUs (Hutama et al., 2017; Shen et al., 2018). Four MOTUs delimitation algorithms were used to delimit species. Two analysis methods based on genetic distance and two analysis methods based on evolutionary tree topology were used to classify molecular clades or MOTUs. Among them, the Barcode of Life Data Systems (BOLD) and the Automatic Barcode Gap Discovery (ABGD) are two analysis methods based on genetic distance, and a Poisson tree processes (PTP) and the general mixed Yule-coalescent (GMYC) are two analysis methods based on evolutionary tree topology. The barcode index number (BIN) system was used to delimit MOTUs automatically in the BOLD workbench (http://v4.boldsystems.org/) (Ratnasingham and Hebert, 2013). Upload the sequence data to Bold Systems and use the Cluster Sequences in sequence analysis to get the MOTUs for the whole sequence, then Taxon ID Tree in sequence analysis is used to generate evolutionary trees delimitation species. The ABGD was used via a web interface (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html) (Puillandre et al., 2012). It is an online analysis software that automatically searches barcode gap location and classifies sequences into assumed species. The parameters Pmin set to 0.001, Pmax set to 0.1, Steps set to 10, X set to 1.5, and Nb bins set to 20. The partition result of P=0.0278 is selected as the final ABGD partition result.

A Poisson tree processes model was used to delimit species through the bPTP server (http://species.h-its.org/ptp/), including a Bayesian likelihood PTP with a maximum likelihood partition and Bayesian implementation (Zhang et al., 2013). ModelFinder in PhySuite is used to select the best Bayesian tree model (Kalyaanamoorthy et al., 2017). The *.fas file was converted into *.nex file by EasyCodeML (Gao et al., 2019), and the corresponding optimal model and corresponding parameters were set. The parameters were set to 10 million generations, and the sample was taken once every 100 generations, and the

outgroup was set. Input the *.nex file into the Mrbays to build the BI evolutionary tree. Check the BI tree parameter, when average standard deviation of split frequencies is less than 0.01, indicating that the results of the two runs have little difference and the parameter has converged. After merging the two *.p files, all the Effective sample size (ESS) statistical parameters are greater than 200, indicating that the parameters have converged. Then, Fig Tree V1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) is used to view the obtained BI Tree and convert it to Nitwick format. It's uploaded to the online bPTP, the root tree was set, the external group was removed, the algebra of reversible Markov chain is 500000, and other parameters are default for analysis.

The GMYC model in the R package Splits 1.0-19 (Fujisawa and Barraclough, 2013) was used to infer the MOTUs. The ModelFinder in Physuite was used to select nucleotide replacement model, and the hypermetric tree was constructed based on Software Beast 2.4.8 (Bouckaert et al., 2014). The assumption of a strict molecular clock was set, the system generation model is birth-death Process and a GTR+G substitution model. MCMC chain is 10 million generations, sampling once every 10 thousand generations, the first 10% of trees were discarded as burn-in. Tracer1.6 was used to evaluate the convergence of the system tree and check the *.log.txt file. All ESS>200 indicates that the system tree has converged. The *.rains.txt file that contains the tree file was import into TreeAnnotator v1.8.2, Burn-in set to 1000, and finally the Maximum clade Credibility tree was generated. Fig Tree V1.4.4 was used to convert the Tree file to Newick format, and the R package Splits 1.0-19 was use to analyse GMYC.

Abbreviations

QTP: the Qinghai-Tibet Plateau; COI: the mitochondrial cytochrome C oxidase subunit I; BOLD: the Barcode of Life Data Systems; NJ: the neighbor-joining method; ML: maximum likelihood method; BI: Bayesian inference method; K2P: Kimura 2-parameter; MOTU: molecular operational taxonomic unit; PTP: A Poisson tree processes model; GMYC: the general mixed Yule-coalescent model; BIN: the barcode index number system; ABGD: Automatic Barcode Gap Discovery.

Declarations

Ethics approval and consent to participate

All the sampling sites were not privately-owned or protected, and field sampling did not involve protected species. The collection of fishes complied with the guidelines of Gansu Fisheries Research Institute. The study was approved by the Laboratory Animal Ethics Committee of Gansu Fisheries Research Institute. All animals and experiments were conducted in accordance with the "Guidelines for Experimental Animals" of the Ministry of Science and Technology (Beijing, China).

Consent for publication

Not applicable

Availability of data and materials

Table S1 Checklist of the plateau loach including their distribution and GenBank accession numbers.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was support by the National Natural Science Foundation of China (Project 31460560); t the Outstanding Youth Foundation of Gansu province (Grant No:18JR3RA021). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in the writing of the manuscript.

Authors' contributions

LZ and DYY conceived and designed the experimental plan. WT, LZ performed experiments. WT, ZYP and YZY analysed and interpreted the sequence data. WT, DYY drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to Wang JZ, Zhang DX and Hu YB for help to obtain the samples.

References

1. Ahmed, M.I., Sabrah, M.M., Heneish, R.A., El-Alwany, M. (2016). DNA barcoding uncover cryptic diversity in goat fishes (Family: Mullidae) across the Egyptian coastal waters. Pakistan Journal of Biological Sciences 19, 65-70.

2. Almerón-Souza, F., Sperb, C., Castilho, C. L., Figueiredo, P. C. C., Goncalves, L. T., Machado, R., et al. (2018). Molecular Identification of Shark Meat From Local Markets in Southern Brazil Based on DNA Barcoding: Evidence for Mislabeling and Trade of Endangered Species. Frontiers in Genetics 9,138.

3. Ajamma, Y. U., Villinger, J., Omondi, D., Salifu, D., Onchuru, T. O., Njoroge, L., et al. (2016). Composition and genetic diversity of mosquitoes (Diptera: Culicidae) on Islands and Mainland shores of Kenya's lakes Victoria and Baringo. Journal of Medical Entomology 53, 1348-1363.

4. Bandelt, H.J., Forster, P., Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16, 37-48.

5. Barley, A. J., Thomson, R. C. (2016). Assessing the performance of DNA barcoding using posterior predictive simulations. Molecular Ecology 25,1944-1957.

6. Barrett, R. D. H., Hebert, P. D. N. (2005). Identifying spiders through DNA barcodes. Canadian Journal of Zoology, 83, 481-491.

7. Beebe, N. W. (2018). DNA barcoding mosquitoes: Advice for potential prospectors. Parasitology, 145, 622-633.

8. Berbel-Filho, W.M., Ramos, T. P. A., Jacobina, U. P., Maia, D. J. G., Torres, R.A., Lima, S. M. Q. (2018). Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil. Journal of fish biology 93, 311-323.

9. Bhattacharjee, M.J., Laskar, B.A., Dhar, B., Ghosh, S.K. (2012). Identification and reevaluation of freshwater catfishes through DNA barcoding. Plos One 7, e49950.

10. Boucaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Suchard, M.A., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. Plos Computational Biology 10, e1003537.

11. Burrows, M., Browning, J. S., Breitbart, M., Murawski, S. A., Peebles, E. B. (2019). DNA barcoding reveals clear delineation between spawning sites for neritic versus oceanic fishes in the Gulf of Mexico. Fisheries Oceanography 12, 228-239.

12. Brown, S.D.J., Collins, R.A., Boyer, S., Lefort, M.C. (2012). SPIDER: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Molecular Ecology Resources 12, 562-565.

13. Cardinale, B., Duffy, J., Gonzalez, A., Hooper, D., Perrings, C., Venail, P., et al. (2012). Biodiversity loss and its impact on humanity. Nature 486, 59–67.

14. Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. Mitochondrial DNA 22, 80-86.

15. Carstens, C., Pelletier, T. A., Reid, N. M., Satler J. D. (2013). How to fail at species delimitation. Molecular Ecology 22, 4369-4383.

16. Chan-Chable, R.J., Martínez-Arce, A., Mis-Avila, P., Ortega-Morales, A.I. (2019). DNA barcodes and evidence of cryptic diversity of anthropophagous mosquitoes in Quintana Roo, Mexico. Ecology and Evolution 9, 4692-4705.

17. Chan-Chable, R. J., Ortega-Morales, A. I., & Martínez-Arce, A. (2016). First record of Psorophora albipes in Quintana Roo, México. Journal of the American Mosquito Control Association 32, 237-239.

18. Chen, J., Yang, L., Zhang, R., Uebbing, S., Zhang, C., Jiang, H., et al. (2020). Transcriptome-Wide Patterns of the Genetic and Expression Variations in Two Sympatric Schizothoracine Fishes in a Tibetan Plateau Glacier Lake. Genome Biology and Evolution 12, 3725-3737.

19. Chen, Y. (1998). Hengduan Mountains fishes. Science and technology press, Beijing (In Chinese).

20. Collins, R. A., Cruickshank, R. H. (2013) The seven deadly sins of DNA barcoding. Molecular Ecology Resources 13, 969.

21. Dhar, B., Ghosh, S. (2017). Mini-DNA barcode in identification of the ornamental fish: a case study from Northeast India. Gene 627, 248-254.

22. Dincă, V., Zakharov, E. V., Hebert, P. D. N., Vila R. (2011). Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. Proceedings of the Royal Society B: Biological Sciences 278, 347-355.

23. Ding, R., Fang, S., Fang, J. (1996). Studies on the DNA fingerprinting in two species of the gunes *Triplophysa* from China with description of a new species. Sichuan Journal of Zoology 15, 10-14.

24. Durand, J.-D., Hubert, N., Shen, K.-N., Borsa, P. (2017). DNA barcoding grey mullets. Reviews in Fish Biology and Fisheries 27, 233-243.

25. Ence, D. D. and Carstens, B. C. (2011). SpedeSTEM: a rapid and accurate method for species delimitation. Molecular Ecology Resources 11, 473-480.

26. Esselstyn, A., Evans, B. J., Sedlock, J. L., Anwarali Khan, F. A., Heaney, L. R. (2012). Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. Proceedings of the Royal Society B: Biological Sciences 279, 3678-3686.

27. Fontenot, B.E., Makowsky, R., Chippindale, P.T., (2011). Nuclear–mitochondrial discordance and gene flow in a recent radiation of toads. Molecular Phylogenetics and Evolution 59,66-80.

28. Frezal, L., Leblois, R. (2008). Four years of DNA barcoding: Current advances and prospects. Infection Genetics and Evolution 8, 727-736.

29. 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 62,707-724.

30. Gao, F. L., Chen, C. J., Arab, D. A., Du, Z. G., He, Y. H., Ho, S.Y.W. (2019). EasyCodeML: A visual tool for analysis of selection using CodeML. Ecology and Evolution 9,3891-3898.

31. Guindon, S., Anisimova, M., Lefort, V., Hordijk, W., Gascuel, O., Dufayard, J.F. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59, 307-321.

32. Gustafsson D. R., Price D. A., Erseus C. (2009). Genetic variation in the popular lab worm Lumbriculus variegatus (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation. Molecular Phylogenetics and Evolution 51,182-189.

33. He, C., Song, Z., Zhang, E. (2011). Triplophysa fishes in China and the status of its taxonomic studies. Sichuan Journal of Zoology 30,156-161 (In Chinese).

34. Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the royal society B: Biological sciences 270, 313-321.

35. Herzenstein, S.M. (1891). Fische. In: wissenschaftliche resultate der von N.M. Przewalski nach Central-Asien unternommenen Reisen. Zoologischer Theil., Band III. 3, 1-262.

36. Hubert, N., Hanner, R., Holm, E., Mandrak, N.E., Taylor, E., Burridge, M., et al., (2008). Identifying Canadian freshwater fishes through DNA barcodes. Plos One 3, e2490.

37. Hutama, A., Dahruddin, H., Busson, F., Sauri, S., Keith, P., Hadiaty, R.K., et al., (2017). Identifying spatially concordant evolutionary significant units across multiple species through DNA barcodes:

Application to the conservation genetics of the freshwater fishes of Java and Bali. Global Ecology and Conservation 12, 170-187.

38. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Haeseler, A.V., Jermiin, L. S. (2017), ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14,587-589.

39. Keskin, E., Atar, H.H. (2013). DNA barcoding commercially important fish species of Turkey. Molecular Ecology Resources 13, 788-797.

40. Khan, G., Zhang, F., Gao, Q., Fu, P., Zhang, Y., Chen, S. (2018). Spiroides shrubs on Qinghai-Tibetan Plateau: Multilocus phylogeography and palaeodistributional reconstruction of *Spiraea alpina* and *Mongolica* (Rosaceae). Molecular Phylogenetics and Evolution 123,137-148.

41. Knowles, L. L. and Carstens, B. C. (2007). Delimiting species without monophyletic gene trees. Systematic Biology 56, 887-895.

42. Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W., Reich, P., et al. (2011). High plant diversity is needed to maintain ecosystem services. Nature 477, 199-202.

43. Lakra, W.S., Verma, M.S., Goswami, M., Lal, K.K., Mohindra, V., Punia, P., et al., (2011). DNA barcoding Indian marine fishes. Molecular Ecology Resources 11, 60-71.

44. Leavitt, D.H., Marion, A.B., Hollingsworth, B.D., Reeder, T.W., (2017). Multilocus phylogeny of alligator lizards (Elgaria, Anguidae): Testing mtDNA introgression as the source of discordant molecular phylogenetic hypotheses. Mol. Phylogenet. Evol. 110, 104-121.

45. Li, S. (2017). Fishes of the Yellow River. China ocean university press, Qingdao (In Chinese).

46. Li, J., Wang, Y., Jin, H., Li, W., Yan, C., Yan, P., et al. (2017). Identification of *Triplophysa* species from the Qinghai-Tibetan Plateau (QTP) and its adjacent regions through DNA barcodes. Gene 20, 12-19.

47. Li, S. and Chang, S. (1974). Two new species and one new subspecies of fishes from the northern part of Kansu province, China. Acta Zoologica Sinica 20, 414-419 (In Chinese).

48. Librado, P. and Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451-1452.

49. Lim G. S. and Meier B. R. (2011) Determining Species Boundaries in a World Full of Rarity: Singletons, Species Delimitation Methods. Systematic Biology 61, 165-169.

50. Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J., Hector, A., et al. (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294, 804-808.

51. Mardulyn, P., Othmezouri, N., Mikhailov, Y. E., & Pasteels, J. M. (2011). Conflicting mitochondrial and nuclear phylogeographic signals and evolution of host-plant shifts in the boreo-montane leaf beetle *Chrysomela lapponica*. Molecular Phylogenetics and Evolution 61, 686-696.

52. Mccusker, M.R., Denti, D., Guelpen, L.V., Kenchington, E., Bentzen, P. (2012). Barcoding Atlantic Canada's commonly encountered marine fishes. Molecular Ecology Resources 13, 177-188.

53. McCloud D. S. (2010). Of least concern? Systematics of a cryptic species complex: *Limnonectes kuhlii* (Amphibia: Anura: Dicroglossidae). Molecular Phylogenetics and Evolution 56, 991-1000.

54. Meier R., Shiyang K., Vaidya G., Ng P.K.L. (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Systematic Biology 55, 715-728.

55. Pecnikar, Z. F., Buzan, E. V. (2014) 20 years since the introduction of DNA barcoding: from theory to application. Journal of Applied Genetics 55,43-52.

56. Pentinsaari, M., Vos, R.A., Mutanen, M. (2017). Algorithmic single-locus species delimitation: Effects of sampling effort, variation and nonmonophyly in four methods and 1870 species of beetles. Molecular Ecology Resources 17, 393-404.

57. 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.

58. Pereira, L. H., Maia, G. M., Hanner, R., Foresti, F., Oliveira, C. (2011). DNA barcodes discriminate freshwater fishes from the Paraíba do Sul River Basin, São Paulo, Brazil. Mitochondrial DNA 22(Suppl. 1), 71-79.

59. Peterson, A. T., Ortega-Huerta, M. A., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, R. H., Stockwell, D. R. B. (2002). Future projections for Mexican faunas under global climate change scenarios. Nature 416, 626-629.

60. Pigliucci, M. (2005). Evolution of phenotypic plasticity: Where are we going now? Trends in Ecology and Evolution 20, 481-486.

61. Prokofiev, A. M. (2007). Materials towards the revision of the genus *Triplophysa* Rendahl, 1933 (Cobitoidea: Balitoridae: Nemacheilinae): A revision of nominal taxa of Herzenstein (1888) described within the species "*Nemachilus*" *stoliczkae* and "" *dorsonotatus*, with the description of the new species *T. scapanognatha* sp. nova. Journal of Ichthyology 47,1-20.

62. Puillandre, N., Lambert, A., Brouillet, S., Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. Molecular ecology 21, 1864-1877.

63. Ramirez, J.L., Birindelli, J.L., Carvalho, D.C., Affonso, P.R.A.M., Venere, P.C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated Neotropical Ichthyofauna: DNA barcoding in the recently described Genus *Megaleporinus* (Characiformes: Anostomidae). Frontiers in Genetics 8,149.

64. Ramirez, J. L., Galetti, P. M. Jr. (2015). DNA barcode and evolutionary relationship within *Laemolyta* Cope 1872 (Characiformes: Anostomidae) through molecular analyses. Molecular Phylogenetics and Evolution 93, 77-82.

65. Ratnasingham, S., Hebert, P.D.N. (2007). BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular ecology notes 7, 355-364.

66. Ratnasingham, S., Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. Plos One 8, e66213.

67. Robinson, B. W., Parsons, K. J. (2002). Changing times, spaces, and faces: tests and implications of adaptive morphological plasticity in the fishes of northern postglacial lakes. Canadian Journal of Fisheries and Aquatic Sciences 11, 1819-1833.

68. Ronquist, F., Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572-1574.

69. Sales, N. G., Stefano, M., Salvador, G. N., Pessali, T. C., Carvalho, D. C. (2018). Hidden diversity hampers conservation efforts in a highly impacted neotropical river system. Frontiers in Genetics 9, 271.

70. Shaw, K. L. (2002). Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proceedings of the National Academy of Sciences of the United States of America, 99, 16122-16127.

71. Shen, Y., Hubert, N., Huang, Y., Wang, X., Gan, X., Peng, Z., et al. (2019). DNA barcoding the ichthyofauna of the Yangtze River: insights from the molecular inventory of a mega-diverse temperate fauna. Molecular Ecology Resources 5,1-14.

72. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28, 2731-2739.

73. Thibert-Plante, X., Hendry, A. P. (2011). The consequences of phenotypic plasticity for ecological speciation. Journal of Evolutionary Biology 2, 326-342.

74. Valdez-Moreno, M., Ivanova, N.V., Elías-Gutiérrez, M., Contera-Balderas, S., Hebert, P.D.N. (2009). Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes. Journal of Fish Biology 74, 377-402.

75. Versteirt, V., Nagy, Z. T., Roelants, P., Denis, L., Breman, F. C., Damiens, D., et al. (2015). Identification of Belgian mosquito species (Diptera: Culicidae) by DNA barcoding. Molecular Ecology Resources 15, 449-457.

76. Vieites D. R., Wollenberg K. C, Andreone F., Kohler J., Glaw F., Vences M. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. Proceedings of the National Academy of Sciences of the United States of America 106, 8267-8272.

77. Wang, L., Wu, Z., Liu, M., Liu, W., Zhao, W., Liu, H., et al. (2018). DNA barcoding of marine fish species from Rongcheng Bay, China. PeerJ-life and Environment 6, e5013.

78. Wang, X. (1991). Vertebrate fauna of Gansu. Gansu science and technology press, Lanzhou (In Chinese).

79. Wang, Y., Shen, Y., Feng C., Zhao, K., Song, Z., Zhang, Y., et al. (2016). Mitogenomic perspectives on the origin of Tibetan loaches and their adaptation to high altitude. Scientific Reports 6, 29690.

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

81. Wu, Y. and Wu, C. (1992). The fishes of the Qinghai-Xizang plateau. Sichuan science and technology press, Chengdu (In Chinese).

82. Wu, Y. (2017). The classification of *Triplophysa stenura*, *hialmari*, *T. alexandrae*, *T. lixianensis*, *T.brevibarba*; *T.orientalis*, *T.hutjertjuensis*, *T.obscura* and a new species of the genus *Triplophysa* from Sichuan. China West Normal University, Nanchong, Sichuan Province, China (In Chinese).

83. Zhang, Y., Qin, G., Zhang, H., 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.

84. Zhang, J. and Hanner, R. (2012). Molecular approach to the identification of fish in the South China Sea. Plos One 7, e30621.

85. Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics 29, 2869-2876.

86. Zhao, K., Duan, Z., Pen, Z., Gan, X., Zhang, R., He, S., et al. (2011). Phylogeography of the endemic *Gymnocypris chilianensis* (Cyprinidae): Sequential westward colonization followed by allopatric evolution in response to cyclical Pleistocene glaciations on the Tibetan Plateau. Molecular Phylogenetics and Evolution 59, 303-310.

87. Zhao, T. (1984). Taxonomic problems of some *Newachilus* Zoological Research 5, 341-346 (In Chinese).

88. Zhu, S. and Wu, Y. (1975). Ichthyo-fauna in Qinghai lake area and biology of *Gymnocypris przewalskii*. Science Press, Beijing (In Chinese).

89. Zhu, S. and Wu, Y. (1981). A new species and a new subspecies of loaches of the genus *Nemachilus* from Qinghai province. Acta Zootaxonomica Sinica 6, 221-224.

90. Zugmayer, E. (1910). Beitrage zur Ichthyologie von Zentral-Asien. Zool. Jahrb. Syst. Geog. U. Biol. 29, 275-298.

Tables

Table 1 Sampling information number of individuals and diversity parameters for the specimens included in this study.

Species	Collection site (River)	Numbure of specimens（N）	Numbure of haplotypes (Nh)	Haplotype diversity (h)	Nucleotide diversity (π)
Triplophysa bleekeri (Sauvage *et* Dabry, 1874)	Jialing River	5	1	——	——
T. dalaica (Kessler, 1876)	Jinghe River	55	9	0.469±0.083	0.00114±0.00102
T. hsutschouensis (Rendahl, 1933)	Heihe River, Shulehe River, Shiyanghe River	46	8	0.731±0.041	0.00440±0.00125
T. leptosoma (Herzenstein, 1888)	Shulehe River	7	3	0.667±0.160	0.00126±0.00095
T. minxianensis (Wang et zhu, 1979)	Yellow River, Jinghe River	253	21	0.385±0.040	0.00074±0.00001
T. minxianensis sp1	Yellow River, Weihe River	20	2	0.526±0.036	0.00087±0.00047
T. obscura (Wang, 1987)	Jialing River, Weihe River, Taohe River	234	27	0.877±0.012	0.00777±0.00145
T. orientalis (Herzenstein, 1888)	Taohe River	19	1	——	——
T. papillosolabiatus (Kessler, 1879)	Heihe River, Shulehe River	95	13	0.603±0.036	0.00201±0.00112
T. pappenheimi (Fang, 1935)	Yellow River, Weihe River	21	4	0.610±0.114	0.00196±0.00011
T. polyfasciata (Ding, 1996)	Jialing River	10	3	0.600±0.131	0.00121±0.00082
T. pseudoscleroptera (Zhu & Wu, 1981)	Yellow River, Xiahe River	9	4	0.583±0.183	0.00138±0.00105
T. robusta (Kessler, 1876)	Yellow River, Jialing River, Shiyanghe River	219	46	0.936±0.006	0.00588±0.00182
T. robusta sp1	Jinghe River	194	15	0.591±0.020	0.00133±0.00012
T. scleroptera (Herzenstein, 1888)	Yellow River, Taohe River	44	3	0.090±0.059	0.00015±0.00005
T. sellaefer (Nichols, 1925)	Jinghe River	22	5	0.338±0.128	0.00089±0.00101
T. shiyangensis (Zhao & Wang, 1983)	Shiyang River	25	11	0.770±0.086	0.00367±0.00180
T. siluroides (Herzenstein, 1888)	Yellow River, Xiahe	28	5	0.529±0.105	0.00108±0.00095

	River, Taohe River				
T. sp1	Yellow River, Jialing River	3	2	0.667±0.314	0.00660±0.00269
T. sp2	Jialing River	67	6	0.172±0.062	0.00029±0.00007
T. stoliczkai (Steindachner, 1866)	Yellow River, Xiahe River, Taohe River, Jinghe River, Jialing River, Shiyanghe River	129	25	0.812±0.028	0.00285±0.00149
T. strauchii (Kessler, 1874)	Heihe River	11	2	0.509±0.101	0.00084±0.00056
T. tenuis (Day, 1877)	Heihe River, Shulehe River	97	11	0.378±0.061	0.00082±0.00010
T. wuweiensis (Li & Chang, 1974)	Shiyanghe River	17	3	0.404±0.130	0.00090±0.00085
Total		1630	230		

Table 2 Genetic K2P distances of the *Triplophysa* species.

species	MOTU	Mean intra-	Maximum intra-	NN Dist	NN
Triplophysa bleekeri	MOTU-1	0.0000	0.0000	0.0731	*T. papillosolabiatus*
T. shiyangensis	MOTU-3	0.0069	0.0130	0.0265	*T. stoliczkae*
T. strauchii	MOTU-4	0.0017	0.0020	0.0271	*T. stoliczkae*
T. orientalis	MOTU-5	0.0000	0.0000	0.0598	*T. stoliczkae*
T. tenuis	MOTU-7	0.0039	0.0070	0.0613	*T. pseudoscleroptera*
T. wuweiensis	MOTU-8	0.0033	0.0050	0.0751	*T. obscura*
T. sp2	MOTU-9	0.0028	0.0030	0.0264	*T. obscura*
T. obscura	MOTU-10	0.0101	0.0200	0.0264	*T. sp2*
T. polyfasciata	MOTU-11	0.0022	0.0030	0.0857	*T. bleekeri*
T. sp1	MOTU-12	0.0099	0.0100	0.0290	*T. papillosolabiatus*
T. leptosoma	MOTU-13	0.0022	0.0030	0.0147	*T. papillosolabiatus*
T. papillosolabiatus	MOTU-14	0.0053	0.0130	0.0147	*T. leptosoma*
T. sellaefer	MOTU-16	0.0036	0.0070	0.0440	*T. minxianensis* sp1
T. hsutschouensis	MOTU-18	0.0063	0.0130	0.0320	*T. robusta*
T. scleroptera	MOTU-6	0.0030	0.0050	0.0040	*T. pseudoscleroptera*
T. pseudoscleroptera	MOTU-6	0.0040	0.0070	0.0040	*T. scleroptera*
T. stoliczkae	MOTU-2	0.0060	0.0120	0.0130	*T. dalaica*
T. dalaica	MOUT-2	0.0040	0.0070	0.0130	*T. stoliczkae*
T. robusta		0.0310	0.0790		
	MOTU-19	0.0082	0.0250	0.0320	*T. hsutschouensis*
	MOTU-15	0.0077	0.0180	0.0000	*T. pappenheimi, T. siluroides*
T. minxianensis		0.0160	0.0740	0.0110	*T. siluroides*
	MOTU-17	0.0016	0.0020	0.0430	*T. robusta*
	MOTU-15	0.0050	0.0100	0.0000	*T. siluroides, T.pappenheimi*
T. siluroides	MOTU-15	0.0030	0.0050	0.0000	*T. pappenheimi, T. robusta*
T. pappenheimi	MOTU-15	0.0020	0.0030	0.0000	*T. siluroides, T. minxianensis*

The mean and the maximum of intra-group distances, the nearest neighbor (NN), and the minimum distance to the NN for the Nominal species.

Figures
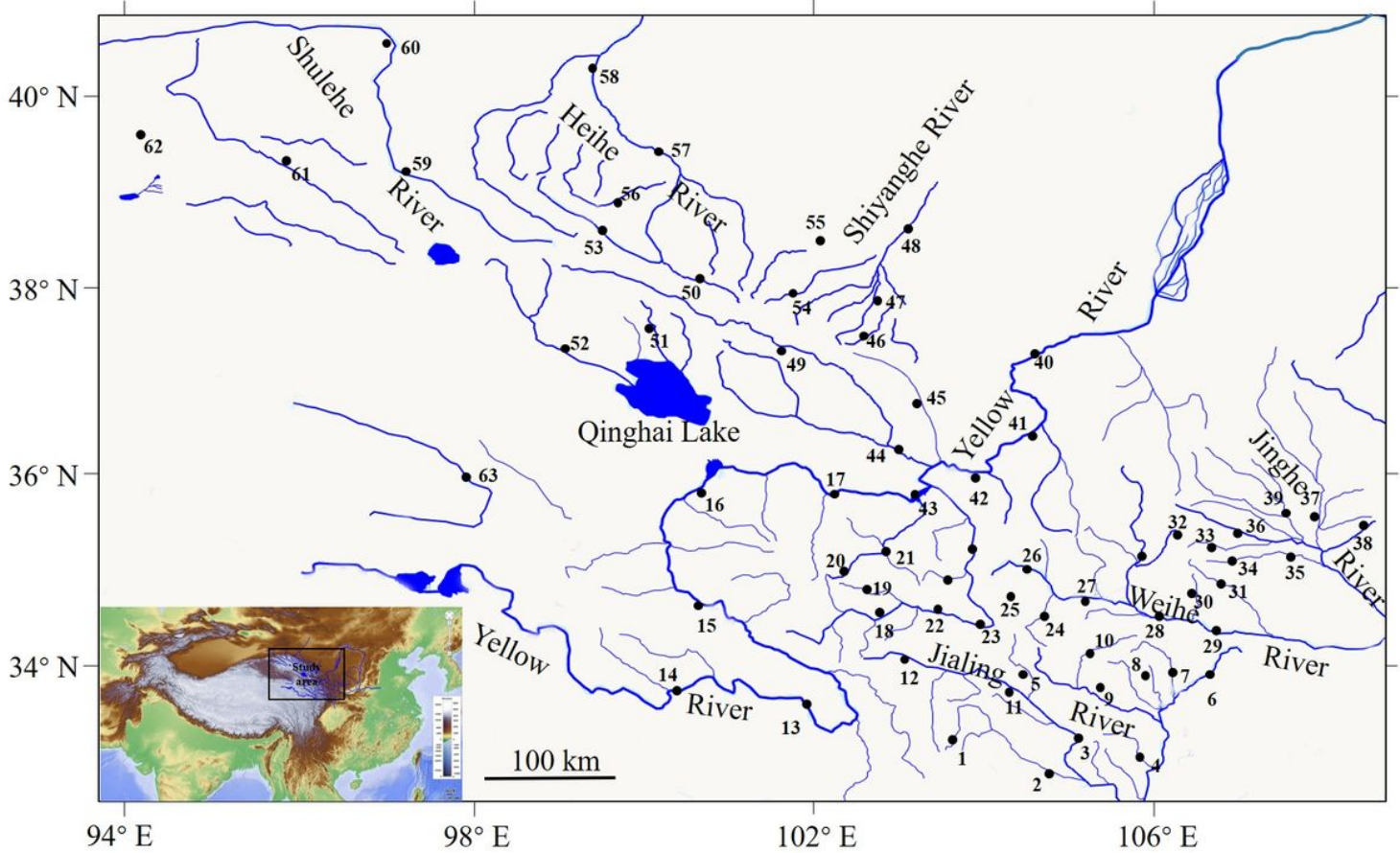
Figure 1

Collection sites. Details of the 111 sites and collected specimens are provided in Table S1. (The sectors name of sampling sites: 1. Jiuzhai; 2.Wenxian; 3.Wudou; 4.Kangxian; 5.Tanchang; 6.Liangdang; 7.Huixian; 8.Chengxian; 9.Xihe; 10.Lixian; 11.Zhouqu; 12.Diebu; 13.Maqu; 14.Hongyuan; 15.Henan; 16.Longyangxia; 17.Jishishan; 18.Luqu; 19.Hezuo; 20.Xiahe; 21.Linxia; 22.Zhuoni; 23.Minxian; 24.Wushan; 25.Zhangxian; 26.Weiyuan; 27.Gangu; 28.Qinzhou; 29.Maiji; 30.Qingshui; 31.Zhangjiachuan; 32.Jingning; 33.Chongxin; 34.Huating; 35.Lingtai; 36.Kongtong; 37.Xifeng; 38.Ningxian; 39.Zhenyuan; 40.Wufo; 41.Pingchuan; 42.Lanzhou; 43.Yongjing; 44.Minhe; 45.Yongdeng; 46.Zhuanglang; 47.Liangzhou; 48.Minqin; 49.Menyuan; 50.Arou; 51.Gangcha; 52.Tianjun; 53.Qilian; 54.Huangcheng; 55.Jinchang; 56.Sunan; 57.Linze; 58.Gaotai; 59.Yumen; 60.Guazhou; 61.Subei; 62.Akesai. This base map is from 91 Vita Assistant software http://www.91weitu.com/index.htm, edited in Adobe Photoshop CS5 software.)

Figure 2

Studied specimens of Triplophysa. (1. T. dalaica G003; 2. T. stoliczkai G0070; 3. T. polyfasciata G0187; 4. T. bleekeri G0195 ;5. T. robusta G0531; 6. T. obscura G0822; 7. T. pappenheimi G0852; 8. T. siluroides G0873; 9. T. hsutschouensis G0915; 10. T. minxianensis GS0213; 11. T. pseudoscleroptera GS0216; 12. T. scleroptera GS0230;13. T. strauchii GS0273; 14. T. papillosolabiatus GS0305; 15. T. wuweiensis GS0381;

16. T. orientalis GS0400; 17. T. shiyangensis GS0432; 18. T. leptosoma GS0441; 19. T. tenuis GS0500; 20. T. sellaefer GS0560; 21. T. sp1 GS562; 22. T. sp2 GS565. Scale bars equal 1 cm)
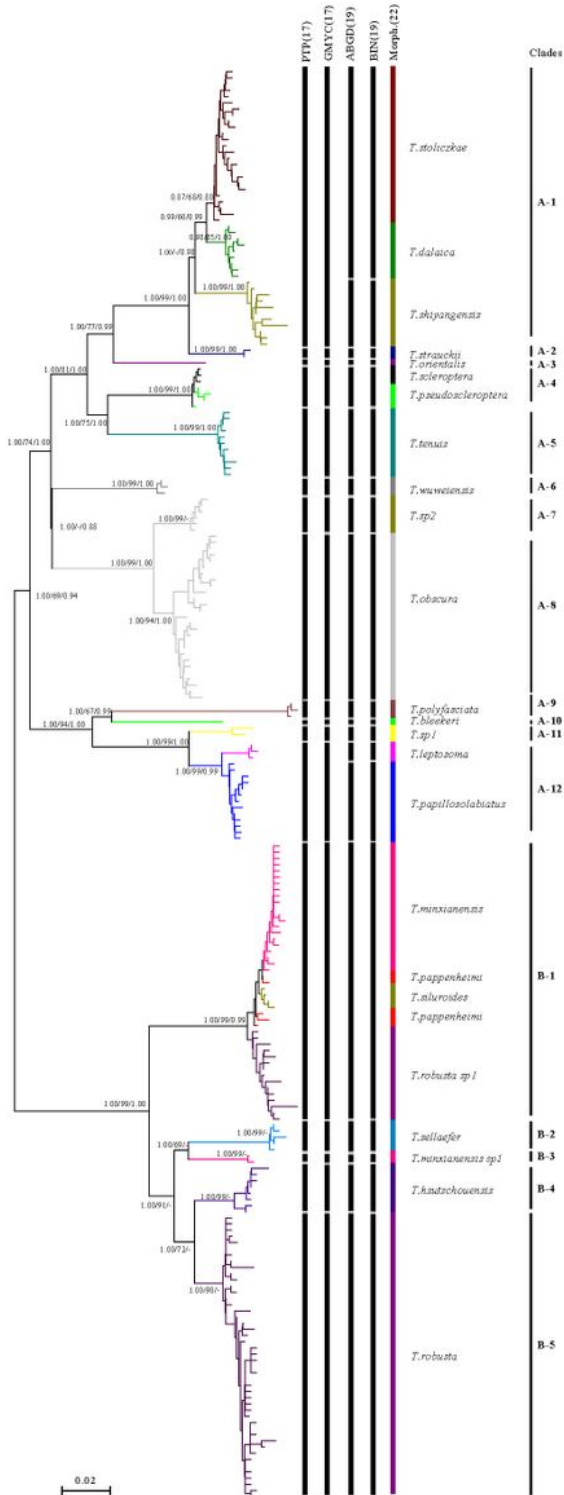
Figure 3

The 50% majority-rule consensus tree showing the clustering of the MOTUs obtained by the four MOTUs delimitation algorithms. (The values at the node represent support values in NJ/ML/BI analysis respectively (NJ bootstrap supports values below 0.50, ML bootstrap support values below 50 and

Bayesian posterior probabilities below 0.50 are not shown). The length of clade indicates the percentage of divergence. The ruler at the bottom represents a 0.02 replacement for each site. The branch colours match those in the right coloured segmented bar, whose coloured segments represent morphological species. And the four black segments bar on the left relative to species delimitation methods delineate the entities detected by each method. The finer black segments bar on the far right represents the division of the main clades.)
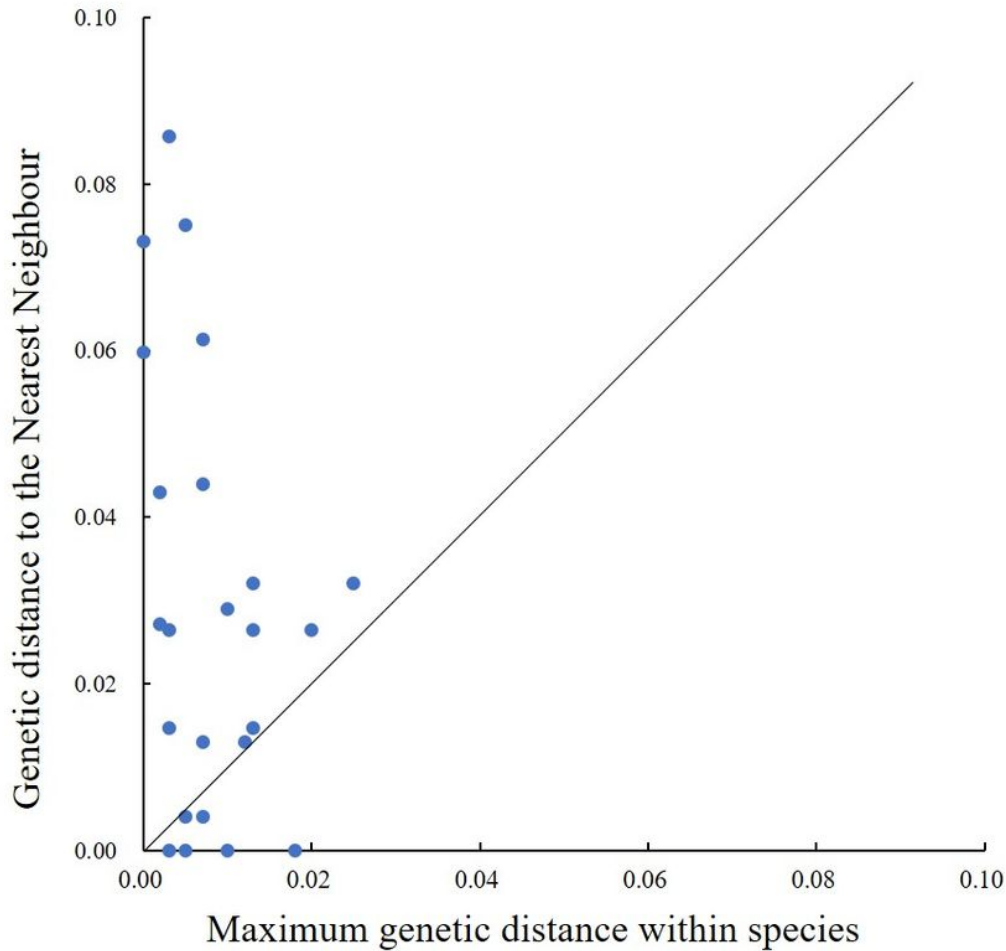
Figure 4

Relationship between maximum genetic distance within species and nearest neighbor genetic distance among species.
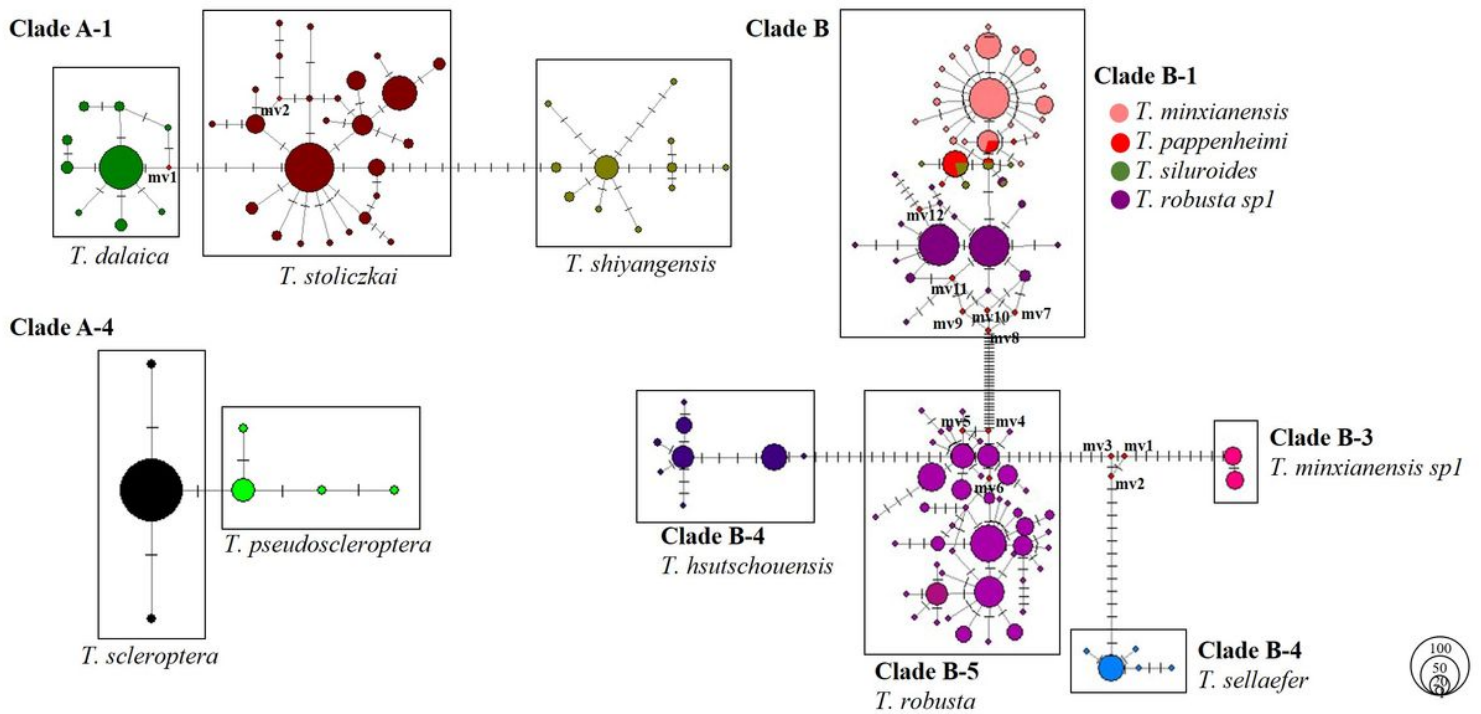
Figure 5

The haplotype networks for the species group involved in mixed genealogies. (The area of circles is proportional to the haplotype frequencies, and mv1-mv12 are missing haplotypes. Lines linking haplotypes indicate the evolutionary paths among haplotypes, vertical bars on the linking lines represent mutation steps between haplotypes)