Species diversity of freshwater shrimp in Henan Province, China, based on morphological characters and COI mitochondrial gene

Chuanjiang Zhou[1], Mengxia Feng[1], Yongtao Tang[1], Changxing Yang[1], Xiaolin Meng[1], and Guoxing Nie[1]

[1]Affiliation not available

January 8, 2021

Abstract

Freshwater shrimp is an extremely rich species group with a long and problematic taxonomic history, attributed to its wide distribution, numerous species and similar morphology. Shrimp diversity and species identification is utmost important for fisheries management. However, identification based on morphological characteristics is difficult and complex for a non-specialist to perform. The water system of Henan Province is relatively abundant, but there are few investigations of freshwater shrimps and no description of molecular features. The aim of this study was to uncover the species diversity and phylogenetic of freshwater shrimp in Henan province by combining morphological identification and molecular species delimitation (barcoding gene: COI gene). About 1,200 freshwater shrimp samples from 46 sampling sites were collected for preliminary traditional morphological identification, 222 samples of these were been further microscopic examination and molecular delimitation. Here we used tree based method (NJ, ML) and distance based method (ABGD, bPTP) mainly to define species, detect the cryptic species and assess the validity of the barcoding in molecular. Comprehensive morphological identification and molecular delimitation results, there were 9 effective species and more than one cryptic species of freshwater shrimp in the province and moreover all of them can be identified by DNA barcoding. The results of morphological identification and molecular identification show high consistency, which proves the high efficiency in freshwater shrimp species identification of the DNA barcoding and the presence of cryptic species.

Introduction

Freshwater prawns mainly refer to certain species of Caridean (Decapoda: Caridea). It is a highly species-rich group with a long taxonomic history, but at the same time its taxonomic status is controversial. There are about 770–800 Caridea species in freshwater, accounting for about one-fifth of the shrimp described (*De Grave et al., 2015*). At present, freshwater shrimps are only exist in 7 families of 38 currently Caridea families (*De Grave, Li, Tsang, Chu, & Chan, 2014*). Among them, the two families of Atyidae and Palaemonidae are quantitatively dominate, comprising 443 and 300 species respectively, which account for 97.4% of all freshwater shrimp species (*De Grave et al., 2015*). Shrimp play an important role not only as an important component of biodiversity but as a very good source of animal protein for people. In addition, freshwater shrimp has high economic value, nutritional value and research significance (*Holthuis, 1980* ;*New & Nair, 2012*). Therefore, the taxonomic identification and species identification of shrimp is one of the most important tasks for all kinds of related biological research and fishery resource conservation and management(*Shen, Guan, Wang, & Gan, 2016*) .

There have been many studies on freshwater prawn fauna, but it is poorly known of Henan Province, China. So far, 8 species of shrimps have been reported. The survey used traditional morphological recognition methods to identify and morphologically describe 352 samples of 8 species that from 15 sampling points (*F. Wang, 1989*). Henan province is located in inland, there are four major river systems, such as the Yellow River, the Yangtze River, the Huaihe River and the Haihe River. But studies on the freshwater prawns of

this area are rather scanty. In view of the above, it is of great significance to enrich shrimp related researches in Henan province to append the list of shrimp species and assess the biodiversity of this area.

The study of species diversity is the basis of all biological research, but at the same time it is also a huge challenge and a harsh burden(*P. D. Hebert, Cywinska, Ball, & deWaard, 2003*). As the main method of species diversity research, traditional morphological identification has high requirements and restrictions on samples and researchers, and the identification results are greatly affected by subjective and objective factors(*Carvalho, Neto, Brasil, & Oliveira, 2011* ; *P. D. Hebert, Cywinska, et al., 2003* ; *Shen et al., 2016*). In recent years, DNA barcode technology has developed rapidly and has gradually become one of the main methods for biological identification. With the implementation of the Barcoding of Life project, DNA barcodes have been widely recognized as a basic tool for species identification, mitochondrial gene cytochrome C oxidase I (*COI*) serves as the core of the global animal biometric system could effectively distinguish Crustacea (*Costa et al., 2007* ;*P. D. Hebert, Cywinska, et al., 2003* ; *P. D. Hebert, Ratnasingham, & deWaard, 2003*). However, the successful application of DNA barcodes is also affected by many factors, as an important part of DNA barcode technology, the establishment of a standard life barcode library is of great significance. In the era of high-throughput sequencing, there is high probability of tentative, incorrect, or low-quality sequences being submitted to databases (*L. L. Wong et al., 2011*). Therefore, compared with the commonly used barcode databases GenBank (NCBI), DDBJ, EMBL, etc., the BOLD database, which conducts strict review and screening of submitted data, is relatively more accurate and applicable(*Macher, Macher, & Leese, 2017* ; *Z. Wang et al., 2009*). In addition, as the fundamental units in biological research, species can be defined on the basis of various operational criteria. At the same time, with the acquisition of a large number of barcodes, there has been growing use of molecular approaches for species delimitation (*P. D. N. Hebert & Gregory, 2005* ; *Luo, Ling, Ho, & Zhu, 2018*). At present, sequence analysis for species delimitation mainly includes tree based method, distance based method and character based method, of which the first two have higher universality and applicability (*Birch, Walsh, Cantrill, Holmes, & Murphy, 2017*). The combined use of multiple molecular methods will improve the accuracy of the species definition results. Therefore, as many different types of molecular methods as possible should be used for comprehensive species definition.

In fact, both morphological identification and molecular definition are supporting tool for species identification. The combination of multiple methods will produce more objective and real species identification results, the species boundaries will be clearer. The main aim of this study was to objective and truthful assessment the shrimp diversity at both taxonomical and molecular level, and provides helpful information for future conservation and fisheries resource management of the shrimp in Henan.

Materials and methods

Ethics statement

The study conformed with the National Institutes of Health guide for the Care and Use of Laboratory Animals (NIH Publication No. 85-23, 1996) (2011).

Sample collection

According to the distribution of water systems in Henan Province, a total of 46 sampling sites were covered in this survey for collecting freshwater shrimp (Fig.1). The sample points cover the main streams or some tributaries of the four major rivers of the province (Table S1). In this study, about 1200 samples from 11 species, 6 genera and 4 families of Amphipoda and Decapoda were collected. As many individuals per species as possible were obtained for this study. However, in some cases, only one few individuals per region per species could be collected. Most of the prawns were collected by the shrimp traps, but a lot of individuals were captured by market buying. The collections were preserved in 95% ethanol for subsequent morphological observation and molecular identification, part of the samples was fixed in 10% formalin, and all samples were stored in the Fisheries College of Henan Normal University.

Morphological identification

Morphological identification was mainly classified in situ by visual inspection, then detailed morphological identification and classification was in laboratory by stereomicroscope microscopic examination. All the collected prawn species were taxonomically classified according to Liu (1955), Liang (2004) and Li et al. (2007) mainly based on the distinguishing morphological characters of the male collected specimens(*X. Li, Liu, Liang, & Chen, 2007* ; *Liang, 2004* ; *Liu, 1955*).

According to the morphological identification results, multiple representative individuals of per taxonomic group were selected for abdominal muscle sampling, and the sampling process should as careful as possible to avoid damaging the complete shape of the specimen. The obtained tissue samples were immediately stored in 95% ethanol and numbered for DNA extraction (While ensuring the specimen coverage of each species, individuals with moderate body size should be selected as far as possible for EP tube preservation and numbering; the larger individuals were marked by winding coils).

DNA extraction, amplification, and sequencing

About 30 μL Genomic DNA was extracted by phenol-chloroform (*Sambrock & Russel, 2001*) from tissue muscle (0.1-0.15g) and verified using 1.0% agarose gel electrophoresis.

The amplification of *COI* gene was carried out by polymerase chain reaction (PCR). A 658bp fragment was amplified using the forward primer (LCO1490: 5'-GGTCAACAAATCA TAAAGATATTGG-3') and reverse primer (HCO2198: 5'-TAAACTTCAGGGTGACCAAAAAA TCA-3') (*Folmer, Black, Wr, & Vrijenhoek, 1994*). PCR reaction was performed in a total volume of 50 μl containing 50-100 ng DNA template, 5 μL of 10× PCR buffer, 1.5 mmol·L^{-1}of $MgCl_2$, 0.2 mmol·L^{-1}of each dNTP , 2 unit(U) of Taq polymerase, 0.2 μmol·L^{-1}of each primer. Thermal cycling began with one cycle pre-denaturation at 94 ^{o}C for 5 min, 35 cycles of denaturation at 94@C for 30 s, annealing at 50@C for 45 s, and extension at 72@C for 45 s, and a final extension holding at 72@C for 7 min, respectively (*Feng, Sun, Cheng, & Li, 2008*). The PCR product was separated by electrophoresis on 1.0% agarose gels.

Primer synthesis and DNA sequencings were conducted in commercial companies. Total 222 specimens were sequenced in one direction (63.51%) and others were in both directions. Except for the sequences obtained of the genomic DNA in this study, the other *COI* sequences were obtained from GeneBank to comparative and analyses.

Sequencing analysis

The chromatogram inspect, align and calibrate of the original sequence obtained by sequencing were using the SeqMan of DNASTAR Lasergene software package (DNASTAR, Inc., Madison, Wisconsin, USA). BioEdit 7.0.9 was used to alignment and sheared sequences.

Integrative taxonomy

In the species delimitation of OTU, traditional morphological identification, phylogenetic analysis, barcoding gap analysis and other different methods were used for comprehensive analysis and identification. The detailed analysis is as follows:

Distance- based approaches

The Kimura two-parameter (K2P) and p-distance model are used to calculate the pairwise genetic distance and construct the neighbor-joining tree (NJ) on the MEGA7.0. The haplotype diversity and nucleotide diversity of *COI* sequences were calculated using DnaSP. Then ML tree analysis was implemented using RaxmlGUI. All analysis used the default parameters and 1,000 bootstraps. In all trees, bootstrap values below 70% are not shown (*Shen et al., 2016*).

According to the phylogenetic trees of the original sequencing sequences and morphological characteristics, no less than 3 DNA sequences of each category were selected for further species confirmation by the IDENTIFICATION of BOLD (Barcode of Life Data System) and the BLAST of NCBI (National Biotechnology Information Center), to evaluate the accuracy of the morphological identification results and obtain reference

sequences with high relative similarity. In the selection of similarity sequences, we have defined 97% as a relatively loose standard to indicate potential species identification (*E. H. K. Wong & Hanner, 2008*).

In this study, a total 42 COI sequence with high similarity is obtained by aligning from GenBank. *Gammarus pisinnus* (GenBank accession number: KF824592) selected as the outer group. All DNA sequences obtained in this study were submitted to GenBank and BOLD; their accession numbers are provided in the electronic appendix (Table S2, Table S3), which can be found on the website of the National Center for Biotechnology Information (NCBI).

In addition to, Automatic Barcode Gap Discovery (ABGD) analysis was implemented on the website (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html), using K80, relative gap width (X = 1.0) and set the rest of the parameters as default values.

Tree- based approach

A large number of tests show that PTP is superior to GMYC on simulated data and its results are comparable to GMYC on real data sets meanwhile PTP requires less for data sets and requires only a simple phylogenetic tree (*Zhang, Kapli, Pavlidis, & Stamatakis, 2013*). Therefore, in this study, we chose PTP analysis to assist in species definition. PTP can delimit species based on the Phylogenetic Species Concept. So the entities output by PTP are in theory species. Bayesian Poisson tree process (bPTP) analyses were run on the web server (https://species.h-its.org/ptp) with 100,000 MCMC generations, a burn- in of 0.1 and other parameters as default. The ML tree generated by RAXML is selected as the input data(*Stamatakis, 2006*).

RESULTS

Morphological identification

According to the morphological characters obtained by the in-situ visual examination and stereomicroscope microscopic examination, the 1200 samples collected in Henan province included 9 species from 2 orders, 4 families and 6 genera(Some of the samples could not be identified morphologically due to damage or missing of distinguishing morphological characters). Among them, the individual morphological variation of*N.denticulate* , *N.davidi* (Bouvier, 1904),*M.maculatum* and *M.nipponense* is obvious, especially the morphology of their frontal angles (FIG.S1). The morphological variation of them may be caused by its widely distributed and geographical separation. In addition, consistent with the general distribution trend of freshwater shrimp, freshwater shrimp in this province are mainly composed of *Macrobrachium* and *Neocaridina* , among which*M.nipponense* is the dominant, followed by *N.denticulate*and *N.davidi* .

Molecular delimitation

Database search

In general, our morphological identification results matched the BLASTN annotations of the NCBI and BOLD databases, with at least 97% identities. According to the identification results of Species Level Barcode Records of the BOLD reference sequence library, the*P.clarkii* , *N.denticulate* , *N.davidi* ,*C.gracilipes* , *M.maculatum* and*M.nipponense*have been identified to the species level. The identification results were relatively reliable, which can basically be judged as effective identification species. The identities of *Macrobrachium sp. 'qilianensis'* and *P.modestus* are all greater than 98%, but the search results show that the sequence identities between the*P.modestus* and the three unpublished*M.sp.'qilianensis'*in the library is also high sequence identities (even sometimes they has the highest identities). At the same time, in the retrieval of*M.sp. 'qilianensis'* , the identities of itself and two unpublished*P.modestus* was also relatively high. After verification, the above*M.sp.'qilianensis'* (Accession: FJ958200, FJ958201) was sourced from GenBank which was direct submission and unpublished by Cheng (2009)(*Chen, Tsai, & Tzeng, 2009*). However, there was no corresponding morphological description and species identification of *M.sp.'qilianensis'* was found, so the *M.sp.'qilianensis'* may be a controversial species. In addition, the search results of the *G.pisinnus* are only 96%-98%, the identities is slightly lower; the search results of*P.sinensis* show no corresponding records. The search results of Species Level Barcode Records are similar to the search results of Species Level Barcode

Records, which will not be repeated here; the NCBI search shows a trend similar to the BOLD result, while the same species sequence has been retrieved from *P.sinensis* (MK994929, MK994930).

Species delimitation

The haplotype analysis result show that all the 222 sequences obtained was divided into 95 haplotypes, of which the widely distributed species such as *M.nipponensis* has greater genetic differentiation (FIG.S2). The NJ tree based on the p-distance model was not shown because of the same topology and similar bootstrap values to those of the K2P model (FIG.2). The NJ tree reveals the phylogenetic clusters formed by the specimens and taxonomic assignment based on morphology. All of the sequences were divided into 11 groups, among which the*M.maculatum* , *N.davidi* and *P.modestus* were further subdivided. By computation, the maximum K2P distances of all species were less than 2%, with the exception of*P.modestus* ,*M.maculatum* and *M.nipponense* which were 2.6%, 2.3% and 2.0% respectively. The maximum p-distances of all species were also less than 2%, with the exception of *P.modestus* , and*M.maculatum* , which were 2.5% and 2.3%, respectively. Mean K2P distances and p-distances within species were 1.6% and1.5%. The results show that both the K2P and the p-distances have similar results in genetic distance and phylogenetic analysis. In addition, the result of ABGD analysis and bPTP analysis shown that, the 222 original sequences were identified as 8 and 9-17 OTUs respectively. However, the 222 sequences were divided into 12 and 8-12 OTUs respectively when the downloaded reference sequences were introduced (FIG.3). All the above analyses supported that *P.modestus* and *M. sp.'qilianensis'*were a distinct monophyletic clade, and ABGD identified them as one species. Due to the failure to identify enough morphological differentiation on the *M.sp. 'qilianensis'* , and there are not sufficient description and evidence in the relevant references and original literature, we tentatively infer that *M. sp.'qilianensis'* may be an invalid species and it may be the synonym of the *P.modestus* . In addition, there is only one sample and the morphology of them are damaged so that it cannot be effectively identified by morphological identification. So, the existence of*N.ikiensis* and *N.palmata* need to be further collected and confirmed.

The comprehensive results of morphological characteristics and molecular delimitation indicate that there are at least 9 freshwater shrimps in Henan Province, among which the species of the *Palaemon* ,*Macrobrachium* and *Neocaridina* have a certain degree of morphological and genetic variation, the *Macrobrachium* was most obvious in them. Specifically, the differentiation of *P.modestus* ,*M.maculatum* and *M.nipponense* is more obvious. Because of the changeable morphology and blurred boundaries, they are only identified to the species level, and further research is needed to refine intraspecies classification. It is possible that these species may be in the middle of the differentiation process that from a species to a new species or there are some cryptic species which have not been discovered. In short, more evidence and studies are needed to determine whether the intermediate process of a species' differentiation to a new species is sufficient to form a new species, and is there a cryptic species.

Discussion

Barcoding success

It is well known that taxonomic identification and species identification of organisms are the most fundamental and important task of all biological research (*Luo et al., 2018*). The early classification identification is mainly based on the detailed morphological characteristics observation and anatomical structure verification by professional taxonomists. However, it takes a lot of time, energy, and has very high requirements for researchers and experimental specimens (*Carvalho et al., 2011* ; *P. D. Hebert, Cywinska, et al., 2003* ; *Shen et al., 2016*). In addition, there is always the demise of existing species and the emergence of new species in the biological world. With the rapid development of science and technology, more and more new species have been discovered one after another, and specialists in alpha taxonomy are not enough to carry out extensive and complex morphological identification (*Coleman, 2015*). Therefore, it is not only prone to misjudgment but also not conducive to widespread implementation. With the development of modern technology and the arrival of the molecular era, traditional morphological identification is gradually replaced by molecular biological identification.

Compared with traditional morphological identification, barcode technology has many advantages and applicability, for example: First, DNA is more stable than morphological characteristics because DNA characters are constant throughout development. However, morphological characteristics vary with age, developmental stage, environment and objective factors. For example, molecular identification of deformed and underdeveloped shrimp larvae has absolute advantages over morphological identification (*Burghart et al., 2014* ; *Lee & Kim, 2014*); Secondly, it can obtain sample DNA through some small parts of tissues, secretions, and even its living environment (*Pont et al., 2018*) of the biological body, which reduces the requirements of the sample (*Chang, Lin, Ren, Lin, & Shao, 2016*); More importantly, it is very friendly to operators and researchers. It is easy to operate, fast and efficient. It can be identified in batches, and requires less professional knowledge (*Takahara, Minamoto, & Doi, 2013* ;*Tinacci et al., 2018*). To be specific, the aims of DNA barcoding are identification unkonwn specimens with DNA barcodes of a priori defined taxonomic entities in databases (*Merckelbach & Borges, 2020*). It is being increasingly utilized to tackle many issues including illegal species exploitation, food fraud and biological invasions, and for biodiversity monitoring(*Bohmann et al., 2014* ; *Collins, Armstrong, Holyoake, & Keeling, 2013* ;*Frantine-Silva, Sofia, Orsi, & Almeida, 2015* ; *Goncalves, Oliveira-Marques, Matsumoto, & Miyaki, 2015* ;*Hubert, Espiau, Meyer, & Planes, 2015* ; *Khaksar et al., 2015* ;*Xiong et al., 2016*). Of course, as mentioned above, barcode technology also has some drawbacks, but these deficiencies are gradually being improved and overcome. Since the first use of cytochrome C oxidase subunit I(COI) for species identification, it has been shown that this gene fragment can be used as a "DNA barcoding" for biological authentication in many invertebrates species (*Barrett & Hebert, 2005* ;*Clare, Lim, Engstrom, Eger, & Hebert, 2007* ; *P. D. Hebert, Cywinska, et al., 2003* ; *Hendrich et al., 2015*). The research of Costa (2007), MAR(2018) et al. further proof that barcode technology is efficient and accurate in species identification of freshwater shrimp(*Costa et al., 2007* ;*Mar, Kang, Mao, & Wang, 2018*). In our results, 100% of the delimitation of species within freshwater shrimps was in accordance with genetic and morphological definitions. It demonstrating that DNA barcoding can help identify new taxa in complex groups and it can be widely used in the identification of unknown species and cryptic species (*Iyiola et al., 2018*).**Phylogenetic relationship**

Phylogeny refers to the formation and development process of a certain taxa, which is usually represented by evolutionary trees. In most cases, the topological structure and monophyletic branches of evolutionary trees can intuitively reflect the evolutionary and taxonomic status of each taxa. Thus, a taxon's taxonomic status may vary due to objective or human factors, but its phylogeny is stable. The taxonomic status of freshwater shrimp has been established as early as the 19th century, but with the development of taxonomy for many years, its taxonomic status has been continuously hit and challenged. As far as Caridea is concerned, Bracken (2009) and Chan (2010) et al. believe that there is a certain controversy in the classification of the current super family (*Bracken, De Grave, & Felder, 2009* ; *Chan, Lei, Li, & Chu, 2010*). The systematic placement of the infraorder Caridea within the decapods has been resolved several years ago, but the past studies have relied mainly on morphological features, leading to conflicting patterns of systematic development (*Bracken et al., 2009*). In recent years, studies have reexamined the phylogenetic relationship based on molecular characteristics, and the phylogenetic relationship tends to be stable. Some of these studies also shed new light on the phylogenetic relationships of Caridea (*Davis, De Grave, Delmer, & Wills, 2018* ; *C. P. Li, De Grave, Chan, Lei, & Chu, 2011*). The infraorder Caridea are a highly diverse group which had been described more than 3,500 species(*Grave & Fransen, 2011*), with a significant number of additions since then. Among them,approximately amounting to nearly a quarter live in freshwater of global caridean(*De Grave et al., 2015*). Due to the lack of high-level cladistics and genetic studies, there is little describe about the phylogeny of freshwater shrimp Except for Pereira (1997) (*De Grave, Cai, & Anker, 2008* ; *Pereira, 1997*).

As one of the most special superfamily in the Caribbean suborder and several taxonomic schemes have been proposed and modified since Palaemonoidea inception. However, its classification is still controversial (*Kou, Li, Chan, Chu, & Gan, 2013*). As most of Palaemoninae are highly conserved in morphological characteristics (*Short, 2004* ;*Walker & Poore, 2003*), the previous studies mostly focused on morphological characteristics, so there were big problems in the classification of genus. With the study of Ashelby (2012)

and Kou (2013) et al. for phylogenetic relationship based on molecular characteristics, the phylogenetic and evolutionary relationship of Palaemoninae has been preliminarily determined(*Ashelby, Page, De Grave, Hughes, & Johnson, 2012* ; *Kou et al., 2013*).. In all of their analyses, *Palaemonetes* ,*Exopalaemon* , *Coutierella* and certain *Palaemon*recover as a strongly supported single monophyletic clade, now they've been reclassified into the *Palaemon* (Ashelby et al., 2012) and be widely accepted (*Ashelby et al., 2012*). Our research also fully proves it. The *Macrobrachium* of Palaemonoidea have been intensively studied because of their rich diversity and wide distribution, it has been the subject of many molecular phylogenetic studies. The results all showed that there was a clear separation of the genetic differentiation of the species within the genus and *Macrobrachium* from other species of the related genera (*Chen et al., 2009* ;*Nicholas P. Murphy & Austin, 2003* ; *N. P. Murphy & Austin, 2004* ; *Nicholas P. Murphy & Austin, 2005* ; *Pileggi & Mantelatto, 2010* ; *Wowor et al., 2009*). In this study, the species of the *Macrobrachium* also showed greater genetic differentiation. In contrast, only a handful of phylogenetic studies have been carried out in the *Neocaridina* of Atyoidea, which is highly abundant and widely distributed. As the widely distributed species north of the Yangtze River, the taxonomy of*N.davidi* has always partially conflicting and unclear (*Klotz, Miesen, Hullen, & Herder, 2013*). With all these taxonomic and nomenclatural uncertainties of it (*Cai, 1996* ;*Liang, 2004* ;*Shih & Cai, 2007*), as the senior synonym the *C. davidi* Bouvier, 1904 has clear priority over*N. heteropoda* Liang, 2002 (article 23 of the ICZN). Thus, hereby we follow Klotz (2013) and retain the name *Neocaridina davidi*(*Klotz et al., 2013*).

Cryptic species

The investigation results show that there are at least 9 species of freshwater shrimp in Henan province. Compared with the existing studies (*F. Wang, 1989*), we have not been collected the *Macrobrachium superbum* , *Macrobrachium asperulum* and *Macrobrachium iusulare* . Although the existing sampling sites completely cover the original sampling sites and we have carried out targeted sampling on some samples. Unfortunately, these species has not yet been found. This may be because they are mostly distributed in the provinces to the south of the Yangtze River, and less distributed in the north, so it is difficult to collect (*X. Li et al., 2007*); or they may have migrated or even become extinct in the province; of course, it may also be because they are similar to *M.maculatum* and*M.nipponense* what has led to the misjudgments in morphological identification. In short, it needs more strong evidence to prove whether they exist in the province.

In this study, a variety of molecular boundary analyses all supported that some species of the *Neocaridina* , *Macrobrachium* and*Palaemom* which are widely distributed and rich in species (*De Grave et al., 2008*), have large genetic distance and more than one OTUs. Although both molecular and morphological characteristics show significant genetic differentiation and morphological differences between some species of these genera, there is no definitive criterion for whether these indicators are sufficient to indicate the emergence of a new species or the existence of an underlying species. In general, phylogenetic tree, genetic distance, haplotype analysis and PTP analysis indicated that the genetic differentiation of *M.maculatum* , *M.nipponense* and*P.modestus* were obvious, all of them had high haplotype diversity and no less than 2% intra-specific genetic distance although each of these groups formed a separate cluster or monophyletic clade. At the same time, most molecular delimitation analyses showed a higher number of species than morphological identification. This suggests that there are likely to be cryptic species within these species that have yet to be identified and described thus potentially explaining the high intraspecific diversity, even if they are not sufficiently differentiated to support the formation of a single new species.

Cryptic species are to a large extent intermediate products or even final products of new speciation. In the process of speciation, the boundaries of new species become more and more obvious over time. However, before the completion of this process (known as gray zone sense), the boundaries between species are often fuzzy and difficult to recognize, which makes the boundaries and identification of new species more subjective and dependent on the concept of species (*De Queiroz, 2007*). It is shown that the ability of DNA barcodes to identify recent speciation and no fully differentiated species is limited. The discovery of new species is usually accompanied by the delineation of molecular species, sometimes referred to as molecular operational

taxonomic units (MOTUs) (*Blaxter et al., 2005*). However, many newly discovered species are undescribed, even when the species hypothesis and species delimitation are highly supported by substantial evidence (*Pante, Schoelinck, & Puillandre, 2015*), which hinders taxonomic progress, identification of species and estimation of biodiversity (*Schlick-Steiner et al., 2007*). If a species is marked as merely presumed rather than formally described and therefore fully established, the taxonomy is still incomplete. In many cases, the transition from species delimitation to species description is a major task to be accomplished (*Jorger & Schrodl, 2013* ;*Merckelbach & Borges, 2020*).

Acknowledgements

We are grateful to the anonymous reviewers for their constructive comments. We thank for Xi Wang, Hui-hui Wu, Ru-yao Liu, and Xue-meng Yan for assistance with field work and their help.

In addition, this work was supported by the following funding: the National Natural Science Foundation of China (31872199, U2004146), this study support was provided by The High Performance Computing Center of Henan Normal University. Thanks also go to these funds and agencies for their support.

Conflict of interest

All authors declare that they have no competing interests.

Author contributions

GN and CZ designed this research. GN, CY, XM, CZ and YT performed sampling. MF conducted all relevant experiments, data analysis, wrote the manuscript. All authors read and approved the final manuscript.

Data availability statement

DNA sequences have been deposited in GenBank under Accession numbers MW069488–MW069709. Details regarding individual samples are available in Table S2.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Reference

Ashelby, C. W., Page, T. J., De Grave, S., Hughes, J. M., & Johnson, M. L. (2012). Regional scale speciation reveals multiple invasions of freshwater in Palaemoninae (Decapoda). *Zoologica Scripta, 41* (3), 293-306. doi:10.1111/j.1463-6409.2012.00535.x

Barrett, R. D. H., & Hebert, P. D. N. (2005). Identifying spiders through DNA barcodes. *Canadian Journal of Zoology, 83* (3), 481-491. doi:10.1139/z05-024

Birch, J. L., Walsh, N. G., Cantrill, D. J., Holmes, G. D., & Murphy, D. J. (2017). Testing efficacy of distance and tree-based methods for DNA barcoding of grasses (Poaceae tribe Poeae) in Australia. *PLoS One, 12* (10), e0186259. doi:10.1371/journal.pone.0186259

Blaxter, M., Mann, J., Chapman, T., Thomas, F., Whitton, C., Floyd, R., & Abebe, E. (2005). Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B-Biological Sciences, 360* (1462), 1935-1943. doi:10.1098/rstb.2005.1725

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., . . . de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution, 29* (6), 358-367. doi:10.1016/j.tree.2014.04.003

Bracken, H. D., De Grave, S., & Felder, D. L. (2009). Phylogeny of the infraorder Caridea based on mitochondrial and nuclear genes (Crustacea: Decapoda). *In J. W. Martin, K. A. Crandall D. L. Felder (Eds) Decapod Crustacean Phylogenetics, 18* , 281-305. doi:10.1201/9781420092592-c14

Burghart, S. E., Van Woudenberg, L., Daniels, C. A., Meyers, S. D., Peebles, E. B., & Breitbart, M. (2014). Disparity between planktonic fish egg and larval communities as indicated by DNA barcoding. *Marine Ecology Progress Series, 503* , 195-204. doi:10.3354/meps10752

Cai, Y. (1996). A revision of the genus Neocaridina (Crustacea: Decapoda: Atyidae). *Acta Zootaxonomica Sinica, 21* , 129-160.

Carvalho, D. C., Neto, D. A., Brasil, B. S., & Oliveira, D. A. (2011). DNA barcoding unveils a high rate of mislabeling in a commercial freshwater catfish from Brazil.*Mitochondrial DNA, 22 Suppl 1* , 97-105. doi:10.3109/19401736.2011.588219

Chan, T. Y., Lei, H. C., Li, C. P., & Chu, K. H. (2010). Phylogenetic analysis using rDNA reveals polyphyly of Oplophoridae (Decapoda : Caridea). *Invertebrate Systematics, 24* (2), 172-181. doi:10.1071/Is09049

Chang, C. H., Lin, H. Y., Ren, Q., Lin, Y. S., & Shao, K. T. (2016). DNA barcode identification of fish products in Taiwan: Government-commissioned authentication cases.*Food Control, 66* , 38-43. doi:10.1016/j.foodcont.2016.01.034

Chen, R. T., Tsai, C. F., & Tzeng, W. N. (2009). Freshwater Prawns (Macrobrachium Bate, 1868) of Taiwan with Special References to Their Biogeographical Origins and Dispersion Routes. *Journal of Crustacean Biology, 29* (2), 232-244. doi:10.1651/08-3072.1

Clare, E. L., Lim, B. K., Engstrom, M. D., Eger, J. L., & Hebert, P. D. N. (2007). DNA barcoding of Neotropical bats: species identification and discovery within Guyana.*Molecular Ecology Notes, 7* (2), 184-190. doi:10.1111/j.1471-8286.2006.01657.x

Coleman, C. O. (2015). Taxonomy in Times of the Taxonomic Impediment - Examples from the Community of Experts on Amphipod Crustaceans. *Journal of Crustacean Biology, 35* (6), 729-740. doi:10.1163/1937240x-00002381

Collins, R. A., Armstrong, K. F., Holyoake, A. J., & Keeling, S. (2013). Something in the water: biosecurity monitoring of ornamental fish imports using environmental DNA. *Biological Invasions, 15* (6), 1209-1215. doi:10.1007/s10530-012-0376-9

Costa, F. O., deWaard, J. R., Boutillier, J., Ratnasingham, S., Dooh, R. T., Hajibabaei, M., & Hebert, P. D. N. (2007). Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences, 64* (2), 272-295. doi:10.1139/F07-008

Davis, K. E., De Grave, S., Delmer, C., & Wills, M. A. (2018). Freshwater transitions and symbioses shaped the evolution and extant diversity of caridean shrimps. *Commun Biol, 1* , 16. doi:10.1038/s42003-018-0018-6

De Grave, S., Cai, Y., & Anker, A. (2008). Global diversity of shrimps (Crustacea : Decapoda : Caridea) in freshwater. *Hydrobiologia, 595* , 287-293. doi:10.1007/s10750-007-9024-2

De Grave, S., Li, C. P., Tsang, L. M., Chu, K. H., & Chan, T. Y. (2014). Unweaving hippolytoid systematics (Crustacea, Decapoda, Hippolytidae): resurrection of several families.*Zoologica Scripta, 43* (5), 496-507. doi:10.1111/zsc.12067

De Grave, S., Smith, K. G., Adeler, N. A., Allen, D. J., Alvarez, F., Anker, A., . . . Wowor, D. (2015). Dead Shrimp Blues: A Global Assessment of Extinction Risk in Freshwater Shrimps (Crustacea: Decapoda: Caridea). *PLoS One, 10* (3). doi:10.1371/journal.pone.0120198

De Queiroz, K. (2007). Species concepts and species delimitation. *Syst Biol, 56* (6), 879-886. doi:10.1080/10635150701701083

Feng, J., Sun, Y., Cheng, X., & Li, J. (2008). Sequence analysis of mitochondrial COI gene of Macrobrachium nipponense from the five largest freshwater lakes in china.*Journal of Fisheries of China* (04), 23-31.

Folmer, O., Black, M., Wr, H., Lutz, R. , & Vrijenhoek, R. (1994). Dna primers for amplification of mitochondrial cytochrome c oxidase subunit i from diverse metazoan invertebrates. . . *Molecular marine biology and biotechnology, 3* (5), 294-299.

Frantine-Silva, W., Sofia, S. H., Orsi, M. L., & Almeida, F. S. (2015). DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring.*Molecular Ecology Resources, 15* (5), 1226-1237. doi:10.1111/1755-0998.12385

Goncalves, P. F. M., Oliveira-Marques, A. R., Matsumoto, T. E., & Miyaki, C. Y. (2015). DNA Barcoding Identifies Illegal Parrot Trade. *Journal of Heredity, 106* , 560-564. doi:10.1093/jhered/esv035

Grave, S. D., & Fransen, C. H. J. M. (2011). Carideorum catalogus: the recent species of the dendrobranchiate, stenopodidean, procarididean and caridean shrimps (crustacea: decapoda). *Zoologische Mededelingen Lden, 89* , 195-589.

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc Biol Sci, 270* (1512), 313-321. doi:10.1098/rspb.2002.2218

Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci, 270 Suppl 1* , S96-99. doi:10.1098/rsbl.2003.0025

Hebert, P. D. N., & Gregory, T. R. (2005). The promise of DNA barcoding for taxonomy. *Systematic Biology, 54* (5), 852-859. doi:10.1080/10635150500354886

Hendrich, L., Moriniere, J., Haszprunar, G., Hebert, P. D. N., Hausmann, A., Kohler, F., & Balke, M. (2015). A comprehensive DNA barcode database for Central European beetles with a focus on Germany: adding more than 3500 identified species to BOLD. *Molecular Ecology Resources, 15* (4), 795-818. doi:10.1111/1755-0998.12354

Holthuis, L. B. (1980). FAO Species Catalogue. Vol. 1—Shrimps and Prawns of the World. An Annotated Catalogue of Species of Interest to Fisheries. *125* , 1-217.

Hubert, N., Espiau, B., Meyer, C., & Planes, S. (2015). Identifying the ichthyoplankton of a coral reef using DNA barcodes. *Molecular Ecology Resources, 15* (1), 57-67. doi:10.1111/1755-0998.12293

Iyiola, O. A., Nneji, L. M., Mustapha, M. K., Nzeh, C. G., Oladipo, S. O., Nneji, I. C., . . . Adeola, A. C. (2018). DNA barcoding of economically important freshwater fish species from north-central Nigeria uncovers cryptic diversity.*Ecol Evol, 8* (14), 6932-6951. doi:10.1002/ece3.4210

Jorger, K. M., & Schrodl, M. (2013). How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front Zool, 10* (1), 59. doi:10.1186/1742-9994-10-59

Khaksar, R., Carlson, T., Schaffner, D. W., Ghorashi, M., Best, D., Jandhyala, S., . . . Amini, S. (2015). Unmasking seafood mislabeling in US markets: DNA barcoding as a unique technology for food authentication and quality control. *Food Control, 56* , 71-76. doi:10.1016/j.foodcont.2015.03.007

Klotz, W., Miesen, F. W., Hullen, S., & Herder, F. (2013). Two Asian fresh water shrimp species found in a thermally polluted stream system in North Rhine-Westphalia, Germany.*Aquatic Invasions, 8* (3), 333-339. doi:10.3391/ai.2013.8.3.09

Kou, Q., Li, X., Chan, T.-Y., Chu, K. H., & Gan, Z. (2013). Molecular phylogeny of the superfamily Palaemonoidea (Crustacea : Decapoda : Caridea) based on mitochondrial and nuclear DNA reveals discrepancies with the current classification.*Invertebrate Systematics, 27* (5). doi:10.1071/is13005

Lee, S. J., & Kim, J. K. (2014). Identification of Trichiurus (Pisces: Trichiuridae) Eggs and Larvae from Korea, with a Taxonomic Note. *Fisheries and Aquatic Sciences, 17* (1), 137-143. doi:10.5657/FAS.2014.0137

Li, C. P., De Grave, S., Chan, T. Y., Lei, H. C., & Chu, K. H. (2011). Molecular systematics of caridean shrimps based on five nuclear genes: Implications for superfamily classification. *Zoologischer Anzeiger, 250* (4), 270-279. doi:10.1016/j.jcz.2011.04.003

Li, X., Liu, R., Liang, X., & Chen, G. (2007). Fauna Sinica, Invertebrata Vol. 44, Crustacea, Decapoda, Palaemonoidea. *Beijing (China): Science Press.*

Liang, X. (2004). Fauna Sinica. Invertebrata Vol. 36. Crustacea, Decapoda, Atyidae. Beijing (China):Science Press.

Liu, R. (1955). The Economic Shrimps in Northern of China. Beijing (China):Beijing: Science Press.

Luo, A., Ling, C., Ho, S. Y. W., & Zhu, C. D. (2018). Comparison of Methods for Molecular Species Delimitation Across a Range of Speciation Scenarios. *Syst Biol, 67* (5), 830-846. doi:10.1093/sysbio/syy011

Macher, J., Macher, T., & Leese, F. (2017). Combining NCBI and BOLD databases for OTU assignment in metabarcoding and metagenomic datasets: The BOLD_NCBI _Merger.*Metabarcoding and Metagenomics 1: e22262.*doi:10.3897/mbmg.1.22262

Mar, W. I. N., Kang, P.-F., Mao, B. I. N., & Wang, Y.-F. (2018). Morphological and molecular features of some freshwater prawn species under genus Macrobrachium Spence Bate, 1868 (Crustacea: Decapoda: Palaemonidae) from Myanmar. *Zootaxa, 4388* (1). doi:10.11646/zootaxa.4388.1.9

Merckelbach, L. M., & Borges, L. M. S. (2020). Make every species count: FastaChar software for rapid determination of molecular diagnostic characters to describe species.*Mol Ecol Resour* . doi:10.1111/1755-0998.13222

Murphy, N. P., & Austin, C. M. (2003). Molecular Taxonomy and Phylogenetics of Some Species of Australian Palaemonid Shrimps. *Journal of Crustacean Biology, 23* (1), 169-177. doi:10.1651/0278-0372(2003)023[0169:Mtapos]2.0.Co;2

Murphy, N. P., & Austin, C. M. (2004). Multiple origins of the endemic Australian Macrobrachium (Decapoda : Palaemonidae) based on 16S rRNA mitochondrial sequences.*Australian Journal of Zoology, 52* (5), 549. doi:10.1071/zo03062

Murphy, N. P., & Austin, C. M. (2005). Phylogenetic relationships of the globally distributed freshwater prawn genus Macrobrachium (Crustacea: Decapoda: Palaemonidae): biogeography, taxonomy and the convergent evolution of abbreviated larval development. *Zoologica Scripta, 34* (2), 187-197. doi:10.1111/j.1463-6409.2005.00185.x

New, M. B., & Nair, C. M. (2012). Global scale of freshwater prawn farming. *Aquaculture Research, 43* (7), 960-969. doi:10.1111/j.1365-2109.2011.03008.x

Pante, E., Schoelinck, C., & Puillandre, N. (2015). From Integrative Taxonomy to Species Description: One Step Beyond. *Systematic Biology, 64* (1), 152-160. doi:10.1093/sysbio/syu083

Pereira, G. (1997). A cladistic analysis of the freshwater shrimps of the family Palaemonidae (Crustacea, Decapoda, Caridea). *Acta Biologica Venezuelica 17(Suppl.)* , 1-69.

Pileggi, L. G., & Mantelatto, F. L. (2010). Molecular phylogeny of the freshwater prawn genus Macrobrachium (Decapoda, Palaemonidae), with emphasis on the relationships among selected American species. *Invertebrate Systematics, 24* (2), 194-208. doi:10.1071/Is09043

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., . . . Dejean, T. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports, 8* . doi: 10.1038/s41598-018-28424-8

Sambrock, J., & Russel, D. (2001). Molecular cloning: a laboratory manual, 3rd ed. . *Immunology, 49* (1), 895-909. doi:10.2307/1309366

Schlick-Steiner, B. C., Seifert, B., Stauffer, C., Christian, E., Crozier, R. H., & Steiner, F. M. (2007). Without morphology, cryptic species stay in taxonomic crypsis following discovery. *Trends in Ecology & Evolution, 22* (8), 391-392. doi:10.1016/j.tree.2007.05.004

Shen, Y., Guan, L., Wang, D., & Gan, X. (2016). DNA barcoding and evaluation of geneticdiversity in Cyprinidae fish in the midstream of the Yangtze River. *Ecol. Evol., 6* (9), 2702-2713. doi:10.1002/ece3.2060

Shih, H. T., & Cai, Y. (2007). Two new species of the land-locked freshwater shrimps genus, Neocaridina Kubo, 1938 (Decapoda : Caridea : Atyidae), from Taiwan, with notes on speciation on the island. *Zoological Studies, 46* (6), 680-694.

Short, J. W. (2004). A Revision of Australian River Prawns, Macrobrachium (Crustacea: Decapoda: Palaemonidae). *Hydrobiologia, 525* (1-3), 1-100. doi:10.1023/b:hydr.0000038871.50730.95

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

Takahara, T., Minamoto, T., & Doi, H. (2013). Using Environmental DNA to Estimate the Distribution of an Invasive Fish Species in Ponds. *PLoS One, 8* (2). doi:10.1371/journal.pone.0056584

Tinacci, L., Guidi, A., Toto, A., Guardone, L., Giusti, A., D'Amico, P., & Armani, A. (2018). DNA barcoding for the verification of supplier's compliance in the seafood chain: How the lab can support companies in ensuring traceability.*Ital J Food Saf, 7* (2), 6894. doi:10.4081/ijfs.2018.6894

Walker, T. M., & Poore, G. C. B. (2003). Rediagnosis of Palaemon and differentiation of southern Australian species (Crustacea: Decapoda: Palaemonidae). *Memoirs of Museum Victoria 60* , 243-256. doi:10.24199/j.mmv.2003.60.25

Wang, F. (1989). A Preliminary study on shrimp from Henan Province. . *Henan Fisheries, 000* (001), 16-18.

Wang, Z., Guo, Y., Chen, R., He, X., Liu, C., & Liu, Y. (2009). COI barcode sequences of teleosts in the South China Sea. *Oceanologia et Limnologia Sinica, 40* (005), 308-614.

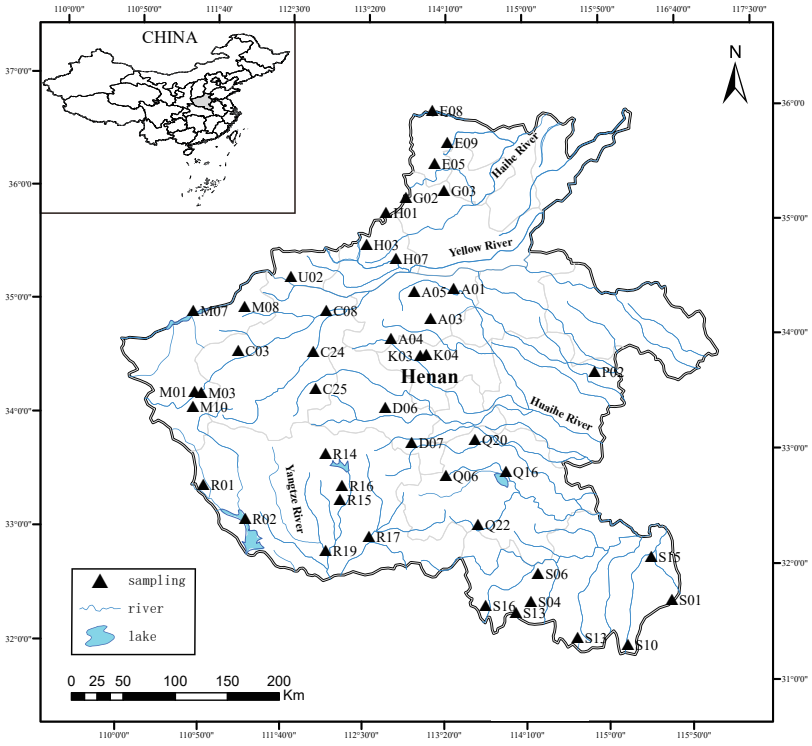
Wong, E. H. K., & Hanner, R. H. (2008). DNA barcoding detects market substitution in North American seafood. *Food Research International, 41* (8), 828-837. doi:10.1016/j.foodres.2008.07.005

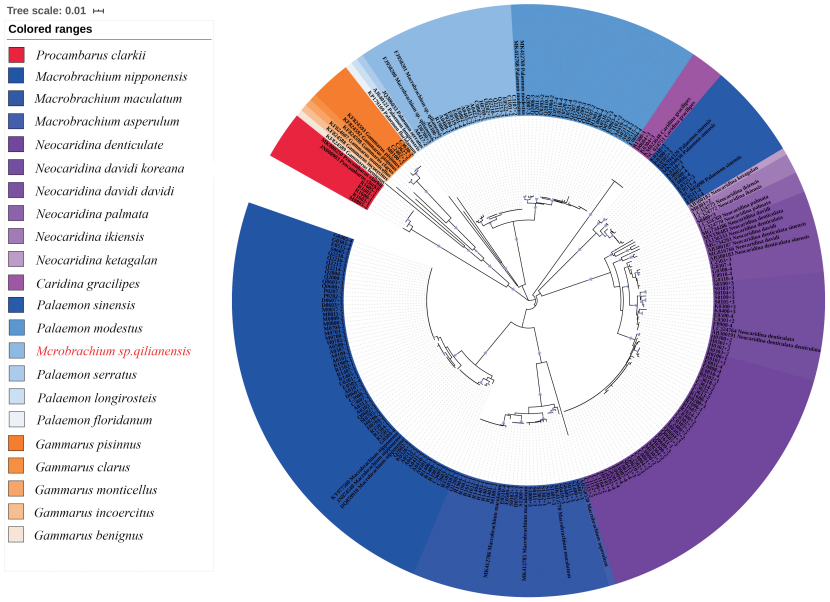
Wong, L. L., Peatman, E., Lu, J., Kucuktas, H., He, S., Zhou, C., Na-nakorn, U., Liu, Z. (2011). DNA barcoding of catfish: species authentication and phylogenetic assessment. *PLoS One, 6* (3), e17812. doi:10.1371/journal.pone.0017812

Wowor, D., Muthu, V., Meier, R., Balke, M., Cai, Y., & Ng, P. K. (2009). Evolution of life history traits in Asian freshwater prawns of the genus Macrobrachium (Crustacea: Decapoda: Palaemonidae) based on multilocus molecular phylogenetic analysis. *Mol Phylogenet Evol, 52* (2), 340-350. doi:10.1016/j.ympev.2009.01.002

Xiong, X., Guardone, L., Giusti, A., Castigliego, L., Gianfaldoni, D., Guidi, A., & Andrea, A. (2016). DNA barcoding reveals chaotic labeling and misrepresentation of cod (鳕, Xue) products sold on the Chinese market. *Food Control, 60* , 519-532. doi:10.1016/j.foodcont.2015.08.028

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





Hosted file

`Figure 3 Classification of freshwater shrimp with molecular methods.pdf` available at https://authorea.com/users/343092/articles/503007-species-diversity-of-freshwater-shrimp-in-henan-province-china-based-on-morphological-characters-and-coi-mitochondrial-gene