Exploring the species diversity of *B. weberi* and the phylogenetic relationship of *Barbronia* within Suborder Erpobdelliformes (Clitellata: Annelida) (#82323)

First submission

Guidance from your Editor

Please submit by **23 Mar 2023** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files
Download and review all files from the materials page.

5 Figure file(s)
3 Table file(s)
3 Raw data file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Exploring the species diversity of *B. weberi* and the phylogenetic relationship of *Barbronia* within Suborder Erpobdelliformes (Clitellata: Annelida)

Yingkui Liu [Corresp., 1, 2], **Xinxin Fu** [1], **Yu Wang** [1], **Jing Liu** [1], **Yong Liu** [1, 2], **Chong Li** [1, 2], **Jiajia Dong** [Corresp. 1, 2]

[1] Jiangsu Key Laboratory of Brain Disease Bioinformation, Research Center for Biochemistry & Molecular Biology, Xuzhou Medical University, Xuzhou, Jiangsu, People's Republic of China

[2] College of Life Science, Xuzhou Medical University, Xuzhou, Jiangsu, People's Republic of China

Corresponding Authors: Yingkui Liu, Jiajia Dong
Email address: yingkui.liu@outlook.com, jiajia.dong@xzhmu.edu.cn

Background. The freshwater *Barbronia* is ~~one kind~~ of macrophagous ~~leech~~ belonging to Erpobdelliformes (Salifidae: Clitellata: Annelida), and *B. weberi* is a well-known invasive leech with a worldwide distribution. However, the systematics of *Barbronia* has not yet been adequately resolved due to a few molecular ~~marks,~~ and only 20 *Barbronia* sequences were available in the GenBank database. This gap severely limits our understanding of the *Barbronia* species identification, as well as the phylogenetic placement of the genus *Barbronia* within Salifidae.

Methods. The species boundary of *Barbronia* species was estimated using bGMYC and bPTP methods based on all available *Barbronia* COI sequences, and the uncorrected COI p-distance was calculated in MEGA. A data matrix consisting of four loci (COI, 12S, 18S, and 28S rDNA) for 49 Erpobdellid~~s~~ leeches representing 8 genera within the Suborder Erpobdelliformes was aligned using MAFFT, and was used to reconstruct the phylogenetic relationship of *Barbronia* via Bayesian inference (BI) and the maximum likelihood (ML) method. In addition, simultaneously obtaining entire mitochondrial genome and completely 18S/28S rDNA from valuable specimens using NGS was attempted via GetOrganelle.

Results. Both bGMYC and bPTP results were generally congruent and suggested that the previously proposed taxa (*B. arcana*, *B. weberi* formosana, and *B. wuttkei* or *Erpobdella wuttkei*) were synonyms of *B. weberi* rather than valid species. The specimens listed in the *B. gwalagwalensis* group, however, were split into at least two Primary Species Hypotheses (PSHs). The p-distance of the first PSH was less than 1.3% but increased to 4.5% when including the secondary PSH (i.e., *B.* cf. *gwalagwalensis*). In comparison, the interspecific p-distance between *B. weberi* group and *B. gwalagwalensis* group ranged from 6.4% to 8.7%, and the intraspecific p-distance within *B. weberi* group was less than 0.8%. Considering the species delimitation results and the sufficient large p-distance, the specimen sampled in China was treated as *B.* cf. *gwalagwalensis*. Full lengths of the newly assembled mitochondrial genome, 18S and 28S rDNAs of *B.* cf. *gwalagwalensis*, were 14847 bp, 1876 bp, and 2863 bp, respectively. Phylogenetic relationships within suborder Erpobdelliformes were well constructed in ML and BI analysis, the monophyly of Salifidae, Orobdellidae, and Gastrostomobdellidae *sensu stricto* was all well supported. Within the Salifidae, a well-supported *Barbronia* was closely related to a clade containing *Odontobdella* and *Mimobdella*, and these three genera were sister to a clade consisted of *Salifa* and *Linta*. According to the results of this study, the strategy of simultaneous obtaining both whole mitochondria and nuclear markers from extensively sampled Salifids species using NGS was expected to fathom both the species diversity of *B. gwalagwalensis* and the evolutionary relationship of Salifidae.

Exploring the species diversity of *B. weberi* and the phylogenetic relationship of *Barbronia* within Suborder Erpobdelliformes (Clitellata: Annelida)

Yingkui Liu[1, 2], Xinxin Fu[1], Yu Wang[1], Jing Liu[1], Yong Liu[1, 2], Chong Li[1, 2], Jiajia Dong[1, 2]

[1] Jiangsu Key Laboratory of Brain Disease Bioinformation, Research Center for Biochemistry & Molecular Biology, Xuzhou Medical University, Xuzhou, Jiangsu, People's Republic of China

[2] College of Life Science, Xuzhou Medical University, Xuzhou, Jiangsu, People's Republic of China

Corresponding Author:
Jiajia Dong[1, 2]
209 Tongshan Road, Xuzhou, Jiangsu, 221004, People's Republic of China.
Email address: jiajia.dong@xzhmu.edu.cn
Yingkui Liu[1, 2]
209 Tongshan Road, Xuzhou, Jiangsu, 221004, People's Republic of China.
Email address: yingkui.liu@outlook.com

Abstract

Background. The freshwater *Barbronia* is ~~one kind~~ of macrophagous ~~leech~~ belonging to Erpobdelliformes (Salifidae: Clitellata: Annelida), and *B. weberi* is a well-known ~~invasive~~ leech with a worldwide distribution. However, the systematics of *Barbronia* has not yet been adequately ~~resolved~~ due to a few molecular marks, and only 20 *Barbronia* sequences were available in the GenBank database. This gap severely limits our understanding of the *Barbronia* species identification, as well as the phylogenetic placement of the genus *Barbronia* within Salifidae.

Methods. The species boundary of *Barbronia* species was estimated using bGMYC and bPTP methods based on all available *Barbronia* COI sequences, and the uncorrected COI p-distance was calculated in MEGA. A data matrix consisting of four loci (COI, 12S, 18S, and 28S rDNA) for 49 Erpobdellids leeches representing 8 genera within the Suborder Erpobdelliformes was aligned using MAFFT, and was used to reconstruct the phylogenetic relationship of *Barbronia* via Bayesian inference (BI) and the maximum likelihood (ML) method. In addition, simultaneously obtaining entire mitochondrial genome and completely 18S/28S rDNA from valuable specimens using NGS was attempted via GetOrganelle.

Results. Both bGMYC and bPTP results were generally congruent and suggested that the previously proposed taxa (*B. arcana*, *B. weberi* formosana, and *B. wuttkei* or *Erpobdella wuttkei*) were synonyms of *B. weberi* rather than valid species. The specimens listed in the *B. gwalagwalensis* group, however, were split into at least two Primary Species Hypotheses (PSHs). The p-distance of the first PSH was less than 1.3% but increased to 4.5% when including the secondary PSH (i.e., *B.* cf. *gwalagwalensis*). In comparison, the interspecific p-distance between *B. weberi* group and *B. gwalagwalensis* group ranged from 6.4% to 8.7%, and the intraspecific p-distance within *B. weberi* group was less than 0.8%. Considering the species delimitation results and the sufficient large p-distance, the specimen sampled in China was treated as *B.* cf. *gwalagwalensis*. Full lengths of the newly assembled mitochondrial genome, 18S and 28S rDNAs of *B.* cf. *gwalagwalensis*, were 14847 bp, 1876 bp, and 2863 bp, respectively. Phylogenetic relationships within suborder Erpobdelliformes were well constructed in ML and BI analysis, the monophyly of Salifidae, Orobdellidae, and Gastrostomobdellidae *sensu stricto* was all well supported. Within the Salifidae, a well-supported *Barbronia* was closely related to a clade containing *Odontobdella* and *Mimobdella*, and these three genera were sister to a clade consisted of *Salifa* and *Linta*. According to the results of this study, the strategy of simultaneous obtaining both whole mitochondria and nuclear markers from extensively sampled Salifids species using NGS was expected to fathom both the species diversity of *B. gwalagwalensis* and the evolutionary relationship of Salifidae.

Introduction

Barbronia ~~was~~ a genus of freshwater predaceous ~~leech~~ belonging to the family Salifidae (Erpobdelliformes, Clitellata, Annelida), only three species were previously reported in China,

Barbronia weberi, *Barbronia zhejiangica,* and *Barbronia yunnanensis* (Yang 1996; Yang et al. 1997). The former one ~~was~~ a well-known invasive species, however, the latter two were rarely known endemic *Barbronia* species. ~~*B.*~~ *weberi* was originally described by Blanchard (1897) in Indonesia, but now it was broadly found in China (Ta-Hsiang 1974; Zhao et al. 2020), India (Bandyopadhyay & Mandal 2005; Chandra & Mahajan 1971; Ghate 1991), Myanmar (Eriksen et al. 2022), United States (Rutter & Klemm 2001; Sawyer & Sawyer 2018), Brazil (Pamplin & Rocha 2000), Mexico (~~de Oca~~ et al. 2016; ~~Figueroa~~ et al. 2005), United Kingdom (Sawyer 1986; Sawyer & Sawyer 2018), Germany (Kutschera 2004; Nehring 2006), Hungary (Ludanyi et al. 2019), Italy (Genoni & Fazzone 2008), Spain (Pavluk et al. 2011), Netherlands (Van Haaren et al. 2004), Australia (Govedich et al. 2003; Govedich et al. 2002), New Zealand (Mason & Julia 1976), and South Africa (Nakano & Nguyen 2015). *B. weberi* has attained its current broad distribution, which was not only due to human activities (Govedich et al. 2003), but also because it can reproduce cocoons without cross-fertilization (Sawyer 2020). Moreover, *B. weberi* was the second intermediate host of some parasitic trematodes (Blasco-Costa et al. 2016).

The taxonomic status of this widely distributed invasive species *B. weberi* needed to be evaluated on a global scale (Oceguera-Figueroa et al. 2011), which involved *B. weberi formosana, B. wuttkei, B. arcana,* and *Barbronia* sp ~~(GenBank accession: AY786457)~~. Van Haaren et al. (2004) described one Dutch specimen as *B. assiuti*/*weberi*-complex, since neither species discrimination nor identification based on morphological characteristics of one specimen was impossible. Molecular identification provides a valuable complement to morphological taxonomy in the past two decades, a partial mitochondrial gene COI sequence has been used to recognize *Barbronia* species. *B. wuttkei* was originally ~~named~~ *Erpobdella wuttkei* based on morphological traits only (Kutschera 2004), but later it was ~~proposed~~ as a valid taxon by Grosser & Trontelj (2008) based on COI-based molecular phylogenetic evidence. Both *B. wuttkei* and *B. arcana*, however, now were lumped into the taxon *B. weberi* using COI sequence-based species delimitation approach Poisson Tree Process (PTP) (Klass et al. 2021). One *Barbronia sp* (GenBank accession: MN503261), as well as another specimen (GenBank accession: MF458701, marked as *Erpobdella* sp), were recently recognized as *B. gwalagwalensis* on basis of COI data (Klass et al. 2021).

The phylogeny of *Barbronia* was firstly ~~estimated~~ based on 18S ribosomal sequences only, *B. weberi* *was* closely related to a clade of two Erpobdellids ~~(~~*Dina lineata* and *Erpobdella octoculata*~~)~~ (Trontelj et al. 1999). Although one ~~partial~~ of ~~a~~ mitochondrial marker 12S rDNA was also tried to amplify from the *B. weberi* specimen in ~~that study~~, it failed using the primer pair 12S-A/12S-B (Siddall 2002b~~; Trontelj et al. 1999~~). The unique 18S rDNA sequence of *B. weberi* was thus used in the following phylogenetic study (Borda & Siddall 2004a; ~~F~~igueroa et al. 2005). The phylogenetic analysis of combined data of partial 18S rDNA, 28S rDNA, 12S rDNA ~~(amplified by modified primers)~~, and cytochrome c oxidase subunit I (COI) sequences ~~recover~~ *Barbronia* species (*B. weberi*, *B. weberi formosana, B. gwalagwalensis,* and one undescribed *Barbronia* species) as one monophyletic ~~clade,~~ which was closely related with *Linta be* ~~(Borda & Siddall 2004b)~~. Another phylogenetic result supported that *B. weberi* and *Barbronia arcana* were lumped into a Salifidae clade, which was sister group of Erpobdellidae (~~F~~igueroa et al. 2005). Six *Barbronia* specimens, represented an expanded dataset of all mentioned *Barbronia* species and *B. wuttkei*, formed a monophyletic group, which was sister to a clade of *Salifa* and *Linta* species (Oceguera-Figueroa et al. 2011).

The implementation of next-generation sequencing (NGS) technologies has led to a substantial increase in the number of whole mitochondrial genomes in the GenBank database (Kuang & Yu 2019), and has gradually become a mainstay of species identification and phylogeny research (Franco-Sierra & Diaz-Nieto 2020; Kortsinoglou et al. 2020).To date, however, only about 20 sequences of *Barbronia* species ~~were~~ available in the GenBank database. Eight of these sequences ~~were~~ the mitochondrial COI barcodes, six partial 18S rDNA sequences, four short 28S rDNA sequences, and only two 12S rDNA sequences, representatively. Neither one whole mitochondrial genome of *Barbronia* species nor even the family Salifidae has been published. This gap severely limits our understanding of either the species identification or phylogenetic relationships of *Barbronia* within Salifidae.

In ~~this~~ study, the whole mitochondrial genome, 18S rDNA, and 28S rDNA of *Barbronia* cf. *gwalagwalensis* from China was assembly and annotated; the species boundary of *Barbronia* species was evaluated by two kinds of single-locus based species delimitation methods; the

phylogenetic relationships of *Barbronia* within the Suborder Erpobdelliformes were conducted based on both mitochondrial and nuclear data.

Materials and methods

DNA extraction and assembly of the mitochondrial genome, and rDNAs

The total DNA material was extracted from a *B.* cf. *gwalagwalensis* specimen sampled from Xuzhou (Jiangsu province, China) using Ezup Column Animal Genomic DNA Purification Kit (Sangon Biotech Company, Shanghai, China) following the manufacturer's protocols. The quality of DNA was evaluated via gel electrophoresis and visualized on a ChemiDoc XRS+ (Bio-Rad®). The quantification of DNA was checked via the Spectrophotometer NanoDrop 2000. DNA was eluted in ddH_2O and stored at −20 °C until used for next-generation sequencing (NGS). A sequencing library was constructed using a Nextera XT DNA Library Preparation Kit (Illumina™), and sequencing reads were generated on a MiSeq System (Illumina™) to generate 150 bp paired-end reads. Low-quality bases in raw FastQ reads were trimmed using Trimmomatic v.0.36, the remaining reads were assembled as contig and used to reconstruct the mitochondrial genome via GetOrganelle V1.1.7 (Jin et al. 2020). The annotation of mitochondrial protein-coding genes, Transfer RNAs, and ribosomal DNAs of *B.* cf. *gwalagwalensis* was performed using MITOS2 Webserver, and was also manual check by comparing nucleotide sequences with those from the published available closely related leech species to refine these annotations when possible using Geneious 2021.

Single-locus based species delimitation

Species delimitation of *Barbronia* species should be assessed with different molecular species delimitation approaches using data of single-locus or multiple loci. Only 20 *Barbronia* sequences were available in GenBank, and half of these sequences were partial mitochondrial gene COI sequences (see Table 1). Two specimens (accession number: DQ009666 and MF458701~~labled~~) labeled as *Erpobdella* sp were also included, because both were used in previous studies of *Barbronia*.

Therefore, the species delimitation of *Barbronia* was conducted only on basis of a single-locus dataset of COI, using both a Bayesian implementation of the general mixed yule-coalescent model (bGMYC) and an updated version of the maximum likelihood Poisson Tree Processes model (bPTP) in the current study. The bGMYC can infer the transition between the Yule model (species level) and coalescent model (population level within species)(Fujisawa & Barraclough 2013), and posterior probabilities account for phylogenetic uncertainty(Reid & Carstens 2012). A phylogenetic gene tree was required in bPTP analysis, and putative species were estimated with Bayesian support under a simulation of Poisson Tree Processes (Zhang et al. 2013).

For species delimitation analyses in bGMYC and bPTP, the input COI phylogenetic tree was performed with BEAST Version 1.10.4, using HKY+I+G substitution model with two partitions (one partition comprised the 1st and 2nd codon positions of COI , the other one was the 3rd codon position) with a yule speciation prior and a strict clock. Ten million MCMC generations were performed, sampling every 1000 generations. After burn-in, a maximum clade credibility tree was built using TreeAnnotator using the maximum clade credibility method. The bGMYC analyses were implemented generally by flowing the proposal of Liu et al. (2017). The bGMYC analyses consisted of the total 50,000 MCMC generations, with a ~~thinning~~ every 100 generations, discarding the first 2000 generations as burn-in, and setting the upper and lower bounds on the threshold parameter from 1 (the minimum number of species) to 11(the maximum number of tips in COI tree). The maximum clade credibility tree bPTP analyses were performed with default setting via an online server (https://species.h-its.org/). In addition, GMYC models may be not able to delineate species properly in data sets composed of one or two species (Dellicour & Flot 2015), therefore, *Mimobedlla japonica* and *Odontobdella blancharidi* were added in the analyses of species delimitation.

Phylogenetic analysis using multiple loci

To corroborate previous overarching phylogenetic frameworks, 249 sequences of four molecular markers (COI, 12S rDNA, 18S rDNA, and 28S rDNA) representing 49 ~~E~~rpobdellids leeches were retrieved from the GenBank database (see Table 2). In addition, the genera *Dina* and *Mooreobdella* were formally synonymized under the genus *Erpobdella* (Siddall 2002a), thus the name of both *Dina* and *Mooreobdella* were instead with *Erpobdella* in the current study. COI

sequences were aligned using MAFFT 7.45 (Katoh & Standley 2014), and sequences of rDNA were aligned with LocARNA (Will et al. 2012). The individual alignments of four markers were concatenated in PhyloSuite (Zhang et al. 2020) as a matrix used for flowing analyses. Phylogeny analyses were performed both by Bayesian inference (BI) and the maximum likelihood (ML) method via MrBayes V3.2.7(Altekar et al. 2004) and IQ-TREE (Nguyen et al. 2015), separately. According to the parsimony results of Erpobdelliformes (Nakano & Nguyen 2015; Oceguera-Figueroa et al. 2011), *Erpobdella* species were lumped into the basal clade of Erpobdelliformes, thus *Erpobdella* species were set as outgroups. The best-fit evolutionary model of each molecular marker (the first, second, and third codon position sites of COI) was determined using PartitionFinder 2 (Lanfear et al. 2017) based on the AICc score. For Bayesian analysis, two independent runs, with four Markov Chain Monte Carlo (MCMC) chains each, were simultaneously carried out for 4 million generations and sampled every 10,000 generations. The analysis was assumed to have reached stationarity when the potential scale reduction factor value (PSRF) approached 1.0 and the Effective Sample Size value > 100. After discarding the 25% samples as burn-in, the 50% majority-rule consensus tree was built. For the ML analysis, the reliability of bootstrap values and tree topology was assessed by ultrafast bootstrap using 1000 replicates.

Results

The outcome of Single-locus-based species delimitation

The *B. weberi* groups and the *B. gwalagwalensis* group listed in the table 1 ~~were corresponded~~ to two main well-supported (PP > 0.99) clades in COI phylogenetic analysis (see, Figure. 1). All members of *B. weberi* group were consistently recognized as a well-supported PSH, including the previou described taxa *B. weberi*, *B. weberi* formosana, *B.arcana*, ~~*Erpobdella wuttkei*~~ ~~or~~ *Barbronia wuutttkei* (see clade A in Figure. 1). The Primary Species Hypotheses (PSHs) of *B. weberi* groups proposed by bPTP were consistent with bGMYC, but the difference between two results was the species delimitation of individuals listed in the *B. gwalagwalensis* group. (See clade B in Figure. 1). In the bPTP results, the five individuals of *B. gwalagwalensis* group in clade B were divided into two PSHs. One well supported PSH comprised only *B.* cf.

gwalagwalensis, and the other moderately supported PSH encompassed the paratype of *B. gwalagwalensis* (GenBank accession number: AY786455) and the remaining specimens (MF458701, MN295405 and MN503261, see Figure. 1 and Table 1). . However, the five individuals of *B. gwalagwalensis* group was either split into four well-supported PSHs or pooled into one moderate PSH in bGMYC (see Figure. 1). In the four well-supported PSHs, one PSH contains MF458701 and AY786455, and each of the remaining PSHs was formed separately from MN295405, MN503261and OQ339201. ~~*B.*~~ cf *gwalagwalensis* (OQ339201) alone was consistently recognized as a separately well-supported PSH in both bGMYC and bPTP analyses (threshold 0.95-1), although it was also somewhat classified with the paratype *B. gwalagwalensis* (AY786455) in a moderately supported PSH (threshold 0.5-0.9) by bGMYC. Moreover, the uncorrected p-distance of COI sequences between *B.* cf *gwalagwalensis* (OQ339201) and the paratype *B. gwalagwalensis* (AY786455) was 4.4% (see Figure. 2), which was substantially high in compared to the p-distance (ranged from 0.2% to 1.2%) among four individuals in clade B of figures 1 collected from Myanmar (MN295405), South Korea (MN503261), South Africa (AY786455) and France (MF458701) (see Figure. 2 and Table 1). The p-distance between above four individuals and *B.* cf *gwalagwalensis* varied from 3.8% to 4.5%. In contrast, the p-distances of seven *B. weberi* individuals in clade A (see figure 1) sampled from South Africa, the United States, Mexico, Costa Rica, and Germany were less than 0.8%, and the interspecific uncorrected p-distance between *B. weberi* group and *B. gwalagwalensis* group (clade A and clade B, respectively, see Figure 1) was started from 6.4% to 8.7%. Considering the result of species delimitation and the sufficiently large p-distance mention above, this specimen sampled in China (OQ339201) was treated as *B.* cf. *gwalagwalensis*.

Phylogenetic Relationships within Erpobdelliformes

The topologies of both the BI tree and ML tree ~~were~~ well supported with only minor differences in the placement of Orobdellidae and Gastrostomobdellidae *sensu stricto* (see Figure. 3). With setting a well-supported clade (PP = 1.00, and BS = 100) of 18 *Erpobdella* species as outgroups, both BI and ML analyses show a clear separation of the three main monophyletic groups with strong nodal support values (PP = 1.00, and BS = 100). The first clade was Salifidae represented by 7 species from five genera (*Barbronia*, *Mimobdella*, *Odontobdella*, *Linta,* and *Salifa*), the second clade was Orobdellidae consisted of 18 *Orobdella* species, and the last one was

Gastrostomobdellidae *sensu stricto* on behalf of 3 *Gastrostomobdella* species. *Barbronia* was a well-supported (PP = 1.00, BS = 100) monophyletic group, and its closely related clade consisted of *Mimobdella* and *Odontobdella*. The genus *Salifa* was recovered as paraphyletic with the inclusion of *Linta be*. Gastrostomobdellidae *sensu stricto* (containing genus *Orobdella* only), but not Gastrostomobdellidae *sensu lato* (*Orobdella* + *Gastrostomobdella*), was a well-supported monophyletic group.

General Features of Mitochondrial Genomes and nuclear rDNAs

The newly assembled mitochondrial genome of *B.* cf. *gwalagwalensis* was AT-rich (71.9%) circular mapping molecule (total length, 14847 bp), with an average 209-fold coverage (Figure. 4 and Table 3). The components of *B.* cf. *gwalagwalensis* mitogenome consist of 13 protein-coding genes (PCGs), 22 ~~transport~~ RNAs (tRNAs), 2 rDNAs (12S and 16S rDNA), and a possible control region (451 bp, the longest no-coding reign located between tRNA R and tRNA H). All mitochondrial genes of *Barbronia* were encoded on the same strand, and gene order was generally consistent with previous published erpobdellids mitogenomes. Among 13 *Barbronia* PCGs, both NAD2 and NAD5 genes were inferred to use ATT as an initiation codon, the COX3 gene was initiated with ATA, and the remaining genes used ATG as a start codon. The predicted secondary structures of 22 tRNAs had a similar clover leaf shape (see Figure. 5), and these tRNAs range in size from 60 to 69 bp, the shortest tRNA gene was tRNA L and the longest one was tRNA Q. The complete length of nuclear 18S rDNA and 28S rDNAs were 1876 bp and 2863 bp, respectively, the former one was identical but longer than the particle 18S rDNA sequence (AY786462) of *B. gwalagwalensis*.

Discussion

The last two decades have witnessed that species identification based on molecular data was indeed provided a valuable complement to morphological taxonomy (Mahadani et al. 2022), aided by the increasing availability of genetic techniques (Dellicour & Flot 2015). In the current study, seven individuals of *B. weberi* constituted a well-supported clade in the COI phylogenetic results, this clade was also recognized as one valid species both in bGMYC and bPTP analyses

(see Figure. 1). The current species delimitation results supported the previous proposed three taxa (*B. arcana, B. weberi* formosana and *B. wuttkei* or *Erpobdella wuttkei*) were synonyms of *B. weberi* rather than valid species, which was generally consistent with previous studies (Klass et al. 2021; Oceguera-Figueroa et al. 2011). The COI sequences similarity of seven individuals in *B. weberi* group (clade A in Figure 1) collected from South Africa, the United States, Mexico, Costa Rica, and Germany were nearly identical. The lack of molecular differentiation (small p-distance) between them can be explained by the large effective population size but low speciation ratio, which was caused by both a relatively large distribution of invasive *B. weberi* and the potentially low speciation rate (i.e., reproduce cocoon without cross-fertilization as mentioned in the introduction). However, due to a gene tree was not always consistent with a species tree when population sizes were large and speciation rates were high (Dellicour & Flot 2015), single locus-based methods may not be sufficient for delimiting specimens listed in the group of *B. gwalagwalensis*, especially considering the relatively high genetic diversity between specimens sampled from southern Africa, France, Korea, and China (see Figure. 2). Whereas, on basis of all available *Barbronia* data listed in Table 1, all *Barbronia* individuals were tended to lump into one taxon by applying multiple loci-based species delimitation method. This was likely to cause by insufficient data. In such a case, adding data from both the mitochondrial genome and nuclear loci collected from a few *B. gwalagwalensis* specimens (listed in the table 1), rather than only partial mitochondrial gene COI, was likely to promise for solving the species delimitation problem of *B. gwalagwalensis*.

Although cryptic speciation is common among clitellates (Erséus & Gustafsson 2009; Liu et al. 2017; Martinsson & Erséus 2021), the conclusion that treated *Barbronia* cf *gwalagwalensis* as a cryptic species rather than a species can be drawn until using both multiple loci-base approaches and morphological discriminations (Dellicour & Flot 2015; Jorger & Schrodl 2013; Stengel et al. 2022; Sukumaran et al. 2021). This specimen named *Barbronia* cf. *gwalagwalensis* was first recognized as *B. weberi* due to the observation of accessory gonopores (the anterior and posterior ones separately close to the male and female gonophores). This external morphological character was obviously similar to the invasive species *B. weberi*, but different from the other two known endemic *Barbronia* species in China (Yang et al. 1997). Accessory gonophores were also found on *B. gwalagwalensis* (Westergren & Siddall 2004), however, there was no records about *B.*

gwalagwalensis in China yet. This specimen was also not possible to be either *B. zhejiangica* or *B. yunnanensis* due to the presence of two accessory gonopores, and it was closely related to *B. gwalagwalensis*, including the Korean specimen (accession number: KF966549, voucher SOKN017). Although this Korean one was tentatively reassigned to *Barbronia* cf. *zhejiangica* by Klass et al. (2021), the specimen found in China was geographically close to the type locality of the nominal taxon *B. zhejiangica* than the Korean specimen. Unfortunately, the morphological traits of this specimen now were impossible to check in detail due to the whole body had been used for DNA extraction in NGS. Consequently, this specimen was lumped into the taxon *B.* cf *gwalagwalensis* rather than described as new species, further studies on more samples with morphological delineation were needed before drawing reliable conclusions. It was worth noting that *B. gwalagwalensis*, first reported in South Africa but now also recorded in Myanmar, was possibly distributed widely (Klass et al. 2021), and the current study confirms the need of further explore its species distribution not only through extensive fieldwork in China but also through increased collecting efforts across Asia and Europe.

The entire mitochondrial genomes, as well as the full length of nuclear rDNAs, have been now routinely applied to assess species boundaries and deep relationships in many phylogenetic studies due to obtaining sufficient data efficiently through NGS (Jia et al. 2023; Moreno-Carmona et al. 2023; Prada et al. 2023). However, neither one whole mitochondrial genome of *Barbronia* species nor even the family Salifidae has been published until now. In the current study, the first complete mitochondrial genome of salifid leeches was assembled, and compared with the four whole or incomplete mitochondrial genome sequences of Erpobdellids (see Table 3), representing 3 nominal species and 1 unrecognized species in the GenBank. The specimen named *Erpobdella octoculata*, KC688270 or NC_023927, was not included here, since it is likely to be a misidentified specimen (Oceguera-Figueroa et al. 2016). Clearly, it was far from insufficient using these data to estimate either the specie delimitation of *Barbronia* species or the phylogenetic relationships of *Barbronia* within Erpobdelliformes (Arhynchobdellida, Hirudinea). Therefore, four conventional molecular markers used in many previous studies of Erpobdelliformes were collected and analyzed, including full-length 18S and 28S rDNA sequences and COI and 12S rDNA sequences extracted from the newly acquired mitochondrial genome. The four well-supported clades, *i.e.*, Salifidae, Orobdellidae, Gastrostomobdellidae

sensu stricto, and *Erpobdella*, were clearly distinguished in the current molecular phylogenetic results. The *Barbronia* clade was sister to a clade consisting of *Odontobdella blanchardi* and *Mimobdella japonica* with well supported in the current study (see Figure. 3), and the monophyly of the *Barbronia* was congruent with previous research (Klass et al. 2021; Nakano et al. 2018; Nakano & Nguyen 2015; Oceguera-Figueroa et al. 2011). However, the *Salifa* genus represented by two species was not monophyletic in the current study and in Nakano et al. (2018) result, which was not consistent with previous studies (Klass et al. 2021; Nakano & Nguyen 2015). The Orobdellidae clade was well supported, and it was paraphyletic to the clade of *Gastrostomobdella sensu lato*, which was congruent with previous studies (Nakano 2016a; Nakano 2016b; Nakano 2022; Nakano et al. 2018; Nakano & Nguyen 2015; Nakano et al. 2012). In addition, the basal phylogenetic placement of Erpobdella within the Erpobdelliformes when using different outgroups is maybe controversial (Nakano et al. 2018; Nakano & Nguyen 2015). To solve these problems, further studies of additional species or at least simultaneously obtaining more molecular data (i.e., mitochondrial genomes and rDNAs) from a few valuable specimens using NGS were urgently needed. The coming systematics study of *B. gwalagwalensis* as well as other closely related species will continue to benefit from the methods used in the current study, which simultaneously acquire the mitochondrial genome and full-length 18S and 28S rDNAs. This approach bridges the gaps between sequences amplified with different primers, especially when the conventional primers do not perform well (Siddall 2002b; Trontelj et al. 1999).

Conclusions

The species delimitation results supported that the previously proposed taxa (*Barbronia arcana*, *B. weberi* formosana, and *B. wuttkei* or *Erpobdella wuttkei*) were synonyms of *B. weberi*, but the taxonomic status of *B. gwalagwalensis* and *B.* cf *gwalagwalensis* need to be further studied through extensive fieldwork in China as well as increased collection efforts across from Asia to Europe. *Barbronia* represented by *B. weberi*, *B. gwalagwalensis*, and *B.* cf *gwalagwalensis* formed a well-supported clade, and was sister to a well-supported clade of *Odontobdella* and *Mimobdella*. They constituted a strongly supported monophyletic group and were incorporated into Salifidae. Additionally, the families Salifidae, Gastrostomobdellidae *sensu stricto* containing only the genus *Orobdella*, and Orobdellidae were all well-supported monophyletic clades. Moreover, the first complete mitochondrial genome, full-length 18S and 28S rDNAs of salifid

leeches within the family Salifidae were provided. According to the results of this study, the strategy of obtaining both whole mitochondria and nuclear markers from extensively sampled Salifids species using NGS was expected to fathom both the species diversity of *B. gwalagwalensis* and the evolutionary relationship of Salifidae.

Acknowledgments

no.

References

Altekar G, Dwarkadas S, Huelsenbeck JP, and Ronquist F. 2004. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20:407-415. 10.1093/bioinformatics/btg427

Bandyopadhyay PK, and Mandal CK. 2005. On the Occurrence of Five Leeches Barbronia weberi, Glossiphonia annandalei, Glossiphonia heteroclita, Glossiphonia reticulata and Placobdella undulata in West Bengal, India. *Records of the Zoological Survey of India* 105:93-101.

Blanchard R. 1897. *Hirudinées des Indes Néerlandaises*: Zoologische Ergebnisse in Indes Neerlandaises.

Blasco-Costa I, Poulin R, and Presswell B. 2016. Species of *Apatemon* Szidat, 1928 and *Australapatemon* Sudarikov, 1959 (Trematoda: Strigeidae) from New Zealand: linking and characterising life cycle stages with morphology and molecules. *Parasitology Research* 115:271-289. 10.1007/s00436-015-4744-0

Borda E, and Siddall ME. 2004a. Arhynchobdellida (Annelida: Oligochaeta: Hirudinida): phylogenetic relationships and evolution. *Molecular Phylogenetics and Evolution* 30:213-225. 10.1016/j.ympev.2003.09.002

Borda E, and Siddall ME. 2004b. Review of the evolution of life history strategies and phylogeny of the Hirudinida (Annelida: Oligochaeta). *Lauterbornia* 52:25.

Chandra M, and Mahajan KK. 1971. Leeches of Simla Hills, Himachal Pradesh. *Records of the Zoological Survey of India* 69:255-259.

~~de Oca UGM,~~ Martinez-Flores G, Contreras-Miron S, Jimenez-Armenta J, and Oceguera-Figueroa A. 2016. First record of the introduced leech *Barbronia weberi* (Annelida: Clitellata) in Nuevo Leon, Mexico. *Revista Mexicana De Biodiversidad* 87:1124-1126. 10.1016/j.rmb.2016.06.008

Dellicour S, and Flot JF. 2015. Delimiting Species-Poor Data Sets using Single Molecular Markers: A Study of Barcode Gaps, Haplowebs and GMYC. *Syst Biol* 64:900-908. 10.1093/sysbio/syu130

Eriksen TE, Jacobsen D, Demars BOL, Brittain JE, Søli G, and Friberg N. 2022. Effects of pollution-induced changes in oxygen conditions scaling up from individuals to ecosystems in a tropical river network. *Science of The Total Environment* 814:151958. https://doi.org/10.1016/j.scitotenv.2021.151958

Erséus C, and Gustafsson D. 2009. Cryptic Speciation in Clitellate Model Organisms. *Annelids in Modern Biology*, 31-46.

~~Figueroa AO,~~ Régagnon VL, and Siddall ME. 2005. Phylogeny and revision of Erpobdelliformes (Annelida, Arhynchobdellida) from Mexico based on nuclear and mithochondrial gene sequences. *Revista Mexicana De Biodiversidad* 76:191-198.

Franco-Sierra ND, and Diaz-Nieto JF. 2020. Rapid mitochondrial genome sequencing based on Oxford Nanopore Sequencing and a proxy for vertebrate species identification. *Ecology and Evolution* 10:3544-3560. 10.1002/ece3.6151

Fujisawa T, and Barraclough TG. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst Biol* 62:707-724. 10.1093/sysbio/syt033

Genoni P, and Fazzone A. 2008. Barbronia weberi (R. Blanchard, 1897) (Hirudinea: Salifidae), an Asian leech species new to Italy. *Aquatic Invasions* 3:77-79.

Ghate HV. 1991. Extension of range of freshwater leech *Barbronia weberi* (Blanchard) (Arhynchobdellae: Erpobdellidae) from Pune, Maharashtra.

Govedich FR, Bain BA, Burd M, and Davies RW. 2003. Reproductive biology of the invasive Asian freshwater leech *Barbronia weberi* (Blanchard, 1897). *Hydrobiologia* 510:125-129. DOI 10.1023/B:HYDR.0000008638.87536.b0

Govedich FR, Bain BA, and Davies RW. 2002. First record of the Asian freshwater leech *Barbronia weberi* (Blanchard, 1897) (Euhirudinea: Erpobdellidae) in Australia. *Victorian Naturalist (Blackburn)Victorian Naturalist (Blackburn)* 119.

Grosser C, and Trontelj P. 2008. On the taxonomic status of *Barbronia wuttkei* (Kutschera, 2004) n. comb.–a leech species (Hirudinea: Salifidae) from a German aquarium.

Jia Y, Qiu G, Cao C, Wang X, Jiang L, Zhang T, Geng Z, and Jin S. 2023. Mitochondrial genome and phylogenetic analysis of Chaohu duck. *Gene* 851:147018. https://doi.org/10.1016/j.gene.2022.147018

Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, and Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* 21. 10.1186/s13059-020-02154-5

Jorger KM, and Schrodl M. 2013. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front Zool* 10:59. 10.1186/1742-9994-10-59

Katoh K, and Standley DM. 2014. MAFFT: Iterative Refinement and Additional Methods. *Multiple Sequence Alignment Methods* 1079:131-146. 10.1007/978-1-62703-646-7_8

Klass AL, Kondakov AV, Vikhrev IV, Bespalaya YV, Lunn Z, Chan N, Gofarov MY, and Bolotov IN. 2021. Is the South African leech *Barbronia gwalagwalensis* Westergren & Siddall, 2004 (Hirudinida: Erpobdelliformes: Salifidae) a Paleotropical species? *Zootaxa* 4974:585-595. 10.11646/zootaxa.4974.3.7

Kortsinoglou AM, Saud Z, Eastwood DC, Butt TM, and Kouvelis VN. 2020. The mitochondrial genome contribution to the phylogeny and identification of Metarhizium species and strains. *Fungal Biology* 124:845-853. 10.1016/j.funbio.2020.06.003

Kuang WM, and Yu L. 2019. Mitogenome assembly strategies and software applications in the genome era. *Yi Chuan* 41:979-993. 10.16288/j.yczz.19-227

Kutschera U. 2004. Description of a new leech species, *Erpobdella wuttkei* nov. sp. (Hirudinea: Erpobdellidae). *Lauterbornia* 52:147-151.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, and Calcott B. 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. *Molecular Biology and Evolution* 34:772-773. 10.1093/molbev/msw260

Liu YK, Fend SV, Martinsson S, and Erseus C. 2017. Extensive cryptic diversity in the cosmopolitan sludge worm *Limnodrilus hoffmeisteri* (Clitellata, Naididae). *Organisms Diversity & Evolution* 17:477-495. 10.1007/s13127-016-0317-z

Ludanyi M, Balla D, Muller Z, and Kiss B. 2019. The first occurrence of *Barbronia weberi* (Blanchard, 1897) (Hirudinea: Arhynchobdellida: Erpobdelliformes: Salifidae) in Hungary. *Bioinvasions Records* 8:633-639. 10.3391/bir.2019.8.3.20

Mahadani P, Dasgupta M, Vijayan J, Kar CS, and Ray S. 2022. DNA Barcoding in Plants: Past, Present, and Future. In: Singh RL, Mondal S, Parihar A, and Singh PK, eds. *Plant Genomics for Sustainable Agriculture*. Singapore: Springer Nature Singapore, 331-350.

Martinsson S, and Erséus C. 2021. Cryptic Clitellata: Molecular Species Delimitation of Clitellate Worms (Annelida): An Overview. *Diversity-Basel* 13. 10.3390/d13020036

Mason, and Julia. 1976. Studies on the freshwater and terrestrial leeches of New Zealand. 2. Orders Gnathobdelliformes and Pharyngobdelliformes. *Journal of the Royal Society of New Zealand* 6:255-276.

Moreno-Carmona M, Montaña-Lozano P, Prada Quiroga CF, and Baeza JA. 2023. Comparative analysis of mitochondrial genomes reveals family-specific architectures and molecular features in scorpions (Arthropoda: Arachnida: Scorpiones). *Gene* 859:147189. https://doi.org/10.1016/j.gene.2023.147189

Nakano T. 2016a. Four new species of the genus Orobdella from Shikoku and Awajishima island, Japan (Hirudinida, Arhynchobdellida, Orobdellidae). *Zoosystematics and Evolution* 92:79-102.

Nakano T. 2016b. A new quadrannulate species of *Orobdella* (Hirudinida, Arhynchobdellida, Orobdellidae) from western Honshu, Japan. *ZooKeys* 553:33-51.

Nakano T. 2022. A new sexannulate species of *Orobdella* (Hirudinea, Arhynchobdellida, Orobdellidae) from Kii-Oshima Island, Japan. *Evolutionary Systematics* 6:135-142.

Nakano T, Eto K, Nishikawa K, Hossman MY, and Jeratthitikul E. 2018. Systematic revision of the Southeast Asian macrophagous leeches, with the description of two new gastrostomobdellid species (Hirudinida: Arhynchobdellida: Erpobdelliformes). *Zoological Journal of the Linnean Society* 184:1-30. 10.1093/zoolinnean/zlx097

Nakano T, and Nguyen ST. 2015. A new predatory leech from Vietnam (Hirudinida: Arhynchobdellida: Salifidae): its phylogenetic position with comments on the classification of the family. *Invertebrate Systematics* 29:473-486.

Nakano T, Ramlah Z, and Hikida T. 2012. Phylogenetic position of gastrostomobdellid leeches (Hirudinida, Arhynchobdellida, Erpobdelliformes) and a new family for the genus Orobdella. *Zoologica Scripta* 41:177-185. 10.1111/j.1463-6409.2011.00506.x

Nehring SGS. 2006. National checklist for aquatic alien species in Germany. *Aquatic Invasions* 1:245-269.

Nguyen LT, Schmidt HA, von Haeseler A, and Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32:268-274. 10.1093/molbev/msu300

Oceguera-Figueroa A, Manzano-Marin A, Kvist S, Moya A, Siddall ME, and Latorre A. 2016. Comparative Mitogenomics of Leeches (Annelida: Clitellata): Genome Conservation and Placobdella-Specific trnD Gene Duplication. *PloS one* 11:e0155441. 10.1371/journal.pone.0155441

Oceguera-Figueroa A, Phillips AJ, Pacheco-Chaves B, Reeves WK, and Siddall ME. 2011. Phylogeny of macrophagous leeches (Hirudinea, Clitellata) based on molecular data and evaluation of the barcoding locus. *Zoologica Scripta* 40:194-203. 10.1111/j.1463-6409.2010.00465.x

Pamplin PAZ, and Rocha O. 2000. First report of *Barbronia weberi* (Hirudinea : Salifidae) from South America. *Revista De Biologia Tropical* 48:723-723.

Pavluk T, Pavluk E, and Rasines R. 2011. First record of the Asian leech *Barbronia weberi* (Blanchard, 1897) (Hirudinea: Arhynchobdellida: Erpobdelliformes: Salifidae) in the Iberian Peninsula. *Aquatic Invasions* 6:S61-S64. 10.3391/ai.2011.6.S1.014

Prada CF, Hazzi NA, Hormiga G, Cabarcas F, and Franco LM. 2023. Complete mitochondrial genome of Phoneutria depilata (Araneae, Ctenidae): New insights into the phylogeny and evolution of spiders. *Gene* 850:146925. https://doi.org/10.1016/j.gene.2022.146925

Reid NM, and Carstens BC. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol Biol* 12:196. 10.1186/1471-2148-12-196

Rutter RP, and Klemm DJ. 2001. The presence of an Asian leech, Barbronia weberi, in a home aquarium in South Florida (Hirudinea: Salifidae). *Florida Scientist*:216-218.

Sawyer RT. 1986. *Leech Biology and Behaviour. Volume 2: Feeding Biology, Ecology, and Systematics*. Oxford, UK: Oxford University Press.

Sawyer RT. 2020. Reproduction without cross-fertilisation in the invasive Asian leech *Barbronia weberi* (Blanchard, 1897) (Hirudinea: Arhynchobdellida). *Aquatic Invasions* 15:271-281. 10.3391/ai.2020.15.2.04

Sawyer RT, and Sawyer DR. 2018. The alien Asian leech *Barbronia weberi* (Blanchard, 1897) (Hirudinea: Salifidae) reported from two disjunct localities in North Carolina, United States, with observations on its biology and potential for laboratory research. *Bioinvasions Records* 7:61-64. 10.3391/bir.2018.7.1.09

Siddall ME. 2002a. Phylogeny of the leech family Erpobdellidae (Hirudinida : Oligochaeta). *Invertebrate Systematics* 16:1-6. 10.1071/It01011

Siddall ME. 2002b. Phylogeny of the leech family Erpobdellidae (Hirudinida: Oligochaeta). *Invertebrate Systematics* 16:1-6.

Stengel A, Stanke KM, Quattrone AC, and Herr JR. 2022. Improving Taxonomic Delimitation of Fungal Species in the Age of Genomics and Phenomics. *Front Microbiol* 13:847067. 10.3389/fmicb.2022.847067

Sukumaran J, Holder MT, and Knowles LL. 2021. Incorporating the speciation process into species delimitation. *PLoS Comput Biol* 17:e1008924. 10.1371/journal.pcbi.1008924

Ta-Hsiang S. 1974. Preliminary notes on the behaviour of some Paddy-field leeches from Ninghsien Chekiang Province China. *Acta Zoologica Sinica*.

Trontelj P, Sket B, and Steinbrück G. 1999. Molecular phylogeny of leeches: congruence of nuclear and mitochondrial rDNA data sets and the origin of bloodsucking. *Journal of Zoological Systematics and Evolutionary Research* 37:141-147.

Van Haaren T, Hop P, Soes M, and Tempelman D. 2004. The freshwater leeches (Hirudinea) of the Netherlands. *Lauterbornia* 52:113-131.

519 Westergren S, and Siddall ME. 2004. Two new species of *salifid* leeches (Arhynchobdellida, Erpobdelliformes,
520 Salifidae) from South Africa and Madagascar. *American Museum Novitates* 3456:1-6
521 Will S, Joshi T, Hofacker IL, Stadler PF, and Backofen R. 2012. LocARNA-P: Accurate boundary prediction and
522 improved detection of structural RNAs. *Rna* 18:900-914. 10.1261/rna.029041.111
523 Yang T. 1996. *Fauna Sinica, Annelida: Hirudinea*. Beijing, China: Science Press.
524 Yang T, Wang D-b, and Zhang Y-p. 1997. A New Species and A New Record of Freshwater leeches From China
525 (Arhynchobdellida:Salifidae). - *Zoological Research* - 18:- 383.
526 Zhang D, Gao FL, Jakovlic I, Zou H, Zhang J, Li WX, and Wang GT. 2020. PhyloSuite: An integrated and scalable
527 desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics
528 studies. *Molecular Ecology Resources* 20:348-355. 10.1111/1755-0998.13096
529 Zhang JJ, Kapli P, Pavlidis P, and Stamatakis A. 2013. A general species delimitation method with applications to
530 phylogenetic placements. *Bioinformatics* 29:2869-2876. 10.1093/bioinformatics/btt499
531 Zhao CS, Pan X, Wang XL, Shao NF, Sun CL, and Zhang Y. 2020. Hydrology and water quality factors driving
532 spatiotemporal analyses of zoobenthos in a pilot city in China. *Ecohydrology* 13. 10.1002/eco.2187
533

Figure 1

The results of the Primary Species Hypothesis (PSHs).

The bPTP and bGMYC results were summarized and visualized, the left side was the maximum clade credibility tree from BEAST analyses, and the species delimitation results, using bPTP and bGMYC methods based on COI data, were showed on the right side, with colors corresponding to the posterior probability of same Primary Species Hypotheses (PSHs) under a specific threshold (at the upper left). The accession numbers and taxon names in GenBank presented besides underlines at the tip of the tree. Clade A and clade B referred to *B. weberi* gourp and *B. gwalagwalensis* group, respectively.

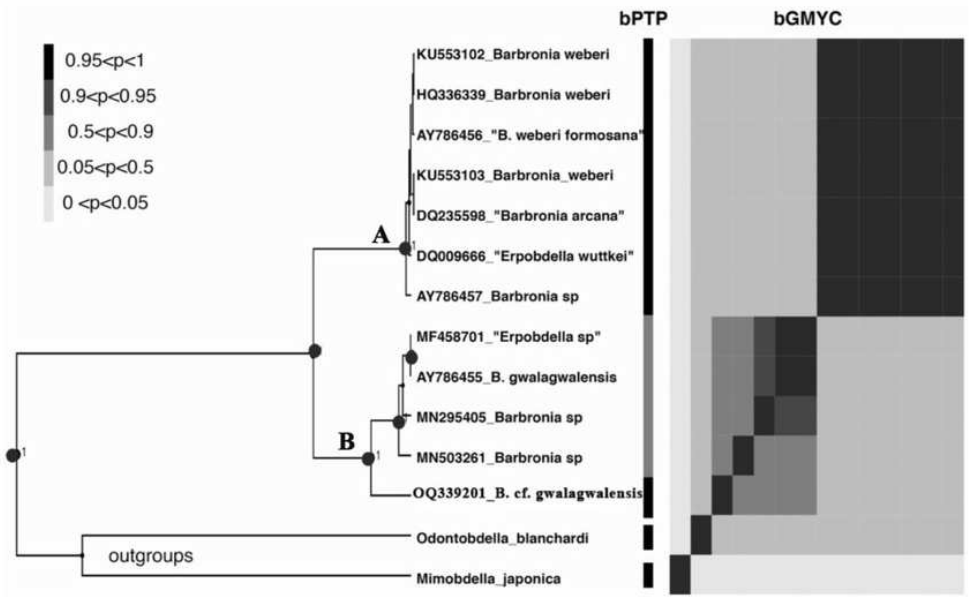

Figure 1. The results of the Primary Species Hypothesis (PSHs). The bPTP and bGMYC results were summarized and visualized, the left side was the maximum clade credibility tree from BEAST analyses, and the species delimitation results, using bPTP and bGMYC methods based on COI data, were showed on the right side, with colors corresponding to the posterior probability of same Primary Species Hypotheses (PSHs) under a specific threshold (at the upper left). The accession numbers and taxon names in GenBank presented besides underlines at the tip of the tree. Clade A and clade B referred to *B. weberi* gourp and *B. gwalagwalensis* group, respectively.

Figure 2

Summary of the bGMYC and bPTP species delimitation methods using the mitochondrial data set.

The uncorrected p-distance matrix corresponding to the alignment of the *Barbronia* COI sequences. The uncorrected p-distance was calculated in MEGA X using the pairwise deletion option, the uncorrected p-distances were visualized in the upper triangular portion of this matrix, with a color bar (0 ~ 10% uncorrected p-distance). The intraspecific and interspecific uncorrected p-distances were represented by the red and blue circles, and the size of circles indicate the value of the corresponding uncorrected p-distances which were listed in triangular portion of this matrix. At the left, the accession numbers and taxon names in GenBank presented besides underlines, corresponding to the numbers from 1 to 12 at the top.

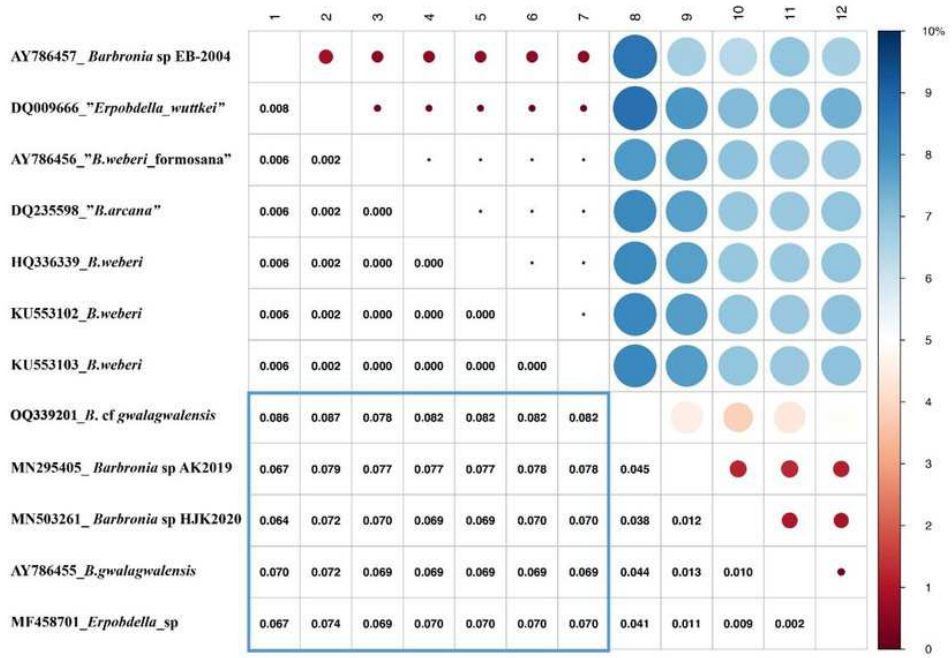

Figure 2. Summary of the bGMYC and bPTP species delimitation methods using the mitochondrial data set. The uncorrected p-distance matrix corresponding to the alignment of the *Barbronia* COI sequences. The uncorrected p-distance was calculated in MEGA X using the pairwise deletion option, the uncorrected p-distances were visualized in the upper triangular portion of this matrix, with a color bar (0 ~ 10% uncorrected p-distance). The intraspecific and interspecific uncorrected p-distances were represented by the red and blue circles, and the size of circles indicate the value of the corresponding uncorrected p-distances which were listed in triangular portion of this matrix. At the left, the accession numbers and taxon names in GenBank presented besides underlines, corresponding to the numbers from 1 to 12 at the top.

Figure 3

The Comparison of BI and ML trees constructed using a concatenated data of both mitochondrial genes (COI, 12S rDNA) and nuclear markers (18S and 28S rDNA), respectively.

The phylogenetic tree was divided into four main clusters, with setting a well-supported clade (PP = 1.00, and BS = 100) of 18 *Erpobdella* species as outgroups, the monophyly of the genus *Barbronia* and three families (Salifidae represented by five genus, Orobdellidae, Gastrostomobdellidae *sensu stricto*) were well supported. Only one specimen of each species was in bold face, the last two *Erpobdella* species name with asterisk were the synonym of *Dina lineata* and *Mooreobdella melanostoma*, respectively. The size of circles indicates that either posterior probabilities or the bootstrap values of corresponding nodes were estimated in BI or ML analyses. The 28S rDNA of *B. gwalagwalensis* (AY786449) was not included in the current analyses, since it was a partial conserved region of 28S rDNA but significantly different from other 28S rDNA sequences of Barbronia listed in the Table 1.

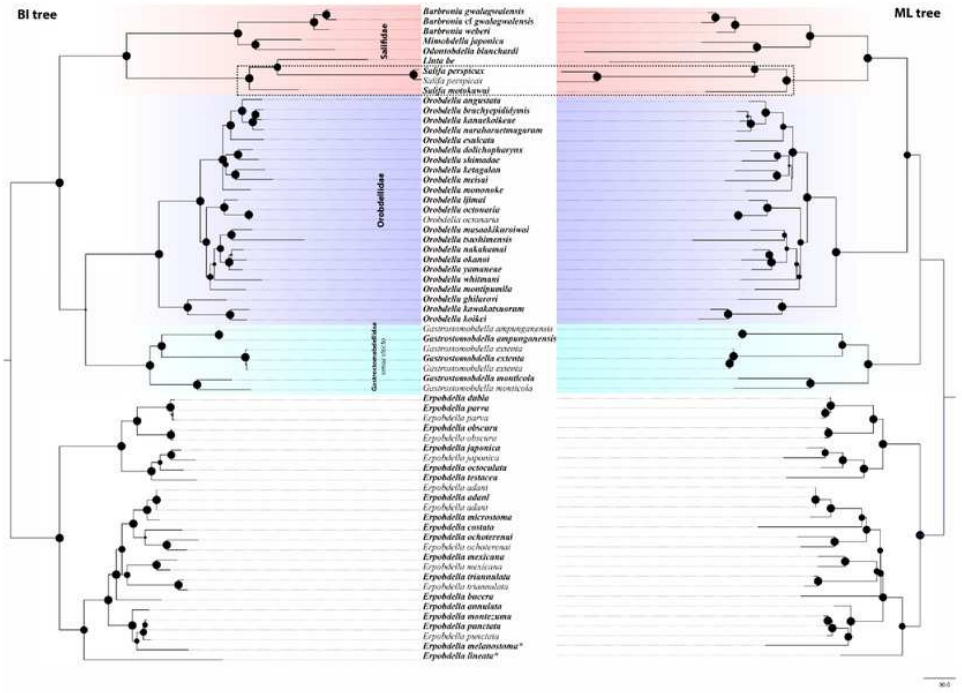

Figure 3. The Comparison of BI and ML trees constructed using a concatenated data of both mitochondrial genes (COI, 12S rDNA) and nuclear markers (18S and 28S rDNA), respectively. The phylogenetic tree was divided into four main clusters, with setting a well-supported clade (PP = 1.00, and BS = 100) of 18 *Erpobdella* species as outgroups, the monophyly of the genus *Barbronia* and three families (Salifidae represented by five genus, Orobdellidae, Gastrostomobdellidae *sensu stricto*) were well supported. Only one specimen of each species was in bold face, the last two *Erpobdella* species name with asterisk were the synonym of *Dina lineata* and *Mooreobdella melanostoma*, respectively. The size of circles indicates that either posterior probabilities or the bootstrap values of corresponding nodes were estimated in BI or ML analyses. The 28S rDNA of *B. gwalagwalensis* (AY786449) was not included in the current analyses, since it was a partial conserved region of 28S rDNA but significantly different from other 28S rDNA sequences of *Barbronia* listed in the Table 1.

Figure 4

Mitochondrial genome organization of *Barbronia* cf. *gwalagwalensis*.

The mitochondrial genome of *Barbronia* cf. *gwalagwalensis* consists of 14,847 base pairs, which encodes 22 tRNA, the 12 S and the 16 S rRNA, and 13 coding genes (ATP6 and ATP8: subunits of ATP synthase, COX1-3: Cytochrome c oxidase subunits, CYTB: Cytochrome b. ND1-6: NADH dehydrogenase subunits).

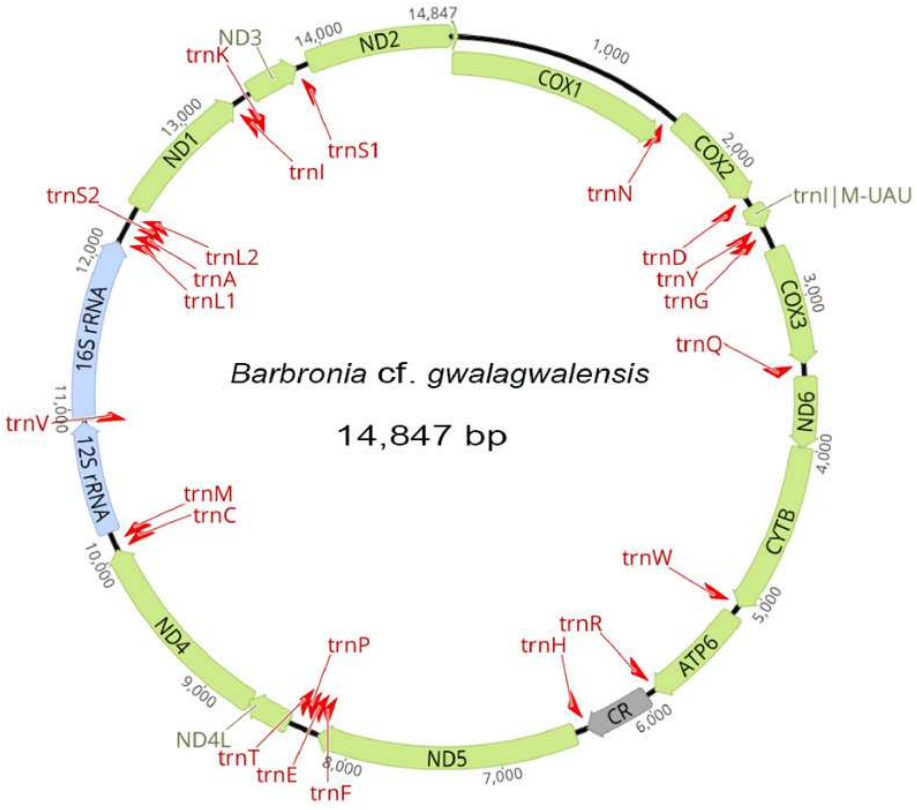

Figure 4. Mitochondrial genome organization of *Barbronia* cf. *gwalagwalensis*. The mitochondrial genome of *Barbronia* cf. *gwalagwalensis* consists of 14,847 base pairs, which encodes 22 tRNA, the 12 S and the 16 S rRNA, and 13 coding genes (ATP6 and ATP8: subunits of ATP synthase, COX1-3: Cytochrome c oxidase subunits, CYTB: Cytochrome b. ND1-6: NADH dehydrogenase subunits).

Figure 5

Secondary structures of tRNA genes in the mitogenome of *Barbronia* cf. *gwalagwalensis.*

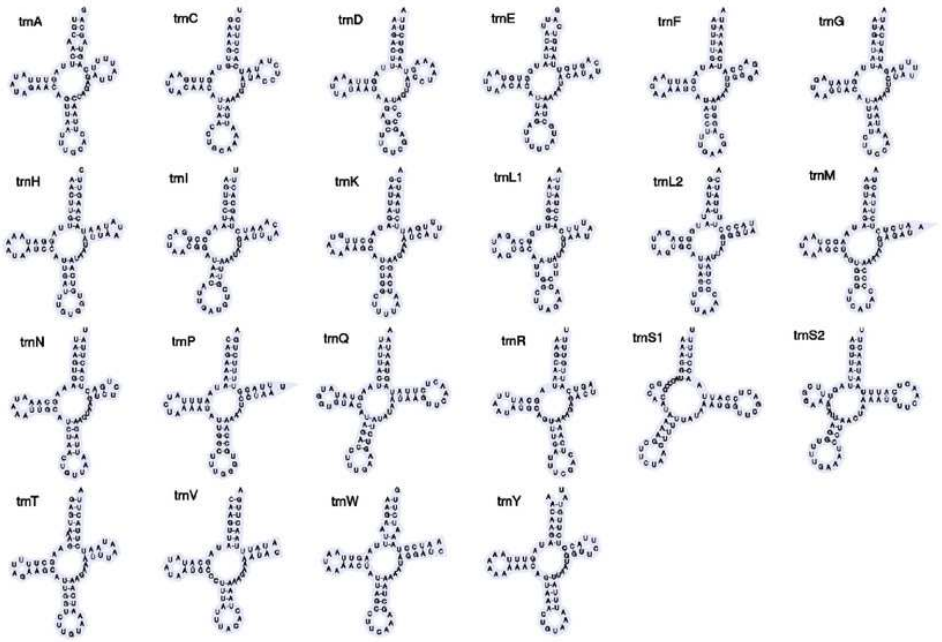

Figure 5. Secondary structures of tRNA genes in the mitogenome of *Barbronia* cf. *gwalagwalensis*.

Table 1(on next page)

The sequences list of all available *Barbronia* and two outgroups in the GenBank.

All COI sequences list here were used in bGMYC and bPTP analyses. The groups referred to the clade A, clade B and outgroups of the COI tree in Figure 1, respectively. Taxon name with quotations means these taxa were likely to be misidentified. The accession number with * represents new data, and "-" means no available sequences.

Table 1. The sequences list of all available *Barbronia* and two outgroups in the GenBank.

Groups	Taxon	Country	COI accession	COI size(bp)	12S accession	18S accession	28S accession
Barbronia weberi	*B. weberi*	Costa Rica	HQ336339	649	-	AF099951	HQ336356
	B. weberi formosana	United States	AY786456	600	-	AY786461	AY786448
	"B. arcana"	Mexico	DQ235598	649	DQ235588	DQ235608	-
	B. weberi	Mexico	KU553102	645	-	-	-
	B. weberi	Mexico	KU553103	645	-	-	-
	"Erpobdella wuttkei"	Germany	DQ009666	584	-	-	-
	Barbronia sp	South Africa	AY786457	360	-	AY786463	AY786450
Barbronia gwalagwalensis	*B.* cf. *gwalagwalensis*	China	OQ339201*	1534	OQ339201*	OQ269483*	OQ269482*
	B. gwalagwalensis	South Africa	AY786455	480	-	AY786462	AY786449
	Barbronia sp	Myanmar	MN295405	660	-	-	-
	Barbronia sp	South Korea	MN503261	710	-	MT010330	-
	"*Erpobdella*. sp"	France	MF458701	634	-	-	-
OUTGROUPS	*Mimobedlla japonica*	Japan	AB679658	1267	-	-	-
	Odontobdella blancharidi	Japan	AB938004	1267	-	-	-

All COI sequences list here were used in bGMYC and bPTP analyses. The groups referred to the clade A, clade B and outgroups of the COI tree in Figure 1, respectively. Taxon name with quotations means these taxa were likely to be misidentified. The accession number with * represents new data, and "-" means no available sequences.

Table 2(on next page)

A list of erpobdellids leech specimens and GenBank accession number of four molecular marks used in the current phylogeny reconstruction.

The accession number with * represents new data, and "-" means no available sequences.

Table 2. A list of erpobdellids leech specimens and GenBank accession number of four molecular marks used in the current phylogeny reconstruction.

Family	Species	COI	12S	18S	28S
Salifidae	*Barbronia* cf *gwalagwalensis*	OQ339201*	OQ339201*	OQ269483*	OQ269482*
Salifidae	*Barbronia gwalagwalensis*	AY786455	-	AY786462	-
Salifidae	*Barbronia weberi*	HQ336339	DQ235588	AF099951	HQ336356
Salifidae	*Odontobdella blanchardi*	AB675016	AB675017	AB663651	AB663671
Salifidae	*Mimobdella japonica*	AB675014	AB675015	AB663650	AB663670
Salifidae	*Linta be*	AY786460	-	AY786466	AY786453
Salifidae	*Salifa motokawai*	LC029431	LC029432	LC029434	LC274548
Salifidae	*Salifa perspicax*	HQ336341	HQ336349	HQ336375	HQ336359
Salifidae	*Salifa perspicax*	HQ336343	HQ336351	HQ336377	HQ336360
Orobdellidae	*Orobdella angustata*	LC323139	LC323141	LC323140	LC431606
Orobdellidae	*Orobdella brachyepididymis*	LC106320	LC106318	LC106319	LC274535
Orobdellidae	*Orobdella dolichopharynx*	AB675028	AB675029	AB663665	AB663666
Orobdellidae	*Orobdella esulcata*	AB675020	AB675021	AB663655	AB663656
Orobdellidae	*Orobdella ghilarovi*	LC431609	LC431616	LC431608	LC431607
Orobdellidae	*Orobdella ijimai*	AB675030	AB675031	AB663659	AB663660
Orobdellidae	*Orobdella kanaekoikeae*	LC184548	LC184547	LC184551	LC274533
Orobdellidae	*Orobdella kawakatsuorum*	AB675032	AB675033	AB663661	AB663662
Orobdellidae	*Orobdella ketagalan*	AB704787	AB704788	AB704785	LC274546
Orobdellidae	*Orobdella koikei*	AB679688	AB679689	AB698883	LC274543
Orobdellidae	*Orobdella masaakikuroiwai*	AB938006	AB937997	AB938003	LC274530
Orobdellidae	*Orobdella meisai*	LC314424	LC314422	LC314423	LC431605
Orobdellidae	*Orobdella mononoke*	AB698866	AB698867	AB698868	LC274547
Orobdellidae	*Orobdella montipumila*	LC616663	LC616667	LC616674	LC616673
Orobdellidae	*Orobdella nakahamai*	LC106331	LC106329	LC106330	LC274534
Orobdellidae	*Orobdella naraharaetmagarum*	LC087144	LC087142	LC087143	LC274531
Orobdellidae	*Orobdella octonaria*	AB675024	AB675025	AB663667	AB663668
Orobdellidae	*Orobdella octonaria*	HQ336338	HQ336348	HQ336372	HQ336355

Orobdellidae	*Orobdella okanoi*	LC106342	LC106340	LC106341	LC274532
Orobdellidae	*Orobdella shimadae*	AB675026	AB675027	AB663663	AB663664
Orobdellidae	*Orobdella tsushimensis*	AB675018	AB675019	AB663653	AB663654
Orobdellidae	*Orobdella whitmani*	AB675022	AB675023	AB663657	AB663658
Orobdellidae	*Orobdella yamaneae*	LC106350	LC106348	LC106349	LC274536
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella ampunganensis*	LC274551	LC274564	LC274517	LC274516
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella ampunganensis*	LC274559	LC274568	LC274525	LC274524
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella extenta*	LC274553	LC274565	LC274519	LC274518
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella extenta*	LC274555	LC274566	LC274521	LC274520
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella extenta*	LC274557	LC274567	LC274523	LC274522
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella monticola*	AB675011	AB675010	AB663649	AB663669
Gastrostomobdellidae *sensu stricto*	*Gastrostomobdella monticola*	LC274549	LC274563	LC274514	LC274513
Erpobdellidae	*Dina lineata*	-	AF099952	AF099950	-
Erpobdellidae	*Erpobdella adani*	MG745144	MG745141	MG745138	-
Erpobdellidae	*Erpobdella adani*	MG745145	MG745142	MG745139	-
Erpobdellidae	*Erpobdella adani*	MG745146	MG745143	MG745140	-
Erpobdellidae	*Erpobdella annulata*	HQ336345	-	HQ336379	HQ336362
Erpobdellidae	*Erpobdella bucera*	MN612829	MN613043	MN613063	MN613084
Erpobdellidae	*Erpobdella costata*	AY425460	AY425442	AY425478	AY425406
Erpobdellidae	*Erpobdella dubia*	AF116023	AF462022	AF115997	AY425365
Erpobdellidae	*Erpobdella japonica*	AB675012	AB675013	AB663648	AB663652
Erpobdellidae	*Erpobdella japonica*	AF116026	AF462023	AF116000	AY425366
Erpobdellidae	*Erpobdella mexicana*	DQ235595	DQ235585	DQ235605	HQ336364
Erpobdellidae	*Erpobdella mexicana*	DQ235597	DQ235587	DQ235607	HQ336365
Erpobdellidae	*Erpobdella microstoma*	MN612934	MN613044	MN613065	MN613086
Erpobdellidae	*Erpobdella montezuma*	GQ368760	GQ368820	GQ368802	-
Erpobdellidae	*Erpobdella obscura*	MN613005	MN613045	MN613066	MN613087
Erpobdellidae	*Erpobdella obscura*	MN612911	MN613046	MN613067	MN613088

Erpobdellidae	*Erpobdella ochoterenai*	DQ235596	DQ235586	DQ235606	HQ336370
Erpobdellidae	*Erpobdella ochoterenai*	DQ235599	DQ235590	DQ235609	HQ336371
Erpobdellidae	*Erpobdella octoculata*	HQ336344	-	HQ336378	HQ336361
Erpobdellidae	*Erpobdella parva*	MN612997	MN613052	MN613073	MN613094
Erpobdellidae	*Erpobdella parva*	MN612930	MN613053	MN613074	MN613095
Erpobdellidae	*Erpobdella punctata*	HQ336346	HQ336352	HQ336380	HQ336363
Erpobdellidae	*Erpobdella punctata*	MN612994	MN613056	MN613077	MN613098
Erpobdellidae	*Erpobdella testacea*	AF116027	AF462025	AF116003	AY425370
Erpobdellidae	*Erpobdella triannulata*	DQ235602	DQ235592	DQ235612	HQ336366
Erpobdellidae	*Erpobdella triannulata*	HQ336347	HQ336353	DQ235614	HQ336367
Erpobdellidae	*Mooreobdella melanostoma*	AF116025	AF462027	AF115999	AY425395

The accession number with * represents new data, and "-" means no available sequences.

Table 3 (on next page)

General features of the mitochondrial genomes of *B.* cf*. gwalagwalensis* and other Erpobdellids species.

The specimen (KC688270 or NC_023927, named as *Erpobdella octoculata*) was not included here, since it is likely to be a misidentified specimen.

Table 3. General features of the mitochondrial genomes of *B.* cf. *gwalagwalensis* and other Erpobdellids species.

Taxa	Size (bp)	%GC	CDSs	tRNAs	rRNAs	Accession
B. cf. gwalagwalensis	14,847	28.1	13	22	2	OQ339201
Erpobdella testacea	14,495	26.9	13	20	2	MT584166
Erpobdella japonica	14,725	27.9	13	22	2	NC_036150
Erpobdella octoculata	13,035	27.9	12	20	2	MT410851
Erpobdellidae sp	14,746	30.2	13	22	2	MW431582

The specimen (KC688270 or NC_023927, named as *Erpobdella octoculata*) was not included here, since it is likely to be a misidentified specimen.