

Phylogenetic Relationships among Chinese Rice Frogs within the *Fejervarya limnocharis* Species Complex (Amphibia: Dicroglossidae)

Kun YANG^{1*}, Yubin WO^{2*}, Gang SHAO¹, Pinghu LIAO¹, Haojie TONG¹, Richard P. BROWN^{1*} and Yuanting JIN^{1*}

¹ College of Life Sciences, China Jiliang University, Hangzhou 310018, Zhejiang, China

² Ningbo Beilun District Center for Disease Control and Prevention, Ningbo 315000, Zhejiang, China

Abstract We present a molecular assessment of the widely-distributed rice frog (*Fejervarya limnocharis*) which provides many new samples that add to knowledge of their phylogeography within China and considers genetic support for five Chinese species within this complex. Two mtDNA fragments from 270 individuals and eight nuclear DNA loci (105 individuals) were sequenced from specimens sampled from across China. Data from nine specimens from China, Indonesia and Japan were also retrieved from previous studies. MtDNA was informative about population divergence within China and indicated one major clade (with four subclades) from South China and the Zhoushan Archipelago, Zhejiang, China, and a second major clade (with eight subclades) from other parts of China. Recent demographic expansions (less than 50ka ago) were detected within three of these 12 subclades, potentially associated with lowered sea-levels after marine transgressions. Notably, most frogs from the previously unstudied Zhoushan Archipelago (eastern China) were found to be closely related to Japanese populations. BPP and STACEY species delimitation analyses of the multilocus data favoured five candidate species within the complex. Previous work had described *Fejervarya multistriata* and *F. kawamurai* from the Chinese mainland although here

we detected considerable genetic divergence within the latter and found that this may be indicative of two species. One of these corresponds to the Zhoushan Archipelago, Zhejiang, China and Japan, and the other from most parts of Chinese mainland. This study provides a large multilocus dataset that contributes to the systematics of this species complex.

Keywords amphibian, Asia, genetic structure, phylogeography, population expansion, Zhoushan Archipelago

1. Introduction

Systematics studies at the species level are important not only to determine the most fundamental taxonomic units but also to provide important insights into the historical biogeography of different regions. There has been a significant number of phylogeographic and systematics studies of intraspecific genetic variation in eastern Asia over the last 15 years, which have largely focused on specific regions such as the Qinghai-Tibet Plateau or the north-west Pacific (Favre *et al.*, 2015; Muellner-Riehl, 2019; Ni *et al.*, 2014). However, there is generally less known about intraspecific patterns than in some other regions, such as Europe and North America. Moreover, very few studies have addressed widespread terrestrial animals that are found across major geographical areas.

Analyses that use the multispecies coalescent to analyze groups of sequences are useful tools for detecting population or species splits as well as examining the relationships between them (Heled and Drummond, 2010; Leaché *et al.*, 2019; Yang and Rannala, 2010). Irrespective of differences in interpretation about whether groups represent divergent populations or valid species, they allow probabilistic assessment of several important

* Both authors contributed equally to this paper.

* Corresponding authors: Dr. Yuanting JIN, from China Jiliang University, Hangzhou, Zhejiang, China, with his research focusing on evolutionary herpetology; Prof. Richard BROWN, from China Jiliang University, Hangzhou, Zhejiang, China, with his research focusing on phylogenetics and population genomics, mainly on lizards.

E-mail: jinyuanting@126.com (Yuanting JIN), r.p.brown@ljmu.ac.uk (Richard BROWN)

Received: 4 January 2022 Accepted: 11 March 2022

characteristics of the species (or population) tree including its topology, the likely number of species and the times at which different species (rather than DNA sequences) diverged. They have been used to help reveal new species or hidden evolutionary significant units (ESUs) in systematic biology studies (Behrooz *et al.*, 2018; Casacci *et al.*, 2014). They can greatly complement traditional morphological taxonomy studies by providing robust insights into population/species divergence histories, therefore allowing improved assessment of species proposals.

Here we examined the Asian rice frogs, a species complex generally known as *Fejervarya limnocharis* Gravenhorst, 1829. The species is distributed widely across central and southeastern China, southeast Asia and Japan (Zhong *et al.*, 2008; Djong *et al.*, 2011; Kotaki *et al.*, 2010). Interestingly, there is considerable morphological variation within the rice frog which has led to taxonomic debate (Matsui *et al.*, 2007; Veith, 2001). Using mtDNA control region (D-loop) sequences, analyses of this group revealed two distinct lineages within China: the Yangtze River lineage and the southern lineage (Zhong *et al.*, 2008).

Four species have been proposed within the rice frog complex on the basis of morphological or molecular divergence: (1) *F. kawamurai* Djong, Matsui, Kuramoto, Nishioka, Sumida, 2011; found on the Japan and possibly throughout northern and central China, and western Taiwan, China; (2) *F. sakishimensis* Matsui, Toda, and Ota, 2008; from the Iriomote and Ishigaki islands (southern Ryukyus) of Japan and eastern Taiwan, China (Djong *et al.*, 2011); (3) *F. multistriata* Hallowell, 1861; from southeast Asia and adjacent areas in southern China (Djong *et al.*, 2011; Fei *et al.*, 2002; Matsui *et al.*, 2007); (4) populations from the Orchard and Green islands in southeastern and eastern Taiwan, China seem to form a monophyletic group and could represent a distinct species (discussed by Djong *et al.*, 2011 who described them as “*Fejervarya* sp.”).

In addition to their taxonomic uncertainty, very few studies have addressed the distributions of the four recognized lineages (Zhong *et al.*, 2008). Within China, Zhong *et al.* (2008) investigated the mtDNA phylogeography of all Chinese populations, which they considered to be *F. multistriata*, following Fei *et al.* (2002). They did not obtain well-supported phylogenetic trees but detected two lineages: a Yangtze lineage (within the drainage system of the Yangtze River) and a southern China lineage (Zhong *et al.*, 2008). It has since become apparent that the Yangtze lineage could be grouped with central Chinese and Japanese populations (i.e., *F. kawamurai*) while the southern Chinese lineage corresponds to the south Asian *F. multistriata* (see discussion in Djong *et al.*, 2011). Nonetheless, there is still uncertainty about populations from unstudied areas, such as the Zhoushan Archipelago (and adjacent coastal areas in eastern China). These islands are of biogeographical interest

because they harbor differentiated lineages of several groups, e.g., the Zhoushan cobra (Lin *et al.*, 2008), the bat *Pipistrellus abramus* (Wei *et al.*, 2010) and the toad *Bufo gargarizans* (Tong *et al.*, 2017). It is also rather notable that three of the proposed species (*F. kawamurai*, *F. multistriata* and *Fejervarya* sp.) within the complex have been detected within the relatively small region of Taiwan, China, and its neighboring islands (Djong *et al.*, 2011). Hence, despite considerable previous work on the complex (e.g., Djong *et al.*, 2011; Matsui *et al.*, 2007; Zhong *et al.*, 2008), more systematics research is required to establish relationships within this group (Kotaki *et al.*, 2010).

In this study, we contributed a large amount of newly generated mtDNA and nuDNA sequence data from Chinese populations to the existing data from the rice frog complex. The aims were to: 1) examine the phylogeography of the Chinese populations within this complex, 2) assess the overall lineage diversity of these populations relative to other members of complex, 3) examine whether the multilocus data provided statistical support for delimitation of divergent groups into distinct species.

2. Materials and Methods

2.1. Sampling Genetic data were obtained from 270 individuals from sites in China (obtained between July 2016 and September 2017) which comprised (1) Zhoushan Archipelago: 119 individuals from 25 sites, (2) Hainan: 4 individuals from one site and (3) other regions of China: 147 individuals from 20 sites (see Supplementary Table S1 and Figure S1). Tissue samples (mostly muscle from toes) were stored in ethanol immediately after sampling, and subsequently maintained at 4°C. Fieldwork and tissue sampling authorization were performed in accordance with guidelines from the China Council on Animal Care and approved by the Ethics Committee of Animal Experiments at China Jiliang University.

DNA sequences from 9 additional specimens were retrieved from GenBank (from Kotaki *et al.*, 2010), and comprised rice frogs from the following areas: China, Indonesia and Japan, and labelled as *Fejervarya limnocharis* (also referred to as *Fejervarya* sp. by Djong *et al.* 2011) ($n=4$), *F. sakishimensis* ($n=2$), and *F. multistriata* ($n=2$) (see Supplementary Table S1 and Figure S1). Sequences from an outgroup taxon (*Fejervarya triora*) were also used for some analyses based on phylogenetic relationship within the *Fejervarya* genus (Kotaki *et al.*, 2010).

2.2. DNA extraction, amplification, and sequencing Total DNA was extracted from the 270 Chinese rice frog tissue samples with a universal genome extraction kit (Sangon Biotech (Shanghai) Co., Ltd.) following the manufacturers protocol. Two pairs of primer were designed for amplification of two mtDNA regions: i) *ND1-ND2* containing subunits one

and two of the NADH dehydrogenase (ubiquinone) complex and ii) *Cytb*-CR containing cytochrome b, tRNA and control region sequences (see Supplementary Table S2). A cytochrome b region situated within *Cytb*-CR was also available for the Genbank specimens. Eight modified pairs of primers that have been described previously (Bossuyt and Milinkovitch, 2000; Kotaki *et al.*, 2008, 2010) were used for amplification of the following eight nuclear loci: i) C-X-C Motif Chemokine Receptor 4 (*CXCR4*), ii) Tyrosinase (*Tyr*), iii) Rhodopsin (*Rhod*), iv) 3' portions of recombination activating 1 (*3RAG-1*), v) 5' portions of recombination activating 1 (*5RAG-1*), vi) Brain Derived Neurotrophic Factor (*BDNF*), vii) sodium-calcium exchanger (*NCX1*), viii) recombination activating 2 (*RAG-2*) (see Supplementary Table S2). These loci were amplified in 105 specimens selected from all localities to cover a wide geographical area and all mtDNA lineages. PCR products were sequenced by Sangon Biotech (Shanghai) Co., Ltd. These loci were selected as they were homologous with those in nine Genbank specimens (Supplementary Table S1). Sequencing chromatograms were spliced using SeqMam software in DNASTAR v.7.10 (Nystuen, 2001) and aligned using ClustalX (Thompson *et al.*, 1997). Nuclear sequences were phased using DnaSP v5 (Librado and Rozas, 2009) to allow phylogenetic analyses of alleles detected at each locus.

2.3. MtDNA analyses of Chinese rice frogs The mtDNA sequences (*Cytb*-CR and *ND1*-*ND2*, from 270 individuals, plus the outgroup *F. cancrivora* from Genbank EU652694) were used to investigate phylogeographic and demographic patterns within previously poorly sampled regions of the eastern China including Zhoushan Archipelago. Although nuclear loci were available for 105 specimens they showed relatively few variable sites within China and so were not used. The sequences were divided into the following five data partitions: 1st and 2nd (cp1 and 2) and 3rd (cp3) codon positions within the *ND1* and *ND2* gene codon sites, cp1 and 2 and cp3 within the *Cytb* gene, tRNA and control region (see Table S3). Gene trees were inferred using both the Bayesian approach implemented in MrBayes v.3.2.6 (Ronquist and Huelsenbeck, 2003) and the maximum likelihood (ML) approach in RAxML v.8.2.11 (Stamatakis, 2014).

The best DNA substitution model for each partition was assessed using MrModeltest v.2.3a (Nylander, 2004) and then implemented within MrBayes. Details on prior setting and substitution models are listed in Supplementary Table S3. Markov chain Monte Carlo (MCMC) analyses were run for 9 million generations, trees were sampled every 500 generations and the first 1800 sampled trees were discarded as burn-in.

The ML tree was obtained using the raxmlHPC-PTHREADS command in RAxML. Following the use of MrModeltest, the GTRGAMMA model of DNA substitution was the most suitable available model for each partition. A total of 1000

bootstrap replicates were generated.

Historical demographic changes within five identified Chinese mtDNA groups that contained over ten individuals for each group were analyzed using two approaches. First, the mismatch raggedness indices (Excoffier and Schneider, 1999), Tajima's *D* (Tajima, 1989) and Fu's *F_s* values (Fu, 1997) were obtained using 1000 simulations (Program: Arlequin v.3.5) (Excoffier and Lischer, 2010). Second, coalescent Bayesian skyline plots (BSPs) were generated using BEAST v1.8.2. A substitution rate of 0.957% per lineage per million years was used for this analysis based on an estimated divergence date of 10Ma for the *ND2* divergence between European and Asian *Bufo* toads (from Macey *et al.*, 1998). This rate appeared reasonable given the relative homogeneity of *ND2* rates across reptile and amphibian lineages (Macey *et al.*, 2001). Also, for greater reliability, the analyses used only sites that were homologous with the *ND2* fragment analysed by Macey *et al.* (1998).

2.4. Species delimitation analyses The general aim was to use the 279 nuclear and mtDNA sequences for multispecies coalescent species delimitation analyses (sequences are listed in Supplementary Table S1). These analyses examined levels of genetic support for the four previously-proposed rice frog species within the complex (Djong *et al.*, 2011; Kotaki *et al.*, 2010; Matsui *et al.*, 2007; Zhong *et al.*, 2008) although one of these was also separated into two candidate species. We therefore examined support for a total of five candidate species (described below) using the eight nuclear loci plus the mtDNA locus.

Prior to these species tree analyses, we assessed mitonuclear discordance through analyses of nuclear and mtDNA gene trees. The same sequences (collapsed into haplotypes) that were used in the species delimitation analyses described below were used although the outgroup (*Fejervarya triora*: Supplementary Table S1) was also included. ML trees were estimated using haplotypes identified for: i) all eight phased nuclear loci (unpartitioned), ii) homologous mtDNA (*Cytb*, partition by codon position) sequence present in new specimens (note that we only used *Cytb* sequences because they were homologous with sequences used in previous studies that we wished to include). In both cases, partitioning was guided by PartitionFinder (Lanfear *et al.*, 2017). The raxmlHPC-PTHREADS command within RAxML (Stamatakis, 2014) was used, with the GTRGAMMA substitution model. Statistical support for nodes was derived from 1000 bootstrap replicates.

The five specified candidate species that were tested in the multispecies coalescent analyses were mainly based on previously proposed species and the same names are used wherever possible (see Introduction for general details on these species). The candidate species were: i) rice frogs from the Iriomote and Ishigaki islands of Japan, and those

from northeastern Taiwan, China (here, all described as “*F. sakishimensis*”), ii) rice frogs from Orchard and Green islands of southeastern Taiwan, China (*Fejervarya* sp.), iii) rice frogs from Thailand, Cambodia, Malaysia, Indonesia and adjacent areas in southern China (here, *F. multistriata*), iv) populations from the Hiroshima, Japan, and the Chinese Zhoushan Archipelago and some peripheral *F. kawamurai* areas of south-central eastern China (such as Guangzhou and Guangxi), which includes populations that seem to have been assigned to *F. multistriata* (Djong *et al.*, 2011) (here; *F. kawamurai*-ZS), v) north-central China populations, formerly described as *F. kawamurai* (here; *F. kawamurai*-NCC). The justification for splitting *F. kawamurai* into two candidate species (*F. kawamurai*-ZS and *F. kawamurai*-NCC) followed from the initial mtDNA analyses of Chinese populations which showed a major/deep well-supported basal split (see later). Rather than ignore this split, we examined more general genetic support for it using all nine loci in our species tree analyses.

Species trees were analyzed using the Bayesian multispecies coalescent approaches implemented within the programs BPP v4.30 (Yang, 2015) and STACEY v. 1.2.5 (Jones, 2017). All nine loci were used, i.e., the mtDNA (*Cytb*) and eight nuclear loci. The analyses used all new specimens sequenced in this study, for which all loci were available ($n=105$), together with sequences (from eight individuals) that originated from Kotaki *et al.* (2010). BPP and STACEY allow assessment of support for different numbers of species, and the historical relationships among them, but implement quite different algorithms. Both programs allow specified candidate species to be lumped with one another but not split into additional species.

BPP analyses (using the A11 configuration) provided posterior probabilities for different numbers of species and tree topologies. Unlike STACEY, it implements a reversible jump algorithm that allows the MCMC chain to move between species trees with any number of tips, up to the maximum number of candidate species proposed, which leads to changes in the parameterization of the species tree. BPP requires specification of priors on population sizes and divergence times. Inverse-gamma (INVG) priors were used to accommodate scenarios that implied different population sizes (θ) and divergence times (τ). Preliminary estimates suggested the most appropriate prior specification was: θ INVG (1.2, 0.006), τ INVG (1.2, 0.0005) and this was run in triplicate (from different starting positions) to ensure convergence on the same posterior (50000 post burn-in samples taken from a chain run for 2.5 million steps). Additional analyses with different priors were used to ensure that the results were not overly-sensitive to the population size and divergence time priors, as follows: i) smaller population size with more recent divergence time, θ INVG (1.2, 0.0006), τ INVG (1.2, 0.00005); ii) larger population size with

longer divergence time, θ INVG (1.2, 0.06), τ INVG (1.2, 0.005); iii) larger population size but more recent divergence time, θ INVG (1.2, 0.06), τ INVG (1.2, 0.00005); iv) smaller population size with longer divergence time, θ INVG (1.2, 0.0006), τ INVG (1.2, 0.005).

Equivalent analyses were carried out using STACEY combined with the associated SPECIESDELIMITATIONANALYSER (Jones, 2017). STACEY was implemented within BEAST ver2.6 (Bouckaert *et al.*, 2014). A primary difference from BPP is that the problem of different numbers of parameters arising under different species delimitations is overcome using a particular birth-death model for the species tree that allows effective collapse of node heights close to zero. These collapsed nodes continue to exist, leaving the parameter space unchanged despite effective changes to the number of tip species. Another difference is that population size parameters are constant along each branch and integrated out, rather than being sampled. These aspects, together with newly-developed operators seem to allow good convergence, particularly when compared with *BEAST species tree analyses (Jones, 2017). Here, the tree prior was specified from a Birth-Death-Collapse model, with a collapse height of 1×10^{-4} (which specified the upper limit at which the node height is considered to be zero) the difference between birth and death rate, $\mu - \lambda = 100$, and the death rate relate to birth rate, $\mu/\lambda = 0.5$. The Jukes Cantor substitution model (as implemented in BPP) was used for all loci, except mtDNA where the HKY model was used to accommodate greater sequence divergence. The strict clock model was used for all loci, with priors specified from an inverse gamma distribution with parameters $\alpha=10$, $\beta=1$. Collapse weight for species was specified as uniform using a Beta distribution, with both parameters set to 1, giving equal weight to any number of species between 1 and 5. Finally the popPriorScale prior was specified from a lognormal distribution with mean of -7 and standard deviation of 2. The MCMC was run for 2.5×10^8 steps and sampled every 10^4 steps (three replicates).

3. Results

3.1. Sequencing MtDNA and nuclear sequences were obtained for all 270 Chinese specimens. The total aligned lengths/numbers of variable positions for individual loci were as follows: 2596/392 bp for mtDNA (*ND1-ND2* 869/97 bp, *Cytb-CR* 1729/295 bp) and for nDNA: 731/32 bp (*3RAG-1*), 877/64 bp (*5RAG-1*), 441/7 bp (*BDNF*), 651/19 bp (*CXCR4*), 770/9 bp (*NCX1*), 869/41 bp (*RAG-2*), 290bp/1 (*Rhod*) and 661/43 bp (*Tyr*).

3.2. MtDNA history of Chinese rice frogs The long mtDNA fragments that were sequenced for the 270 Chinese specimens provided the resolution necessary to detect phylogeographic and demographic patterns within this region. The mtDNA trees revealed two major well-supported subgroups corresponding to

F. kawamurai.

The first of these included four well-supported groups from the following regions in China: Guangxi (A); Guangzhou (B); Pinghu, Zhejiang (ZJPH) (C); the Zhoushan Archipelago (the majority of samples from this area) and the adjacent Hangzhou Bay area, Zhejiang (two samples) (D) (Figure 1). The second major mtDNA subgroup consisted of eight major subgroups (of which seven were well-supported by posterior probabilities >0.98 from the Bayesian analysis but only two were supported by bootstrap values >70% from the ML analysis) from the following regions in China: Hainan (E); Zhejiang province in eastern China along with some samples from the Zhoushan Archipelago (Jintang Island and the Dongji islands) that did not overlap with subgroup D (F); coastal areas of Zhejiang province and Jintang Island along with two individuals from Henan province, west of Zhejiang (G); Henan and Anui Provinces, and some western areas of Zhejiang province (H); Hunan province (south of Zhejiang and Henan) (I); Guizhou (east of Hunan) (J); Henan (Xinyang) (K); Zhejiang and Shanghai Lucao, Anhui, Jiangxi and Hunan, which are close to Zhejiang, along with an individual from Henan (Tanghe) (L) (Figures 1 and S2). Groups, F, G and H showed overlapping distributions in Zhejiang

province.

Non-coalescent demographic analyses were only carried out on the more diverse mtDNA lineages, B, D, G, H and L, within which the sequence data contained a reasonable number of informative sites. This indicated rapid recent expansion within three of these: D, G and L (see Table 1 for details). The coalescent skyline test also suggested recent population demographic expansions 25–35 ka ago in D and L and a slightly more recent expansion in G (Figure 2).

3.3. Species delimitation analyses Preliminary analyses of gene trees for the individual loci used for the species tree provided generally low levels of support (bootstrap value <50 for most nodes) and did not support general phylogenetic discordance between nuclear and mtDNA markers (Figures S3–S4). Concatenation of rice frog nuclear sequences revealed 92 haplotypes (excluding the *F. triora* outgroup) (Table S4). The nuclear tree using concatenated sequences was poorly resolved with low bootstrap support on the nodes (Figure S3). Only two major nodes had support over 70%: support for the ancestral node for *F. sakishimensis* and *Fejervarya* sp. was 96%, while a large group of sequences with low divergence, comprising all *F. kawamurai*-ZS and the majority of the *F. kawamurai*-NCC

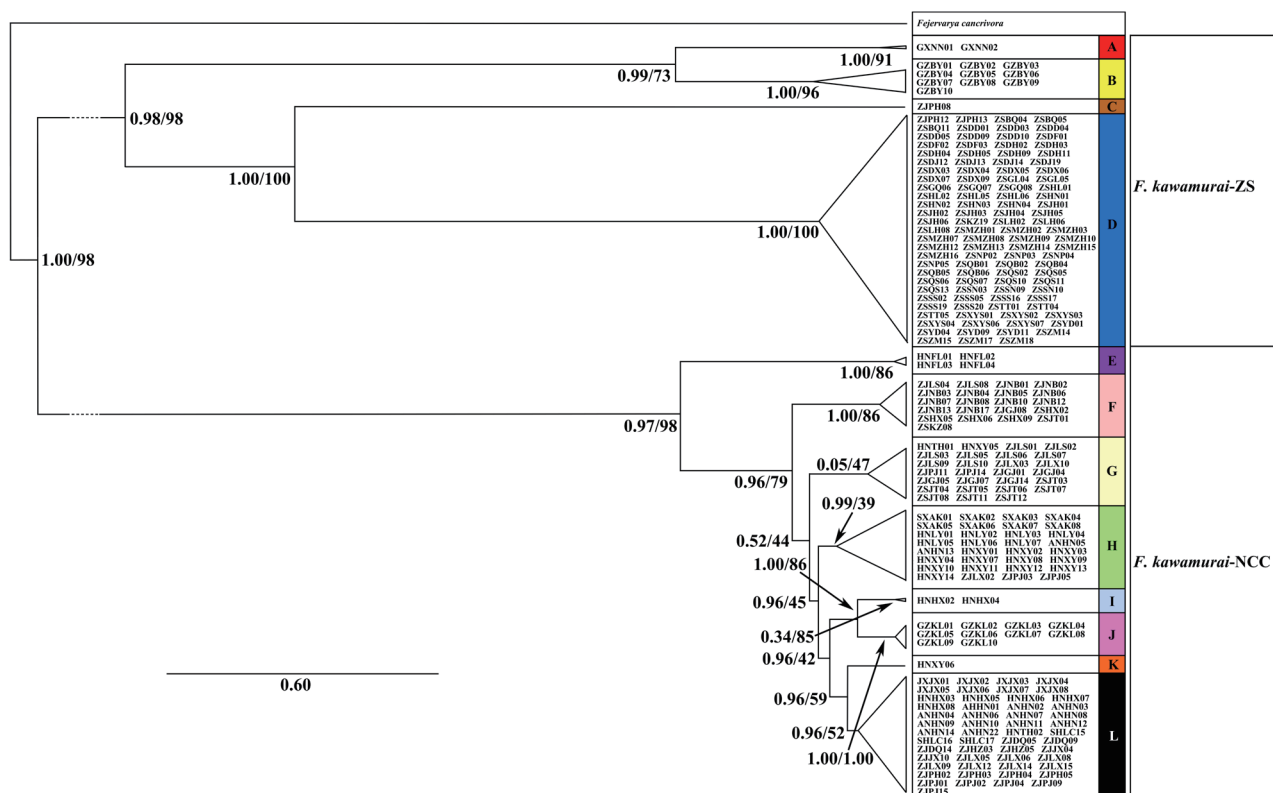


Figure 1 Phylogenetic trees of Chinese *Fejervarya limnocharis* from partitioned analyses using mtDNA haplotypes (*Cytb*-CR and *ND1*-ND2). Specimens found within each mtDNA group (A-L) are labelled and correspond to the full list provided in Table S1. Values on branches represent MrBayes posterior probabilities/RAxML bootstrap support values, respectively.

Table 1 Summary statistics for Chinese mitochondrial clades. Haplotype diversity, nucleotide diversity, F_s , Tajima's D and raggedness indices were estimated from *ND1-ND2* sequences.

Clades	<i>n</i>	Unique haplotypes	Haplotype diversity	Nucleotide diversity	F_s	Tajima's D	Raggedness index (<i>P</i> value)	Tau
B	10	6	0.91 ± 0.06	0.4488 ± 0.3046	-1.53 (0.132)	1.08 (0.79)	0.05 (0.57)	1.54
D	107	13	0.73 ± 0.02	0.1212 ± 0.0805	-4.20 (0.046)	-0.97 (0.12)	0.03 (0.65)	2.02
G	27	3	0.14 ± 0.08	0.0740 ± 0.1208	-2.22 (0.008)	-1.51 (0.02)	0.52 (0.95)	0.1
H	32	4	0.59 ± 0.04	0.2305 ± 0.1996	-0.41 (0.378)	-0.16 (0.35)	0.19 (0.28)	0.31
L	52	9	0.59 ± 0.07	0.1068 ± 0.0890	-4.28 (0.005)	-1.37 (0.05)	0.12 (0.06)	0.94

Values in bold support significant population expansion.

haplotypes had 72% bootstrap support (Figure S3). MtDNA sequences were collapsed into 26 ingroup haplotypes and the ML tree showed generally higher node support than the nuclear tree (Table S5 and Figure S4). The net uncorrected cytochrome b divergences between the five candidate species are provided in 2. In the mtDNA tree (Figure S4), *Fejervarya* sp. (also referred to as *F. limnocharis* from Orchard Island in Djong *et al.*, 2011) mtDNA haplotypes formed an outgroup to all other rice frog haplotypes, with *F. sakishimensis* haplotypes also forming an outgroup relative to *F. kawamurai* and *F. multistriata* haplotypes. Remaining haplotypes fall into two groups with bootstrap support for the respective ancestral nodes of 96% and 82%, although the node ancestral to both these groups is only weakly supported (40%). The first of these groups comprises eleven haplotypes and includes *F. kawamurai*-NCC and all *F. multistriata* individuals, with *F. multistriata* originating from the more basal nodes within the group. There is a tendency for Hainan haplotypes to group with these *F. multistriata* haplotypes, reflecting a weak pattern that is also observed (but with low support) in the concatenated nuclear tree. The second of these two groups contains twelve haplotypes, which include all remaining *F. kawamurai*-NCC and all *F. kawamurai*-ZS haplotypes.

The BPP species tree analysis revealed strong statistical support for the five candidate species that we tested within the rice frog complex, independent of different priors placed on population size and divergence times. Posterior probability (PP) support for a five species delimitation was PP=0.966 for our preferred analysis with θ specified as INVG (1.2, 0.006), and τ specified as INVG (1.2, 0.0005). We favored this analysis in part because prior and posterior means were very similar, both for θ (0.030 and 0.032, respectively) and τ (0.0025 and 0.0029, respectively) (this was not the case for the other exploratory analyses of the sensitivity of the analysis to these priors). The preferred specification provided very high individual delimitation PPs for each of the five proposed species (PP>0.990 for *F. kawamurai*-ZS, *F. kawamurai*-NCC and *F. multistriata*; PP>0.967 for *F. sakishimensis* and *Fejervarya* sp.). A five species delimitation had greatest support (PP=0.967)

in both above analyses and other analyses specifying small population sizes, although we note that overall support for five species was diminished when larger population sizes were specified (PP=0.434 for longer divergence times; PP=0.555 for shorter divergence times). The most strongly favored species tree topology in all analyses presented a ((*F. kawamurai*-ZS, *F. kawamurai*-NCC), *F. multistriata*) lineage that was reciprocally monophyletic with a (*F. sakishimensis*, *Fejervarya* sp.) lineage (Figure 3). The only exception to this finding was the aforementioned long divergence time/large population size analysis, in which *F. sakishimensis* and *Fejervarya* sp. were grouped as one species. Posterior support for the (((*F. kawamurai*-ZS, *F. kawamurai*-NCC), *F. multistriata*), (*F. sakishimensis*, *Fejervarya* sp)) topology was PP=0.750 in our preferred analysis with intermediate priors.

The STACEY analyses also revealed greatest posterior support (PP=0.954) for the five species delimitation. Four species delimitations received much lower support, with a four species delimitation in which *F. sakishimensis* and *Fejervarya* sp. were joined as a single cluster being the most strongly favoured of these, but with a PP of only 0.046. The maximum clade credibility tree revealed an identical species tree topology to that favored by BPP. Highest PPs (1.000) were obtained for two nodes: i) the most basal node, representing the split between *F. multistriata* and the remaining species, and ii) the (*F. kawamurai*-ZS, *F. kawamurai*-NCC) node. Support for the most recent common ancestor (MRCA) of *F. sakishimensis*, *Fejervarya* sp. was PP=0.910. The node with lowest posterior support (PP=0.548) was that representing the MRCA of *F. kawamurai*-ZS, *F. kawamurai*-NCC and *F. multistriata*.

4. Discussion

The analyses presented here throw new light on phylogeographical relationships among rice frog populations in China. Our species tree analyses suggest greater genetic diversity within the *F. limnocharis* species complex than previously described. In particular, our sampling across China allowed detection of a deep mtDNA split within the previously

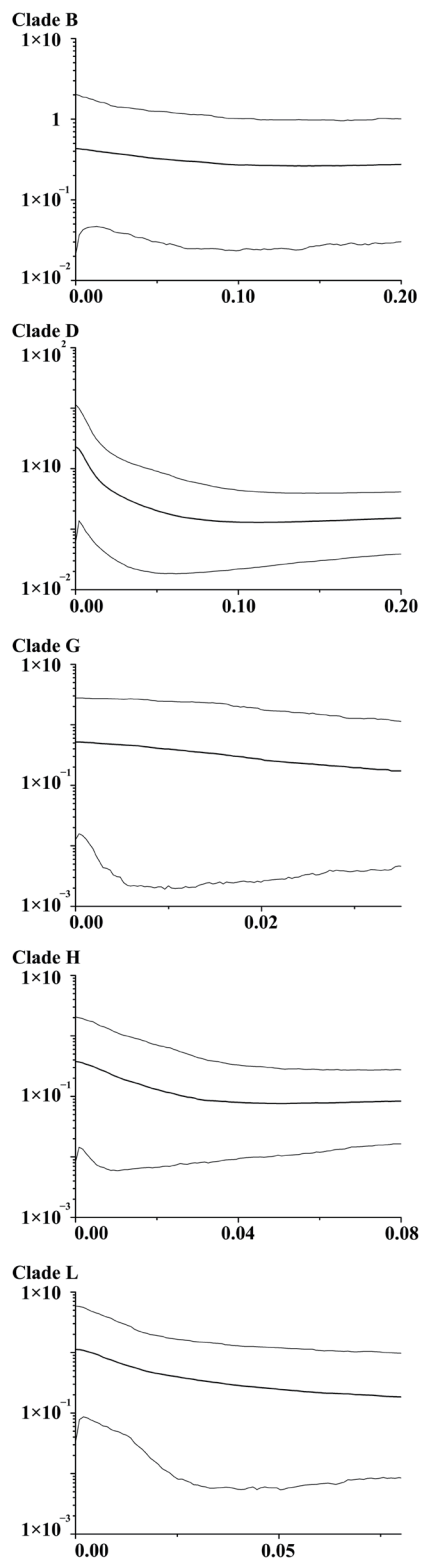


Figure 2 Bayesian skyline plots showing the population growth history of *F. limnocharis*. Horizontal axis corresponds to time (Ma), and vertical axis represents the product of female effective population size (N_e) and generation time (τ , years $\times 10^{-6}$).

proposed species *F. kawamurai* (Djong *et al.*, 2011) and, based on eight nuclear loci and the mtDNA locus, provided initial support for its future delimitation into two new sister species. One of these corresponds to Japan, the Zhoushan Archipelago and some regions close to the coasts of eastern and southern China (here denoted *F. kawamurai*-ZS), and the other is found in the centre and eastern coast of China, including two of the Zhoushan Archipelago (here denoted *F. kawamurai*-NCC).

Genetic data also provided evidence that supported full species recognition of the previously-described *F. sakishimensis* from the Japanese Iriomote and Ishigaki islands, close to Taiwan, China (Djong *et al.* 2011). Furthermore, this group was found to be most closely related to, but distinct from, another proposed species that Djong *et al.* (2011) referred to as "*Fejervarya* sp." from the Orchid and Green islands to the south-east of Taiwan, China. Finally, we find some evidence (albeit weaker) that most Chinese mainland (and Japanese) populations are more closely related to other Asian populations than they are to populations from islands in the East China Sea. Our findings are further supported by the finding of very substantial net pairwise mtDNA divergences of around 7.3% and 9.8% between most pairs of candidate species, with the exception of *F. multistriata*/*F. kawamurai*-NCC (see below).

Mitochondrial control region sequences previously showed that Chinese mainland populations of *F. limnocharis* were split into northern Yangtze River basin and southern groups (Zhong *et al.*, 2008). However, this study does not appear to have included sampled sites within *F. kawamurai*-ZS, which might explain why it detected just two major mtDNA groups in China (the groups we refer to as *F. kawamurai*-NCC and *F. multistriata*). Wider sampling by Zhong *et al.* (2008) would have revealed that their southern mtDNA group also included populations from the Zhoushan Archipelago, Zhejiang, China and Japan.

We are also able to substantially advance the findings of Kotaki *et al.* (2010) who provided an impressively wide sample distribution from across Asia, but did not sample any other sites except Husa (adjacent to southeast Asia) within China. This likely explains why they did not detect the *F. kawamurai*-NCC and *F. kawamurai*-ZS groups. These two sister groups are separated by a net mtDNA divergence of 10% (Table 2), which is greater than expected within most vertebrate species. It is therefore surprising that the non-sister lineages *F. kawamurai*-NCC and *F. multistriata* are separated by only 1% mtDNA divergence (Table 2). This relationship within the species tree appears to be strongly driven by the nuclear loci: *F. kawamurai*-NCC and *F. multistriata* actually show some of the greatest general nuclear divergence between candidate species pairs, similar to that between *F. multistriata* and *F. sakishimensis*. This cytonuclear discordance may have originated from mtDNA

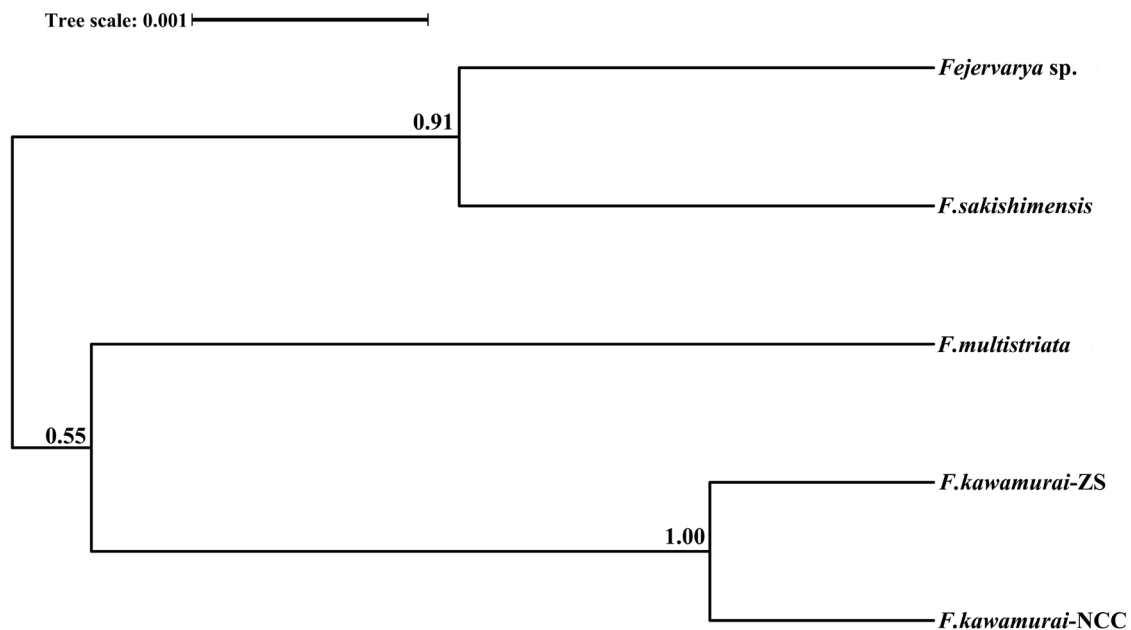


Figure 3 BPP species tree showing posterior support for nodes. The five species delimitation was supported with a posterior probability of PP=0.967.

Table 2 Net mtDNA divergence (cytochrome b) between candidate species (lower diagonal) with associated bootstrap standard error estimates (upper diagonal).

Groups	<i>F. multistriata</i>	<i>F. kawamurai</i> -NCC	<i>F. kawamurai</i> -ZS	<i>F. sakishimensis</i>	<i>Fejervarya</i> sp.
<i>F. multistriata</i>		0.005	0.015	0.015	0.015
<i>F. kawamurai</i> -NCC	0.010		0.016	0.017	0.016
<i>F. kawamurai</i> -ZS	0.089	0.098		0.014	0.014
<i>F. sakishimensis</i>	0.085	0.091	0.079		0.015
<i>Fejervarya</i> sp.	0.083	0.087	0.088	0.073	

introgression between *F. kawamurai*-NCC and *F. multistriata*. This is one of several questions for future studies that are raised here. Future genomic and morphological analyses that examine the different major groups that we have identified in greater depth analyses will provide more definitive conclusions.

Demographic coalescent analyses suggest quite rapid increases in population size within mitochondrial groups from i) neighboring Zhejiang, Jiangxi, Shanghai and Anhui regions close to the east coast of Chinese mainland and neighboring regions and ii) the adjacent Zhoushan Archipelago beginning around 25 000–35 000 ka. There were several late Pleistocene/early Holocene marine transgressions that might have affected the demographics of populations in this area. For example, the Ge Hu marine transgression (starting ~40 ka ago) which flooded the area was followed by a drop in temperature, when sea-levels may have reached an all-time low (Lin *et al.*, 2012; Wang *et al.*, 1980; Zhang *et al.*, 2008). Large areas of marshland

subsequently appeared in what is now the East China Sea and this likely provided opportunities for population expansion for these Zhejiang and Zhoushan groups. Other studies have related marine regression events to the rapid expansion of *Pelophylax* frogs (Dai, 2009) and *Bufo gargarizans* (Tong *et al.*, 2017) in the same region.

5. Conclusion

We have added to knowledge of the biodiversity within the *F. limnocharis* species complex by detecting two lineages within the Chinese mainland. These are found within *F. kawamurai*, although they included some individuals that would previously have been considered to be within the area occupied by *F. multistriata*. Overall the species delimitation approach favours five species within the complex, four of which (including “*Fejervarya* sp.”) have been proposed by previous studies

(including two morphology studies). Areas of uncertainty clearly remain, which could be clarified by future studies. More complete sampling of the entire range of the rice frog and analysis of a larger proportion of the nuclear genome will provide more detailed insights into relationships and introgression between Rice frog groups. Nonetheless, this study establishes a new framework for understanding the population history of the rice frog and may help establish a stable taxonomy within this widely-distributed species complex.

Acknowledgments This work was supported by the Natural Science Foundation of Zhejiang Province, China (LY21C040002), and partially by the National Natural Science Foundation of China (41541002). We wish to thank Guowo DING for local specimen contribution.

References

- Behrooz R., Kaboli M., Arnal V., Nazarizadeh M., Asadi A., Salmanian A., Montgelard C. 2018. Conservation below the species level: Suitable evolutionarily significant units among mountain vipers (the *Montivipera raddei* complex) in Iran. *J Hered*, 109(4): 416–425
- Bossuyt F., Milinkovitch M. C. 2000. Convergent adaptive radiations in Madagascan and Asian ranid frogs reveal covariation between larval and adult traits. *Proc Natl Acad Sci USA*, 97: 6585–6590
- Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C. H., Xie D., Drummond A. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*, 10(4): e1003537
- Casacci L. P., Barbero F., Balletto E. J. 2014. The “Evolutionarily Significant Unit” concept and its applicability in biological conservation. *Ital J Zool*, 81(2): 182–193
- Dai J. H. 2009. Phylogeography of *Pelophylax plancyi* species complex of China. Ph.D. Thesis. Nanjing Normal University, Nanjing (In Chinese with English abstract)
- Djong H. T., Matsui M., Kuramoto M., Nishioka M., Sumida M. 2011. A new species of the *Fejervarya limnocharis* complex from Japan (Anura, Dicroglossidae). *Zool Sci*, 28(12): 922–929
- Excoffier L., Lischer H. E. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res*, 10(3): 564–567
- Excoffier L., Schneider S. 1999. Why hunter-gatherer populations do not show signs of pleistocene demographic expansions. *Proc Natl Acad Sci USA*, 96(19): 10597–10602
- Favre A., Päckert M., Pauls S. U., Jähmig S. C., Uhl D., Michalak I., Muellner-Riehl A. N. 2015. The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biol Rev Camb Philos Soc*, 90(1): 236–253
- Fei L., Ye C., Jiang J., Xie F. 2002. On taxonomic status of *Rana limnocharis* group with revision of nomenclature of the Rice Frog from China. *Herpetologica Sinica*, 9: 88–96
- Fu Y. X. 1997. Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147: 915–925
- Gravenhorst J. L. C. 1829. *Deliciae Musei Zoologici Vratislaviensis. Fasciculus primus. Chelonios et Batrachia*. Leipzig: Leopold Voss
- Heled J., Drummond A. J. 2010. Bayesian inference of species trees from multilocus data. *Mol Biol Evol*, 27(3): 570–580
- Jones G. 2017. Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *J Math Biol*, 74(1): 447–467
- Kotaki M., Kurabayashi A., Matsui M., Khonsue W., Djong T. H., Tandon M., Sumida M. 2008. Genetic divergences and phylogenetic relationships among the *Fejervarya limnocharis* complex in Thailand and neighboring countries revealed by mitochondrial and nuclear genes. *Zool Sci*, 25: 381–390
- Kotaki M., Kurabayashi A., Matsui M., Kuramoto M., Djong, Sumida M. 2010. Molecular phylogeny of the diversified frogs of genus *Fejervarya* (Anura: Dicroglossidae). *Zool Sci*, 27(5): 386–395
- Lanfear R., Frandsen P. B., Wright A. M., Senfeld T., Calcott B. 2017. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol*, 34(3): 772–773
- Leaché A. D., Zhu T., Rannala B., Yang Z. 2019. The spectre of too many species. *Syst Biol*, 68(1): 168–181
- Librado P., Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25(11): 1451–1452
- Lin J. X., Dai L. P., Wang Y., Liu M. 2012. Quaternary marine transgressions in eastern China. *J Palaeogeogr*, 1(2): 105–125
- Lin L. H., Zhao Q., Ji X. 2008. Conservation genetics of the Chinese cobra (*Naja atra*) investigated with mitochondrial DNA sequences. *Zool Sci*, 25(9): 888–893
- Liu L., Pearl D. K. 2007. Species trees from gene trees: Reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Syst Biol*, 56(3): 504–514
- Macey J. R., Li J. A. S., Larson A., Fang Z., Wang Y., Tuniyev B. S., Papenfuss T. J. 1998. Phylogenetic relationships of toads in the *Bufo bufo* species group from the eastern escarpment of the Tibetan Plateau: A case of vicariance and dispersal. *Mol Phylogenet Evol*, 9(1): 80–87
- Macey J. R., Strasburg J. L., Brisson J. A., Vrendenburg V. T., Jennings M., Larson A. 2001. Molecular phylogenetics of western North American frogs of the *Rana boylei* species group. *Mol Phylogenet Evol*, 19(1): 131–143
- Matsui M., Toda M., Ota H. J. C. H. 2007. A new species of frog allied to *Fejervarya limnocharis* from the southern Ryukyus, Japan (Amphibia: Ranidae). *Curr Herpetol*, 26(2): 65–79
- Muellner-Riehl A. N. 2019. Mountains as evolutionary arenas: Patterns, emerging approaches, paradigm shifts, and their implications for plant phylogeographic research in the Tibeto-Himalayan Region. *Front Plant Sci*, 10: 195
- Ni G., Li Q., Kong L., Yu H. 2014. Comparative phylogeography in marginal seas of the northwestern Pacific. *Mol Ecol*, 23(3): 534–548
- Nylander J. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University
- Nystuen A. 2001. Lasergene 5.0. 1. Biotech Software Intern Rep, 2(6): 239–244
- Ronquist F., Huelsenbeck J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12): 1572
- Stamatakis A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9): 1312–1313
- Tajima F. 1989. The effect of change in population size on DNA polymorphism. *Genetics* 123: 597–601
- Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., Higgins D. G. 1997. The CLUSTALX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25(24): 4876–4882

- Tong H. J., Wo Y. B., Liao P. H., Jin Y. T. 2017. Phylogenetic, demographic and dating analyses of *Bufo gargarizans* populations from the Zhoushan Archipelago and Chinese mainland. *Asian Herpetol J*, 8(3):165–173
- Veith M., Kosuch J., Ohler A., Dubois A. 2001. Systematics of *Fejervarya limnocharis* (Gravenhorst, 1829) (Amphibia, Anura, Ranidae) and related species. 2. Morphological and molecular variation in frogs from the Greater Sunda islands (Sumatra, Java, Borneo) with the definition of two species. *Alytes*, 19(1): 5–28
- Wang J. T., Wang P. X. 1980. Relationship between sea-level changes and climatic fluctuation in East China since late Pleistocene. *Acta Geogr Sin*, 35(4): 299–312 (In Chinese with English abstract)
- Wei L. L., Jonr F., Stephen J. R., Cassandram M., Zhang L., Zhang S. 2010. Phylogeography of the Japanese pipistrelle bat, *Pipistrellus abramus*, in China: The impact of ancient and recent events on population genetic structure. *Biol J Lin Soc*, 99(3): 582–594
- Wen J., Zhang J. Q., Nie Z. L., Zhong Y., Sun H. 2014. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front Genet*, 5: 4
- Yang Z. 2015. The BPP program for species tree estimation and species delimitation. *Current Zool*, 61(5): 854–865
- Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci USA*, 107(20): 9264–9269
- Zhang J. Q., Tang L. L., Zou H. 2008. The response to the variety of paleoclimate and sea level in the east China sea after the late Pleistocene. *Trans Oceanol Limnol*, 1: 26–31 (In Chinese with English abstract)
- Zhong J., Liu Z. Q., Wang Y. Q. 2008. Phylogeography of the rice frog, *Fejervarya multistriata* (Anura: Ranidae), from China based on mtDNA D-loop sequences. *Zool Sci*, 25(8): 811–820

Handling Editor: Heling Zhao***How to cite this article:***

Yang K., Wo Y. B., Shao G., Liao P. H., Tong H. J., Brown R. P., Jin Y. T. Phylogenetic Relationships among Chinese Rice Frogs within the *Fejervarya limnocharis* Species Complex (Amphibia: Dicroglossidae). *Asian Herpetol Res*, 2022, 13(4): 232–241. DOI: 10.16373/j.cnki.ahr.210050

Appendix

The Tables S1–S5 and Figures S1–S2 can be downloaded from the website https://pan.baidu.com/s/1DOWuYN7qMFeSiDvCjme_QQ?pwd=1407 (access code: 1407).

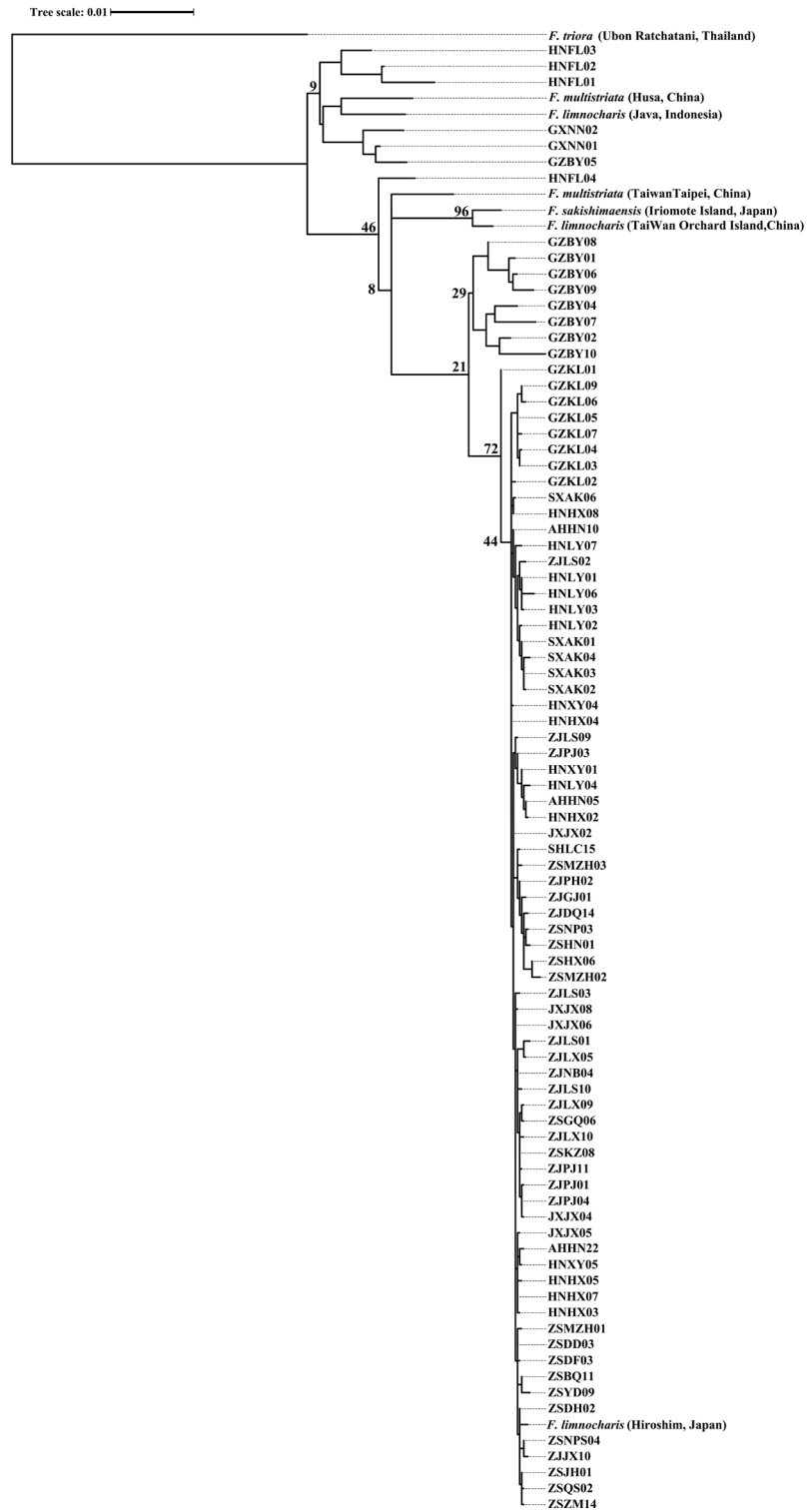


Figure S3 Cladogram of *F. limnocharis* from RAXML based on nuclear DNA. Concatenation of rice frog nuclear sequences revealed 92 haplotypes (excluding the *F. tritora* outgroup). The values labelled at the tree nodes show the RAXML bootstrap values.



Figure S4 Cladogram of *F. limnocharis* from RAxML based on MtDNA sequences. MtDNA sequences were collapsed into 26 ingroup haplotypes. The values labelled at the tree nodes show the RAxML bootstrap values.