Zoologica Scripta

KUNGL. VETENSKAPS-AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes

Qiong-Ying Tang, Si-Qing Liu, Dan Yu, Huan-Zhang Liu & Patrick D. Danley

Submitted: 26 June 2011
Accepted: 16 December 2011
doi:10.1111/j.1463-6409.2011.00530.x

Tang, Q.-Y., Liu, S.-Q., Yu, D., Liu, H.-Z. & Danley, P.D. (2012) Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. —*Zoologica Scripta*, *41*, 233–247.

Understanding the diversification of species is a central goal of evolutionary biological studies. One powerful tool to investigate the speciation process is molecular systematics. Here, we use molecular methods to investigate the evolution of balitorine loaches belonging to two genera, *Lepturichthys* and *Jinshaia*. Both genera contain only two species (*Lepturichthys fimbriata*, *Lepturichthys dolichopterus* and *Jinshaia sinensis* and *Jinshaia abbreviata*), all of which are endemic to China. These species share many morphological and ecological characters and exhibit overlapping distributions in the Upper Yangtze River. In this study, we used two mitochondrial genes (Cyt*b* and COI) and one nuclear gene (RAG1) to investigate the phylogenetic relationships within and between these two genera. Phylogenetic analyses and network construction based on mitochondrial and nuclear genes consistently supported the monophyly of *Jinshaia*. In contrast, the mitochondrial and nuclear genes yielded conflicting results in *Lepturichthys*. The phylogenetic analyses of mitochondrial sequences identify two distinct *Lepturichthys* lineages, *Lepturichthys* A and *Lepturichthys* B. *Lepturichthys* A includes most of *L. fimbriata* individuals from the Upper Yangtze River and is the sister group to all *Jinshaia* species. *Lepturichthys* B consists of the remaining *L. fimbriata* individuals from the Upper and Middle Yangtze River, and all *L. dolichopterus* individuals from the Minjiang River in Southeastern China. However, the analysis of the nuclear sequence indicates that the genus *Lepturichthys* is monophyletic and is only distantly related to *Jinshaia*. This incongruence suggests that introgressive hybridization might have occurred between *L. fimbriata* (*Lepturichthys* A) and *Jinshaia* species. As a result of this hybridization event, *L. fimbriata* captured the mitochondrial genome of the sympatric *Jinshaia* species. This capture event appears to have occurred at least 1.74 million years ago. Additionally, *L. fimbriata* appears to be paraphyletic; the nuclear data indicated that *L. dolichopterus* forms a monophyletic clade nested within *L. fimbriata*. Because *L. dolichopterus* and *L. fimbriata* are allopatric and hybridization may not be possible, we suggest that the observed paraphyly of *L. fimbriata* is a product of incomplete lineage sorting. In addition, the reciprocal monophyly of *J. sinensis* and *J. abbreviata* could not be resolved. This may be the result of interspecific hybridization as these species occur sympatrically. However, incomplete lineage sorting may have caused the observed topology of the *Jinshaia* species. The data presented here illustrate the complex evolutionary history of the balitorine loach species: intergeneric hybridization and interspecific hybridization have likely occurred in this lineage. In addition, possible incomplete lineage sorting may further obscure the evolutionary history of this group. The complex relationships of the balitorine loaches provide a rich evolutionary system to study the creation of sympatric and sister species.

Corresponding author: *Huan-Zhang Liu, The Key Lab of Aquatic Biodiversity and Conservation, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, Hubei, China. E-mail: hzliu@ihb.ac.cn*

Qiong-Ying Tang, Si-Qing Liu, Dan Yu, The Key Lab of Aquatic Biodiversity and Conservation, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, Hubei, China. Emails: tangqy@ihb.ac.cn, yudan@ihb.ac.cn, liusiqing310@gmail.com

Patrick D. Danley, Department of Biology, Baylor University, One Bear Place #97388, Waco, TX 76798, USA. Email: patrick_danley@baylor.edu

Introduction

Understanding the origin of organismal diversity is a central theme of evolutionary biology. The rise of modern molecular techniques in the areas of systematics, population biology and biogeography has made the investigation of evolutionary processes ever more practical. Over the past decades, mitochondrial genes have been widely used in phylogenetic and phylogeographic analyses. The mitochondrial DNA provides a useful phylogenetic marker because it does not recombine, it has a high mutation rate, and universal primers have been available for decades (Kocher *et al.* 1989; Meyer 1993; Xiao & Zhang 2000). Recently, however, a number of studies have shown that mitochondrial and nuclear phylogenies may yield conflicting results if the mitochondrial lineages have experienced introgressive hybridization and/or the lineages have not yet completely sorted (Peters *et al.* 2007; Barbanera *et al.* 2009; Nevado *et al.* 2009; Rheindt *et al.* 2009; Aboim *et al.* 2010; Koblmüller *et al.* 2010). Often then both mitochondrial and nuclear genes are necessary to reveal the true evolutionary history of a lineage.

Lineage sorting is the process by which haplotypes or alleles become fixed within a species such that all alleles within that species coalesce to a single ancestral allele (Heckman *et al.* 2007). In contrast, incomplete lineage sorting and the retention of ancestral polymorphism are observed when multiple haplotypes or alleles persist in a population and are shared across species. This phenomenon is relatively common among recent and rapidly radiating species as these species have not yet had time to fix for alternative haplotypes or alleles (Galtier & Daubin 2008). As a result, the phylogenetic relationships of incipient species typically progress from initial polyphyly through paraphyly and reach monophyly once lineage sorting is complete in both of the two sister species. Thus, relatively young species may appear polyphyletic or paraphyletic owing to incomplete lineage sorting (Avise 2000; Funk & Omland 2003; Heckman *et al.* 2007).

Introgressive hybridization, the transmission of alleles via hybridization from one species into the gene pool of second species (Gompert *et al.* 2008), is increasingly believed to have played a role in the diversification of plants and animals (Dowling & Secor 1997; Kane *et al.* 2009). In nature, unidirectional mitochondrial DNA (mtDNA) introgression is very common (Wirtz 1999; Chan & Levin 2005; Linnen & Farrell 2007; Peters *et al.* 2007; Barbanera *et al.* 2009; Nevado *et al.* 2009; Rabosky *et al.* 2009). Occasionally unidirectional mtDNA introgression leads to the complete replacement of the mtDNA in the recipient species by mitochondrial genome capture (Redenbach & Taylor 2002; Good *et al.* 2008; Barbanera *et al.* 2009; Nevado *et al.* 2009; Rabosky *et al.* 2009).

In animals, cyprinid fishes in the order Cypriniformes have experienced frequent interspecific and intergeneric hybridization (Demarais & Minckley 1992; Demarais *et al.* 1992; Dowling & DeMarais 1993; Freyhof *et al.* 2005; Aguilar & Jones 2009; Aboim *et al.* 2010; Gilles *et al.* 2010). However, reports of hybridization between loach Cypriniformes are rare except in the genera *Cobitis*, *Misgurnus* and *Paramisgurnus* (Saitoh *et al.* 2004, 2010; Šlechtová *et al.* 2008; Tang *et al.* 2008). Hybridization in the hillstream loaches, i.e., the balitorids, has never been documented.

The family Balitoridae (*Sensu* Chen & Tang 2000; referring to hillstream loaches) is a species rich group of Cypriniformes that possess enlarged ventral paired fins that grip the benthos in high-velocity streams. These fishes inhabit mountain torrents and are distributed from India through Southeast Asia including Sumatra, Java and Borneo and west to China and Taiwan (Nelson 2006). The greatest concentration of endemic balitorid fishes occurs in southern China; more than half of the genera of balitorid fishes are found in this region.

Balitorid fishes are one of the most specialized and successful groups of torrent-inhabiting fishes (Chang 1945). As a result, they have become a model system for the study of adaptive evolution in mountain streams. However, the evolutionary history of balitorid fishes is poorly understood which limits inferences made concerning their adaptive evolution.

Morphological and molecular data suggest that the Balitoridae consists of two distinct monophyletic lineages. These lineages correspond to two subfamilies, Gastromyzoninae and Balitorinae (Chen 1980; Chen & Tang 2000; Tang *et al.* 2006). In these subfamilies, 29 genera with at least 170 species are currently recognized (Nelson 2006). The Balitorinae genus *Lepturichthys* is endemic to China and can be easily identified by its elongated and slender tail and numerous barbels around the mouth. Only two species, *L. fimbriata* and *L. dolichopterus*, are found in this genus. *Lepturichthys dolichopterus* has longer paired fins with their standard length being 4.2–4.8 times the length of the pectoral fins (vs. 6.5–7.9 in *L. fimbriata*), and standard length being 4.8–5.7 times the length of the pelvic fins (vs. 6.5–8.3 in *L. fimbriata*). The distance between anus and anal fin is shorter in *L. dolichopterus* compared to *L. fimbriata*, and fewer vertebras are present in *L. dolichopterus* relative to *L. fimbriata* (4 + 31–33 vs. 4 + 35–37) (Dai 1985; Kottelat & Chu 1988; Chen & Tang 2000). *Lepturichthys fimbriata* and *L. dolichopterus* are allopatric. *Lepturichthys fimbriata* is widely distributed in the Upper and Middle Yangtze River, while *L. dolichopterus* is limited to the Upper Minjiang River in Fujian Province, located in the south-eastern part of China (Chen & Tang 2000) (Fig. 1A).

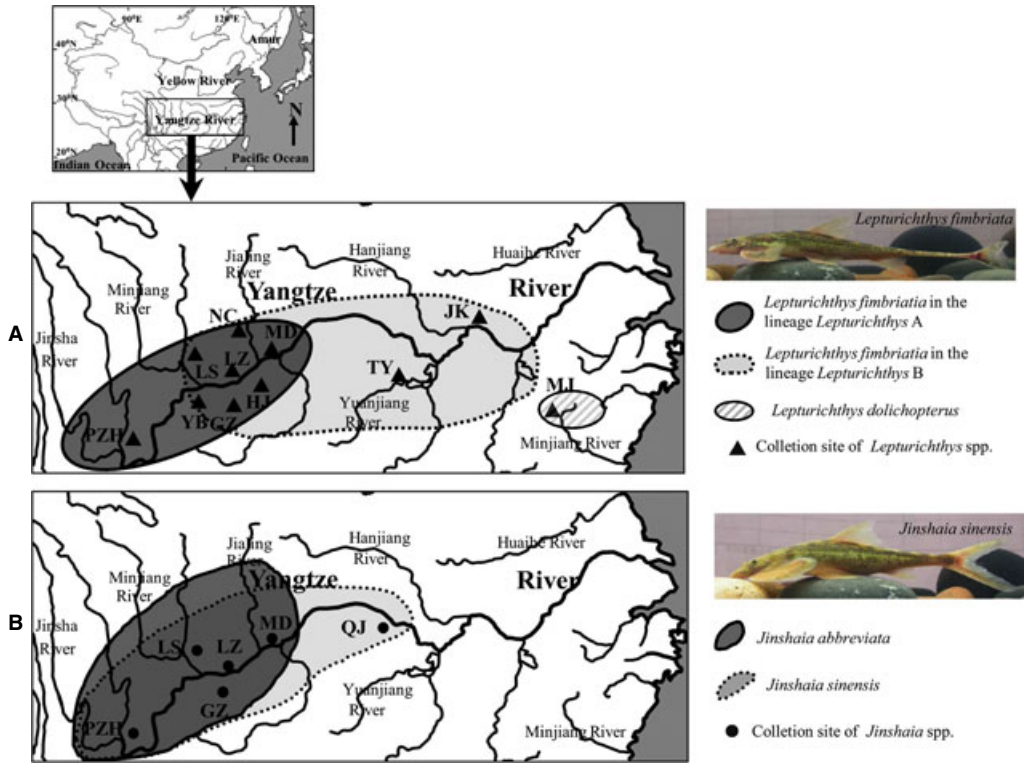

Fig. 1 The distribution and collection sites of *Lepturichthys* spp. —A, and *Jinsha* spp. —B are indicated on the map. The following locality abbreviations are used: PZH, Panzhihua, Sichuan Prov.; LS, Leshan, Sichuan Prov.; YB, Yibin, Sichuan Prov.; LZ, Luzhou, Sichuan Prov.; GZ, Chishui, Guizhou Prov.; NC, Nanchong, Sichuan Prov.; HJ, Hejiang, Chongqing; MD, Mudong, Chongqing; QJ, Qingjiang, Hubei Prov.; JK, Jinkou, Hubei Prov.; TY, Taoyuan, Hunan Prov.; MJ, Jianou, Fujian Prov.

Jinshaia is another Balitorinae genus endemic to China. Fishes in this genus have very long and deeply forked caudal fins. Within this genus, only two species, *J. sinensis* and *J. abbreviata*, are regarded as valid. Morphologically, *J. sinensis* can be distinguished from *J. abbreviata* based on a number of morphological characters. *Jinshaia sinensis* has a greater number of pelvic fin rays compared to *J. abbreviata* (18–19 vs. 14–15). In addition, the ends of the pelvic fins extend to the anus in *J. sinensis*, whereas the pelvic fins do not extend to the anus in *J. abbreviata* (Kottelat & Chu 1988; Chen & Tang 2000). Both these two species are distributed in the Upper Yangtze River where they are largely sympatric (Chen & Tang 2000; Fig. 1B).

Given their value as a model system for understanding morphological adaptation, a reconstruction of the evolutionary relationships within and between *Lepturichthys* and *Jinshaia* is needed. In this study, we investigate the phylogenetic relationships and diversification processes of these two Balitorinae genera using mitochondrial and nuclear genes. The results of this study elucidate the forces shaping the evolutionary history of the Balitoridae and provide greater insight into the mechanisms which gave rise to this pair of sister taxa.

Materials and methods

Sample collection

To investigate the evolutionary relationships of the *Lepturichthys* and *Jinshaia* species, 138 individuals were studied. Of these, 96 belonged to the four focal species (Fig. 1, Table 1) and 33 belonged to related Balitorinae (14 species spanning six genera). Seven samples spanning five species belonging to three genera from the subfamily Gastromyzoninae were used to root the Balitorinae, while two *Myxocyprinus asiaticus* samples were used as an outgroup to the Balitoridae. Detailed information for all of these samples is listed in Table 1. All vouchered specimens were preserved in 95% ethanol and deposited in the Institute of Hydrobiology, Chinese Academy of Sciences. The code for each sample used for the present analyses and their localities are given in Table S1.

DNA extraction, PCR amplification and sequencing

Total DNA was extracted from muscles using the salt-extraction procedure of Tang *et al.* (2008). The targeted regions were amplified using the polymerase chain reaction (PCR). In this study, two mitochondrial genes, Cyt*b* and COI, and one nuclear gene, RAG1, were used for the

Table 1 Summary of the locality, sample size and sample code for specimens used in the present study

Species	Locality	Drainage	Sample size	Sample code	Sample size in different genes		
					Cyt*b*	COI	RAG1
Lepturichthys fimbriata	Jinkou, Hubei Prov.	Middle Yangtze River	16	LfiJK1-16	15	16	5
	Taoyuan, Hunan Prov.	Middle Yangtze River	3	LfiTY1-3	2	2	
	Mudong, Chongqing	Upper Yangtze River	4	LfiMD1-4	4	4	2
	Hejiang, Chongqing	Upper Yangtze River	5	LfiHJ1-5	5	5	1
	Yibin, Sichuan Prov.	Upper Yangtze River	8	LfiYB1-8	8	8	4
	Luzhou, Sichuan Prov.	Upper Yangtze River	1	LfiLZ1	1		
	Panzhihua, Sichuan Prov.	Upper Yangtze River	11	LfiPZH1-11	11	7	6
	Leshan, Sichuan Prov.	Upper Yangtze River	3	LfiLS1-3	3	1	
	Nanchong, Sichuan Prov.	Upper Yangtze River	8	LfiNC1-8	6	8	2
	Chishui, Guizhou Prov.	Upper Yangtze River	5	LfiGZ1-5	5	5	1
	Jianou, Fujian Prov.	Minjiang River	9	LdoMJ1-9	7	5	8
Lepturichthys dolichopterus	Unknown	Unknown	1	LdoAY1	1		
	Panzhihua, Sichuan Prov.	Upper Yangtze River	8	JsiPZH1-8	7	6	5
Jinshaia sinensis	Mudong, Chongqing	Upper Yangtze River	3	JsiMD1-3	3	3	2
	Luzhou, Sichuan Prov.	Upper Yangtze River	1	JsiLZ1	1		
	Qingjiang, Hubei Prov.	Upper Yangtze River	1	JsiQJ1	1		
Jinshaia abbreviata	Panzhihua, Sichuan Prov.	Upper Yangtze River	4	JabPZH1-4	4	3	2
	Mudong, Chongqing	Upper Yangtze River	2	JabMD1-2	2	2	
	Leshan, Sichuan Prov.	Upper Yangtze River	1	JabLS1	1		
	Chishui, Guizhou Prov.	Upper Yangtze River	2	JabGZ1-2	2	2	2
Hemimyzon yaotanensis			1	Hyao	1		1
Hemimyzon taitungensis			2	Htai1-2	2		
Hemimyzon formosanum			3	Hfor1-3	3		
Sinogastromyzon puliensis			3	Spul1-3	3		
Sinogastromyzon szechuanensis			3	Ssze1-3	3	1	1
Sinogastromyzon hsiashiensis			4	Shsi1-4	3	4	2
Sinogastromyzon sichangensis			4	Ssic1-4	4	2	1
Sinogastromyzon wui			3	Swui1-3	3	2	
Sinogastromyzon tonkinensis			2	Ston2-3	2	2	2
Metahomaloptera omeiensis			1	Mome	1	1	1
Balitora elongata			2	Belo1-2	2		
Sinohomaloptera kwangsiensis			4	Skwa1-4	4	3	3
Homaloptera leonardi			1	Hleo	1	1	
Vanmanenia caldwelli			1	Vcal	1		
Crossostoma lacustre			1	Clac	1	1	
Crossostoma cheyiyui			1	Cche	1		
Beaufortia szechuanensis			2	Bsze1-2	2	2	2
Beaufortia liui			2	Bli1-2	2	2	
Myxocyprinus asiaticus			2	Masi1-2	2	1	1
Total			138		130	99	54

Sample code includes abbreviation of species name (and abbreviation of locality in *Lepturichthys* spp. and *Jinshaia* spp.), followed by individual's number.

analysis. The primers L14724 and H15915 were adopted from Xiao *et al.* (2001) to amplify Cyt*b*. The COI fragment was amplified using the following primers: LCOIa (5′-CCT ACC TGT GGC AAT CAC RCG C-3′) and HCOI (5′-GTG AAT AGG GGG AAT CAG TG-3′). Exon 3 of RAG1 was amplified using primers modified from Folmer *et al.* (1994) and López *et al.* (2004): RAG1F1 (5′-AGC TGC AGT CAG TAY CAC AAG ATG T-3′) and RAG1.4090R1 (5′-CTG AAT TCT TGT GAG CCT CCA TRA AC-3′). PCR was performed using the following program: an initial 94 °C denaturation for 3 min, followed by 35 cycles of 94 °C denaturation for 30 s (Cyt*b* and COI) or 40 s (RAG1), annealing at 56 °C (Cyt*b* and COI) or 52 °C (RAG1) for 45 s, extension at 72 °C for 1 min (Cyt*b* and COI) or 1.5 min (RAG1) and a final 72 °C extension for 8 min. PCR products were sent to commercial sequencing companies for purifying and sequencing. The primers for PCR were used for sequencing. All obtained sequences in this study have been deposited in GenBank (Accession numbers are listed in Table S1).

Sequence variation and phylogenetic reconstruction

Multiple alignments of sequences were performed using the CLUSTAL X (Thompson *et al.* 1997) alignment editor and verified by eye in SEAVIEW (Galtier *et al.* 1996). Nucleotide compositions of the sequences were calculated in PAUP* v. 4.0b10 (Swofford 2002). Sequence variations were analysed using MEGA v. 4 (Tamura *et al.* 2007). The analysis of nucleotide saturation was performed by plotting the absolute number of transitions (Ti) and transversions (Tv) against genetic distance values estimated in PAUP*.

The calculation of haplotype diversity (Hd), nucleotide diversity (Pi) and average number of pairwise differences (K) and the generation of the haplotype data file were performed in DnaSPv.5 (Librado & Rozas 2009). Tajima's D test and Fu's *Fs* test were used to check for neutral evolution of the mitochondrial and nuclear genes, which were carried out in ARLEQUIN v. 3.11 (Excoffier *et al.* 2005).

Molecular phylogenetic relationships were estimated using each gene separately in addition to using the concatenated mitochondrial gene data set. For each gene, three analysis methods, neighbor-joining (NJ), maximum parsimony (MP) and Bayesian inference (BI), were used for phylogenetic reconstruction. For the concatenated mitochondrial gene data set, MP and BI analyses were used. The best model of evolution to fit the NJ and BI analyses of the data was identified using the Akaike Information Criterion (AIC) as estimated in Modeltest 3.7 (Posada & Crandall 1998). NJ analysis was carried out in MEGA using the evolutionary model most similar to the one yielded by Modeltest as not all models produced by Modeltest are available in MEGA. MP analysis was conducted with PAUP*: equal weight was employed for all sites; tree searches were performed by heuristic searches using SPR branch swapping; and only minimal step length trees were retained and zero-length branches were collapsed. To assess the relative robustness of the hypothesized clades, bootstrap analyses were performed with 1000 replicates in NJ and MP analyses.

Bayesian analysis was carried out using MrBayesv.3.1.2 (Ronquist & Huelsenbeck 2003). The starting trees were random, and Bayesian posterior probabilities were estimated by using a Metropolis-coupled Markov chain Monte Carlo (MCMCMC) sampling approach. Four simultaneous Markov chains were run for 6 000 000 generations sampling every 1000 generations using two parallel runs. MCMCMC convergence was explored by examining the potential scale reduction factor (PSRF, which is expected to approach 1.0) for all parameters, the average standard deviation of split frequencies (Stdev, which is expected to approach zero), and graphically using the program TRACER v.1.4 (Rambaut & Drummond 2007). Generally, runs became stable within the first 1 000 000 generations of each analysis. We followed a conservative approach by discarding all samples obtained during the first 2 000 000 generations as "burn-in"; therefore, 50% majority-rule consensus trees with posterior probability values for each node were obtained from the remaining 4001 trees.

Furthermore, to learn about the relationship among haplotypes of species in the two genera *Jinshaia* and *Leptu-richthys*, their concatenated mitochondrial (Cyt*b* and COI) and nuclear RAG1 gene sequences were separately used for constructing median-joining networks, which was performed in NETWORK 4.6 (Bandelt *et al.* 1999; freely available at http://www.fluxus-engineering.com) using the MP calculation with the default parameters. For combined mitochondrial gene sequences, we only chose samples whose Cyt*b* and COI gene sequences were both obtained.

Estimation of divergence times

Nucleotide substitution rate constancy was evaluated using a two-cluster test in the LINTREE program (Takezaki *et al.* 1995). Sequences that showed significantly different substitution rates at the 1% level were excluded from further analysis. The Bayesian approach with MCMC analysis was performed using the program BEAST v 1.4.6 (Drummond & Rambaut 2007) to date the focal clades. Because of the absence of fossil or geological data, divergence times were crudely estimated under the strict-clock model. Based on the mitochondrial Cyt*b* gene data set, an average mutation rate of 1% per million years was adopted because it was widely employed in fish mitochondrial Cyt*b* gene analyses (Durand *et al.* 2002; Ketmaier *et al.* 2004; Zhang *et al.* 2008). Posterior estimates were obtained by sampling every 1000 generations from a chain length of 10 000 000 generations with a burn-in of 10%. The effective sample size (ESS) for parameter estimates and convergence was checked using the program Tracer. The software FIGTREE v.1.3.1 (Rambaut 2010) was used to display the summarized and annotated phylogenetic tree with a molecular clock constraint yielded by BEAST.

Results

Sequence data

After trimming, the partial Cyt*b* (1074 bp), COI (1110 bp) and RAG1 exon 3 (1486 bp) gene sequences were used for analyses. For Cyt*b* gene sequences, 101 haplotypes were identified from the sequences of 130 samples. Haplotype diversity (Hd) and nucleotide diversity (Pi) equalled 0.994 and 0.0722, respectively. Only one shared haplotype was observed in both *Jinshaia sinensis* and *J. abbreviata*. Within a given species, haplotypes were shared by individuals from the same or different localities (see Fig. S1). The average nucleotide composition for the haplotype data set

was A = 27.5%, T = 27.3%, G = 14.9% and C = 30.3%. Strong compositional biases against G existed at the second and third positions. This bias was most pronounced at the third position where just 4.7% of all bases at this position were G. The content of A + T (54.8%) was much higher than that of G + C (45.2%). Among the 1074 bp, 466 sites were variable, of which 404 were parsimony informative. The average rate of transitions/transversions (Ti/Tv) was 5.124.

Of the COI gene sequences, 87 haplotypes from 99 samples were identified. The values of Hd and Pi (0.978 and 0.0623, respectively) were slightly lower than those of the Cyt*b* gene. Haplotype distribution information is shown in Fig. S2. No haplotypes were shared among the different species. For all haplotype sequences, the average nucleotide composition was A = 26.1%, T = 28.2%, G = 19.3% and C = 26.4%. The lowest G content was at the third position (10.2%). While this is low, it is much higher than what was observed in the Cyt*b* gene. Similar to Cyt*b* gene sequences, the content of A + T (54.3%) was much higher than that of G + C (45.7%). Among the 1110 bp, 410 sites were variable, of which 313 were parsimony informative. The average rate of transitions/transversions (Ti/Tv) was 7.877.

For nuclear RAG1 exon 3 gene segments, 51 haplotypes were defined from 54 sequences. RAG1 sequences have a similar Hd value (0.994) to the mitochondrial gene sequences. However, the Pi value for RAG1 sequences (0.0218) was lower than that observed in the mitochondrial sequences. In contrast to the mitochondrial genes, the RAG1 gene had an almost equal content of the four bases (A = 23.7%, T = 22.2%, G = 28.5% and C = 25.6%), and the content of G + C (54.1%) was much higher than that of A + T (45.9%). Among the 1486 bp, 319 sites were variable, of which 172 were parsimony informative. The average rate of transitions/transversions (Ti/Tv) was 4.363.

For all three genes, an examination of transitions and transversions vs. TamNei distance showed that both transitions and transversions had not reached saturation (data not shown). Additional summary statistics for the three genes of *Lepturichthys* and *Jinshaia* species are listed in Table 2.

Table 2 Summary statistics for three genes of *Lepturichthys* and *Jinshaia* species

	Cyt b					COI					RAG1				
	Nhap	Hd	Pi	K	Neutrality test	Nhap	Hd	Pi	K	Neutrality test	Nhap	Hd	Pi	K	Neutrality test
Lepturichthys fimbriata	43 (60)	0.986	0.0338	36.331	Tajima's D (D; *P*): −0.656; 0.334.	41 (56)	0.982	0.0357	35.791	Tajima's D (D; *P*): −1.092; 0.189.	18 (21)	0.986	0.0042	6.276	Tajima's D (D; *P*): −0.774; 0.258.
Lepturichthys dolichopterus	4 (8)	0.643	0.0035	3.750	Fu's *Fs* (Fs; *P*): −5.306; 0.305	2 (5)	0.400	0.0015	1.200	Fu's *Fs* (Fs; *P*): −1.780; 0.421	6 (8)	0.893	0.0029	4.250	Fu's *Fs* (Fs; *P*): −4.013; 0.251
Jinshaia sinensis	11 (12)	0.985	0.0066	6.970		9 (9)	1.000	0.0064	5.944		7 (7)	1.000	0.0034	4.524	
Jinshaia abbreviata	6 (9)	0.889	0.0090	9.611		6 (7)	0.952	0.0098	10.476		4 (4)	1.000	0.0050	6.667	
Lepturichthys spp.	47 (68)	0.985	0.0362	39.91		39 (61)	0.966	0.0352	28.836		24 (29)	0.985	0.0049	7.300	
Jinshaia spp.	16 (21)	0.971	0.0081	8.590		12 (16)	0.958	0.0074	6.975		11 (11)	1.000	0.0039	5.255	

Nhap, number of haplotypes (in parenthesis number of individuals sequenced); Hd, Haplotype diversity; Pi, nucleotide diversity; K, average number of pairwise differences.

Phylogenetic reconstruction

mtDNA phylogeny. Separate phylogenetic trees were constructed for the Cyt*b* and COI haplotype data. NJ, MP and BI analyses for these two genes yielded similar topologies; therefore, only the NJ trees are presented in Figs S1 and S2. For the concatenated data set, the Cyt*b* and COI sequences of 137 samples were combined to produce an alignment of 2184 bp in length. Of these 2184 bp, 873 were variable sites that yielded 736 parsimony informative sites. MP and BI analyses yielded similar topologies to

those based on each of the single mitochondrial gene analyses. As a result, only the BI tree of the concatenated data set is presented in Fig. 2.

The phylogenetic trees consistently showed that the genus *Jinshaia* is monophyletic, yet its two species, *J. sinensis* and *J. abbreviata*, were not reciprocally monophyletic. Instead, these species nested within each other. Thus, both species within this genus are polyphyletic. The average genetic distances between these two species are very low (0.008 ± 0.001 for Cyt*b* and 0.006 ± 0.001 for COI) (Table 3).

For the genus *Lepturichthys*, all samples were divided into two distant lineages (referred to as *Lepturichthys* A and *Lepturichthys* B, respectively), and the average genetic distances between them are 0.073 ± 0.008 (Cyt*b*) and 0.076 ± 0.008 (COI). *Lepturichthys* A included most of *L. fimbriata* individuals from the Upper Yangtze River and appears to be the sister clade to all *Jinshaia* species. The average genetic distances between *Lepturichthys* A and *Jinshaia* (0.024 ± 0.004 for Cyt*b* and 0.026 ± 0.004 for COI) are much lower than the genetic distances between the two lineages of *Lepturichthys*. *Lepturichthys* B is composed of a few *L. fimbriata* individuals from the Upper Yangtze River, all *L. fimbriata* individuals from the Middle Yangtze River and all *L. dolichopterus* individuals from the Minjiang River in Fujian Province, Southeastern China. In *Lepturichthys* B, all *L. dolichopterus* samples form a monophyletic group and nest within the *L. fimbriata* samples, suggesting a paraphyletic status of *L. fimbriata* in *Lepturichthys* B. The average genetic divergence between *L. dolichopterus* and *L. fimbriata* in *Lepturichthys* B is 0.009 ± 0.002 at Cyt*b* and 0.005 ± 0.001 at COI. Compared to *Lepturichthys* A, *Lepturichthys* B has a closer relationship with one clade containing *Sinogastromyzon hsiashiensis*, *Sinogastromyzon szechuanensis* and *Hemimyzon yaotanensis*.

Nuclear gene phylogeny. For the nuclear RAG1 gene, all sequences from 54 samples and the resulting haplotype data set were used to construct phylogenetic trees by NJ, MP and BI analyses. The obtained topologies are similar; therefore, only the BI tree for all samples is presented in Fig. 3. The NJ tree for the haplotype data set together with the haplotype distribution information is shown in Fig. S3. The phylogenetic tree based on RAG1 gene sequences indicates that *Jinshaia* and *Lepturichthys* are reciprocally monophyletic with high support values (for *Jinshaia*, 98 in bootstrap, 1.00 in posterior probability; for *Lepturichthys*, 97 in bootstrap, 0.95 in posterior probability, respectively). Within *Jinshaia*, *J. sinensis* and *J. abbreviata* form a polytomy with a very low genetic divergence (0.005 ± 0.001; Table 3). This is consistent with the result of the mitochondrial genes. For the genus *Lepturichthys*, all samples (including *Lepturichthys* A and *Lepturichthys* B defined in mtDNA phylogeny) cluster together and form a monophyletic group. The average genetic distance between the A and B lineages is 0.006 ± 0.001, and they have a distant relationship with *Jinshaia* species (the average genetic distances are 0.015 ± 0.003 and 0.017 ± 0.003, respectively). This finding conflicts with the results of the mitochondrial genes. However, similar to the mitochondrial genes, all *L. dolichopterus* RAG1 sequences group monophyletically and nest within all *L. fimbriata* samples. The average genetic distance between the *Lepturichthys* species is low (0.006 ± 0.001), which is consistent with the mitochondrial data (Table 3).

While the present mitochondrial and nuclear genes did not fully resolve the phylogenetic relationships of other members of the Balitorinae, the results indicate that the genus *Sinogastromyzon* is polyphyletic rather than monophyletic. Furthermore, the current, albeit, limited sampling suggests that the *Hemimyzon* species from Taiwan may have a very distant relationship with the congeneric species distributed in mainland China.

Haplotype networks

Figure 4 shows the haplotype networks of the *Lepturichthys* and *Jinshaia* species. The combined mitochondrial data group the haplotypes of these four species into three lineages corresponding to *Lepturichthys* A, *Lepturichthys* B and *Jinshaia* spp. No haplotypes are shared among different species. *Jinshaia* spp. has a close relationship with *Lepturichthys* A (28 substitutions) but is more differentiated from *Lepturichthys* B (113 substitutions) (Fig. 4A). Thirty-four of 36 individuals (94.4%) of *L. fimbriata* from the Upper Yangtze River contributed 31 haplotypes to *Lepturichthys* A. The remaining two *L. fimbriata* individuals from the Upper Yangtze River have *Lepturichthys* B haplotypes. The remaining constituents of *Lepturichthys* B consist of all 15 haplotypes of 16 *L. fimbriata* individuals from the Middle Yangtze River and two haplotypes from four *L. dolichopterus* individuals from the Minjiang River in Southeastern China. In contrast to the mitochondrial genes, data from the nuclear gene RAG1 grouped all haplotypes into two clades corresponding to *Jinshaia* spp. and *Lepturichthys* spp. (including *Lepturichthys* A and B). These two genera differed by 14 substitutions in the RAG1 gene.

Divergence times

Based on the Cyt*b* gene sequences of all samples, a two-cluster test shows that species in the subfamily Gastromyzoninae and several Balitorinae samples have significantly different substitution rates compared to the remainder of the sequences. These Gastromyzoninae and Balitorinae samples were excluded and the test was repeated. Thus, 87

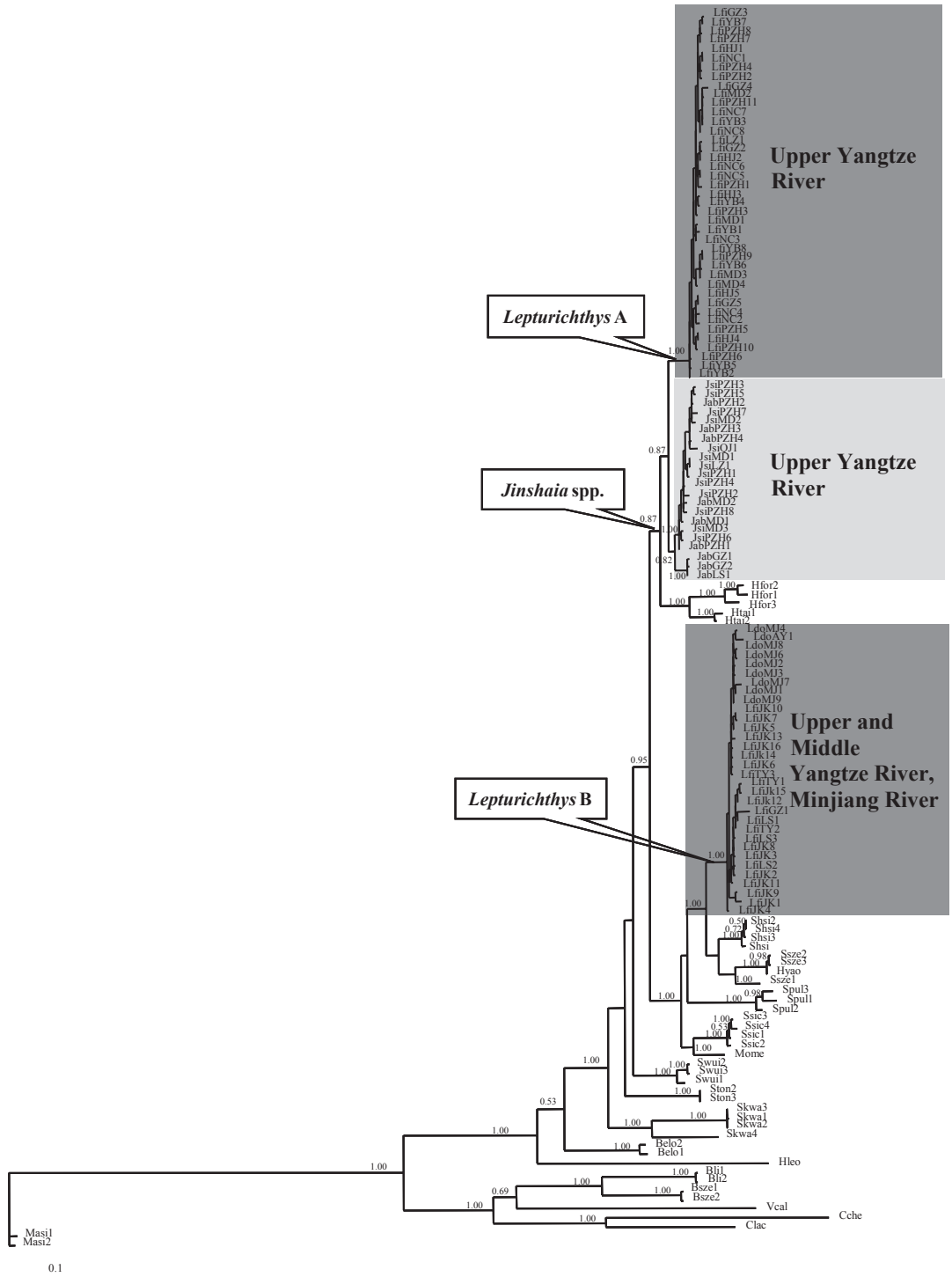

Fig. 2 Bayesian inference tree estimated by MrBayes for all samples based on the concatenated data set of Cyt*b* and COI genes. Posterior probability values are provided for nodes with $a > 0.50$ level of support. Grey rectangular boxes identify the three lineages of *Jinshaia* and *Lepturichthys* species. Species distributions are presented next to the dendogram. Sample codes follow those presented in Table S1.

Table 3 Average genetic distances and standard deviations estimated in MEGA based on Kimura 2-parameter model within each clade, between clades of *Lepturichthys* and *Jinsha* species, and between species within separate genus

	Cyt b (±SD)	COI (±SD)	RAG1 (±SD)
Within *Lepturichthys* A	0.005 ± 0.001	0.006 ± 0.001	0.003 ± 0.001
Within *Lepturichthys* B	0.007 ± 0.001	0.005 ± 0.001	0.006 ± 0.001
Within *Jinshaia* spp.	0.007 ± 0.001	0.007 ± 0.001	0.004 ± 0.001
Lepturichthys A – *Lepturichthys* B	**0.073 ± 0.008**	**0.076 ± 0.008**	**0.006 ± 0.001**
Lepturichthys A – *Jinshaia* spp.	**0.024 ± 0.004**	**0.026 ± 0.004**	**0.015 ± 0.003**
Lepturichthys B – *Jinshaia* spp.	0.071 ± 0.008	0.066 ± 0.008	0.017 ± 0.003
Lepturichthys dolichopterus – *Lepturichthys fimbriata* (in *Lepturichthys* A)	**0.075 ± 0.008**	**0.076 ± 0.008**	**0.006 ± 0.002**
L. dolichopterus – *L. fimbriata* (in *Lepturichthys* B)	0.009 ± 0.002	0.005 ± 0.001	0.006 ± 0.001
L. dolichopterus – *L. fimbriata* (*Lepturichthys* A + B)	**0.055 ± 0.006**	**0.049 ± 0.005**	**0.006 ± 0.001**
Jinshaia abbreviata – *Jinshaia sinensis*	0.008 ± 0.001	0.006 ± 0.001	0.005 ± 0.001

Bold represents the incongruence of genetic divergences between the mitochondrial and nuclear genes.

haplotype sequences were used for estimating the divergence times of the focal clades.

Figure 5 shows the chronogram with divergence times of some of the nodes displayed. The chronogram suggests that about 1.74 million years ago (confidence interval, CI: 1.31–2.51, from late Pliocene to Pleistocene) the mitochondrial lineages of *Lepturichthys* A and *Jinshaia* spp. diverged. *Jinshaia* species began to differentiate about 1.17 MYA (CI: 0.68–1.54, Pleistocene). The diversification within both *Lepturichthys* clades occurred recently (*Lepturichthys* A 0.56 MYA, CI: 0.31–0.67; *Lepturichthys* B 0.57 MYA, CI: 0.43–0.95).

Discussion

Phylogenetic relationships of the genera Lepturichthys and Jinshaia

The genus *Lepturichthys* was erected by Regan (1911) using *Homaloptera fimbriata* Günther as the type species. *Lepturichthys* is defined based on its elongated caudal peduncle (the length of caudal peduncle is more than 10 times its depth) and its large number (up to 50) of barbel-shaped lip papillae, both of which are unique characters among balitorine fishes (Kottelat & Chu 1988). The genus *Jinshaia* was defined by Kottelat & Chu (1988) from the species of *Hemimyzon* possessing the following autapomorphies: widened caudal peduncle, flattened back posterior to the dorsal fin and a very long and deeply forked caudal fin. The elongated caudal peduncle and caudal fin allow species of these two genera to be very good swimmers. Up to now, the phylogenetic position of *Lepturichthys* and *Jinshaia* within the subfamily Balitorinae and their phylogenetic relationships to each other were unclear (Chen 1980; Kottelat & Chu 1988). In order to understand the cladogenic processes that gave rise to these two genera, their phylogenetic relationships must be known.

In the present study, both mitochondrial and nuclear genes indicated that the genus *Jinshaia* is monophyletic. However, the monophyly of its two species (*J. sinensis* and *J. abbreviata*) could not be supported. Morphologically, these two species can be easily distinguished, and their species validity is accepted by most ichthyologists (Kottelat & Chu 1988; Chen & Tang 2000). For the genus *Lepturichthys*, mitochondrial genes show that it is divided into two lineages, one being a sister to *Jinshaia* and the other having a close relationship with the two *Sinogastromyzon* species (*S. szechuanensis* and *S. hsiashiensis*) and *Hemimyzon yaotanensis*. In contrast, the analysis of the nuclear gene RAG1 indicates that all *Lepturichthys* samples form a monophyletic group and are a sister group to a clade containing the *Jinshaia* species. Based on RAG1, all *L. dolichopterus* samples are monophyletic and nest within the *L. fimbriata* samples, suggesting the paraphyly of *L. fimbriata*. The monophyly of the genus *Lepturichthys* is morphologically well supported by its autapomorphies, and the validity of its two species is confirmed by their obvious diagnostic characters (Dai 1985; Kottelat & Chu 1988; Chen & Tang 2000). Therefore, both morphological characters and the analysis of the nuclear gene RAG1 support the monophyly of the genus *Lepturichthys*. As for the phylogenetic position of *Lepturichthys* and *Jinshaia* within the subfamily Balitorinae, no definitive conclusions can be drawn because of the uncertain phylogenetic relationships of the members of the genus *Sinogastromyzon* in the present study. This may be a result of incomplete species sampling: presently 20 *Sinogastromyzon* species are considered as valid with nine species distributed in China. Of these, only six species were included in this study (Froese & Pauly 2011). The limited numbers of molecular makers used in this study may have also contributed to this lack of phylogenetic resolution. However, this does not limit our ability to analyse the cladogenic process of the two genera *Lepturichthys* and *Jinshaia*.

Intergeneric mitochondrial DNA introgression and capture

In the present study, nuclear and mitochondrial genes present conflicting topologies with respect to the

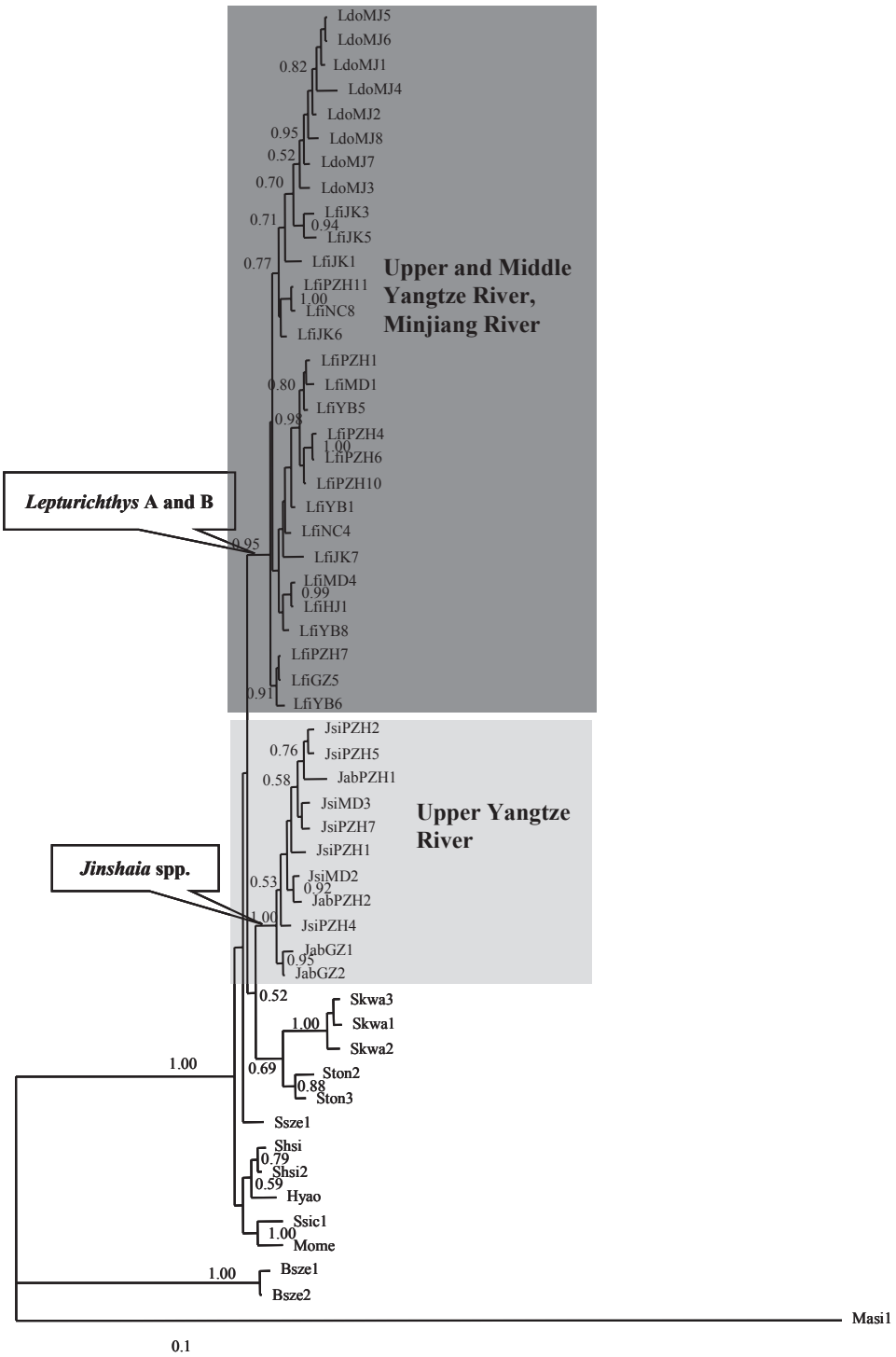

Fig. 3 Bayesian inference tree estimated by MrBayes for all samples based on the nuclear RAG1 gene sequence. Posterior probability values are provided for nodes with $a > 0.50$ level of support. Grey rectangular boxes identify the three lineages of *Jinshaia* and *Lepturichthys* species. Species distributions are presented next to the dendogram. Sample codes follow those presented in Table S1.

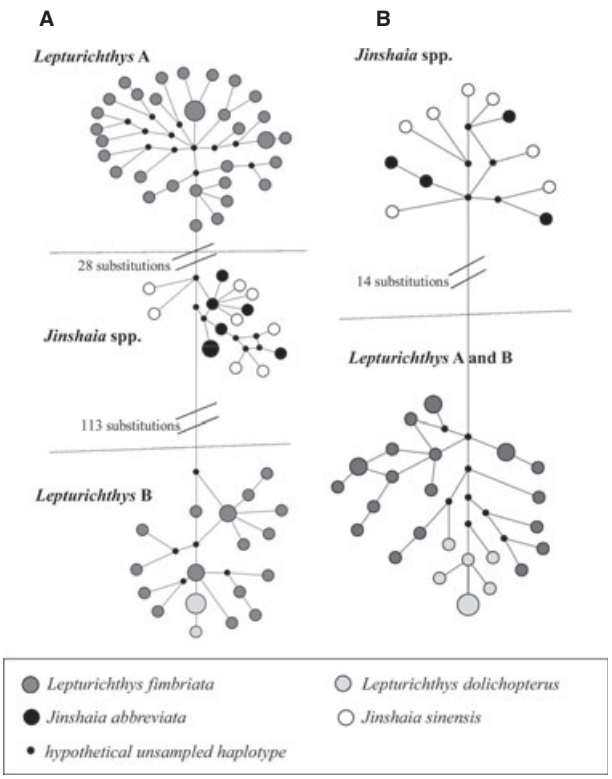

Fig. 4 Haplotype networks for the species of the two genera *Lepturichthys* and *Jinshaia*. The area of each circle is proportional to the number of shared sequences. —A. Based on concatenated mtDNA Cyt*b* and COI gene sequences; —B. based on nuclear RAG1 gene sequences.

relationships within and between *Lepturichthys* and *Jinshaia*. The mitochondrial analysis suggests that the genus *Lepturichthys* is polyphyletic: the *Lepturichthys* samples are divided into two highly differentiated lineages with an average sequence divergence of 7.3–7.6%. One lineage (*Lepturichthys* A) included only *L. fimbriata* samples from the Upper Yangtze River. This lineage is the sister group to the sympatric *Jinshaia* species with an average sequence divergence of 2.4–2.6%. The other lineage (*Lepturichthys* B) including *L. fimbriata* individuals from the Upper and Middle Yangtze River and all *L. dolichopterus* individuals from Upper Minjiang River in Southeastern China grouped with two *Sinogastromyzon* species (*S. szechuanensis* from Upper Yangtze River and *S. hsiashiensis* from Middle Yangtze River) and *Hemimyzon yaotanensis* from the Upper Yangtze River. In contrast, the analysis of the nuclear gene RAG1 indicates that all samples of *Lepturichthys* species are monophyletic with an average sequence divergence from *Jinshaia* species of 1.5–1.7%. The conflicting mitochondrial and nuclear phylogenies might be due to incomplete lineage sorting or mitochondrial introgressive hybridization.

To distinguish these competing hypotheses, we compared the topologies generated by the mitochondrial and nuclear analyses with the expected patterns of coalescence of mitochondrial and nuclear genes. This method requires that all sequences have evolved as a result of neutral processes. To verify this, both Tajima's D and Fu' s *Fs* test were used to demonstrate that none of the *Lepturichthys* or *Jinshaia* sequences deviated from the neutral model (all *P* values are more than 0.1; Table 2). Assuming the neutral evolution of these sequences and excluding several exceptions (Birky 1991), one expects that monophyly in the mitochondrial genes will be achieved four times faster than in nuclear genes (Palumbi *et al.* 2001). Further, if a nuclear locus produces a monophyletic topology, the mitochondrial locus is expected to have been monophyletic for a long time interval (Palumbi *et al.* 2001; Heckman *et al.* 2007). As the current analysis of the nuclear gene RAG1 supports the monophyly of the genus *Lepturichthys*, we speculate that rather than incomplete lineage sorting, it was introgressive hybridization between sympatric *L. fimbriata* and *Jinshaia* species that causes the conflict between mitochondrial and nuclear phylogenies.

It is worth noting that the relationship between *Lepturichthys* B, several *Sinogastromyzon* and a *Hemimyzon* species also conflicted when comparing the mitochondrial and nuclear genes. *Lepturichthys* B clustered with two *Sinogastromyzon* species and one *Hemimyzon* species in the mitochondrial phylogenetic trees, but they had a distant relationship based on the nuclear RAG1 gene. Once again, this conflict may be the result of incomplete lineage sorting or introgressive hybridization. The distributions of these species partially overlap which may have permitted introgressive hybridization. However, we cannot adequately evaluate the alternative hypotheses of introgression vs. incomplete linage sorting because of the incomplete sampling for *Sinogastromyzon* species. Future studies with additional members of *Sinogastromyzon* are needed.

Before the completion of reproductive isolation, hybridization may be a frequent occurrence among incipient species (Egger *et al.* 2007; Koblmüller *et al.* 2010). In fact, hybridization is thought to have played an important role in the adaptive radiation of some animals (Dowling & Secor 1997; Willis *et al.* 2006). In the Australian fruit fly (*Dacus tryoni*), adaptation to extreme warmth has occurred through introgressive hybridization with *Dacus humeralis* (Lewontin & Birch 1966). Finch hybrids on the Galápagos islands may exhibit higher fitness than the parental species (Grant & Grant 1992, 1994). A subspecies of the asp viper, *Vipera aspis hugyi*, frequently hybridizes with *Vipera aspis francisciredi*, which, in some cases, has produced

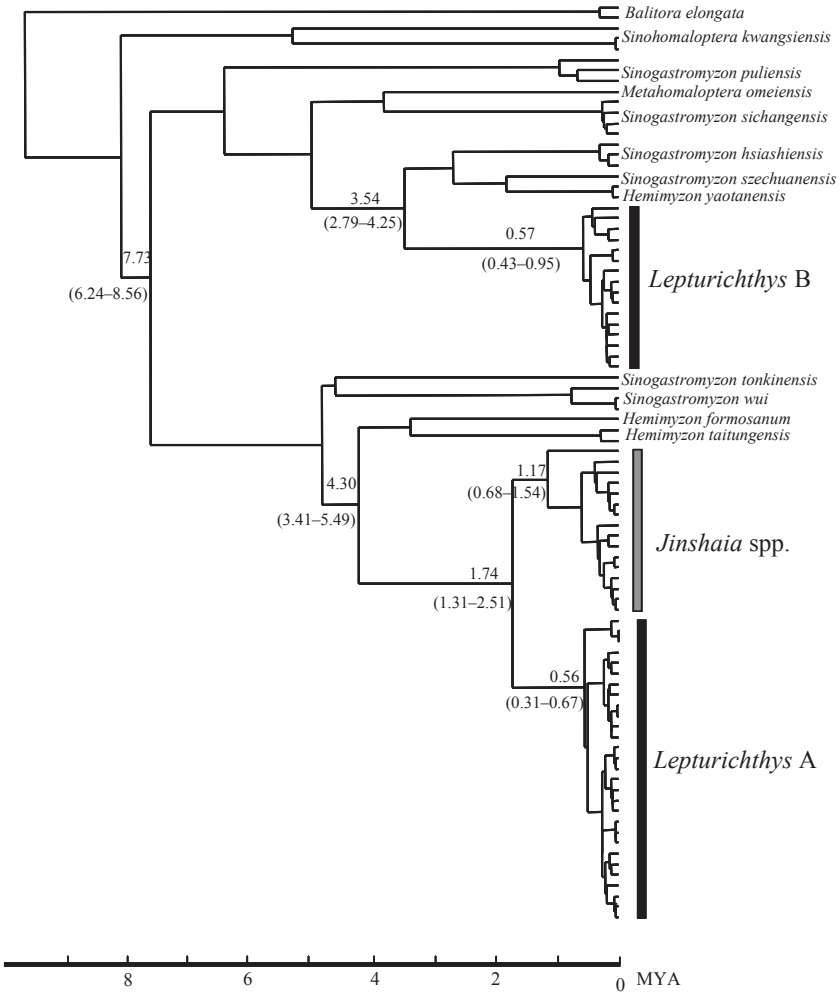

Fig. 5 Chronogram obtained using the software BEAST based on the mitochondrial Cyt*b* gene sequences. Some sequences that showed significantly different substitution rates (1% level) were excluded from this analysis. Values at the nodes represented the node ages, and their confidence interval was given in the parentheses.

highly fit hybrids (Barbanera *et al.* 2009). Among the cichlids of Lake Malawi, a hybridization event is believed to have given rise to one of the most species-rich lineages in that system (Genner & Turner 2011). The hybridization that occurred between *L. fimbriata* and *Jinshaia* species might have increased their adaptive ability leading to the fixation of the *Jinshaia* haplotype within *L. fimbriata*. Alternatively, random genetic drift can also lead to the fixation of introgressed mitochondrial haplotypes (Ferris *et al.* 1983; Barbanera *et al.* 2009; Koblmüller *et al.* 2009). With the current data set, the alternative hypotheses of adaptation and drift leading to the introgression of mitochondrial haplotypes cannot be distinguished.

As mentioned above, the introgression of mtDNA haplotypes was unidirectional: a *Jinshaia* haplotype was fixed within *L. fimbriata*. The timing of this mtDNA capture event appears to have occurred about 1.74 MYA. This would predate the diversification of *Jinshaia* species. Therefore, this introgression might have occurred between a group of *L. fimbriata* and the ancestor of *Jinshaia* spp. Unidirectional mtDNA introgression is common in nature, and occasionally it will lead to the capture of the mitochondrial genome (Wirtz 1999; Chan & Levin 2005; Linnen & Farrell 2007; Peters *et al.* 2007; Barbanera *et al.* 2009; Nevado *et al.* 2009; Rabosky *et al.* 2009). While unidirectional introgression can be explained by a number of factors (Wirtz 1999), the reason that hybridization was unidirectional in these species is unknown.

Speciation and incomplete lineage sorting

The present mtDNA and nuclear gene sequences consistently support that *J. sinensis* and *J. abbreviata* are polyphyletic. For *Lepturichthys* species, *L. dolichopterus* appears to be monophyletic and nested within *L. fimbriata*. Thus, *L. fimbriata* appears paraphyletic. As these four species have distinct diagnostic characters in morphology and species boundaries among them are quite clear, the proposed polyphyletic and paraphyletic topologies are believed to reflect

the true relationships of these groups. Funk & Omland (2003) identify six causes of species-level paraphyly and polyphyly. Among these causes, incomplete lineage sorting or interspecific hybridization likely accounts for the above-mentioned paraphyletic or polyphyletic relationships. For the two *Lepturichthys* species, interspecific hybridization can be rejected because they are allopatric (*L. fimbriata* from the Upper and Middle Yangtze River, while *L. dolichopterus* from the Upper Minjing River in Southeastern China). Therefore, the observed paraphyly is likely the result of incomplete lineage sorting. Avise (2000) proposed that incipient species would first form a polyphyletic group which would later evolve into a paraphyletic group and finally form a monophyletic clade. The paraphyletic status of *L. fimbriata* in the present study suggests that the divergence of species in the genus *Lepturichthys* is still in an early stage and the monophyletic resolution of these species is limited by the retention of ancestral polymorphism.

For the two *Jinshaia* species, the observed polyphyletic relationships may be the result of interspecific hybridization, incomplete lineage sorting or both. As these two species are sympatric and have a similar ecological and morphological characters, interspecific hybridization may possibly explain the observed patterns even if direct evidence for hybridization has not been found. Incomplete lineage sorting could also lead to the polyphyletic relationships between *J. sinensis* and *J. abbreviata*. The data suggest that the genus *Jinshaia* began to differentiate about 1.17 MYA. This relatively young age may contribute to the observed polyphyly owing to incomplete lineage sorting between *J. sinensis* and *J. abbreviata*.

Conclusion

The current study investigated the phylogenetic relationships and speciation process of the two balitorid genera *Lepturichthys* and *Jinshaia* and found that unidirectional mitochondrial introgressive hybridization had occurred between these genera. *Lepturichthys fimbriata* captured the mitochondrial genome from the sympatric *Jinshaia* species, and this capture event occurred about 1.74 MYA. Within the genus *Lepturichthys*, incomplete lineage sorting might contribute to the paraphyletic status of *L. fimbriata*, while either interspecific bidirectional hybridization or incomplete lineage sorting could lead to the polyphyletic relationships of the two species of *Jinshaia*. The observations of these patterns significantly improve our understanding of the evolution and diversification of balitorid fishes.

Acknowledgements

We thank Yu-Yu Xiong, Cai-Ping Liao, Yu Zeng, De-Qing Tan, Zhuo-Cheng Zhou and Zhi-Fu Tian for help in collecting samples and providing photographs. Thanks are also to two anonymous reviewers for their valuable advice on revising our manuscript. Funding support was provided by grants from the National Natural Science Foundation of China (NSFC 3070072 and 31061160185) and the Knowledge Innovation Program of the Chinese Academy of Sciences (Y05E08).

References

Aboim, M. A., Mavarez, J., Bernatchez, L. & Coelho, M. M. (2010). Introgressive hybridization between two Iberian endemic cyprinid fish: a comparison between two independent hybrid zones. *Journal of Evolutionary Biology*, *23*, 817–828.

Aguilar, A. & Jones, W. J. (2009). Nuclear and mitochondrial diversification in two native California minnows: insights into taxonomic identity and regional phylogeography. *Molecular Phylogenetics and Evolution*, *51*, 373–381.

Avise, J. C. (2000). *Phylogeography, the History and Formation of Species*. England: Harvard University.

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

Barbanera, F., Zuffi, M. A. L., Guerrini, M., Gentilli, A., Tofanelli, S., Fasola, M. & Dini, F. (2009). Molecular phylogeography of the asp viper *Vipera aspis* (Linnaeus, 1758) in Italy: evidence for introgressive hybridization and mitochondrial DNA capture. *Molecular Phylogenetics and Evolution*, *52*, 103–114.

Birky, C. W. J. (1991). Evolution and population genetics of organelle genes: mechanisms and models. In R. K. Selander, A. G. Clark & T. S. Whittam (Eds) *Evolution at the molecular level* (pp. 112–134). Sunderland, MA: Sinauer Associates.

Chan, K. M. A. & Levin, S. A. (2005). Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution*, *59*, 720–729.

Chang, H. W. (1945). Comparative study on the girdles and their adjacent structures in Chinese Homalopterid fishes with special reference to the adaptation to torrential stream. *Sinensia*, *16*, 9–26.

Chen, Y. Y. (1980). Systematic studies on the fishes of the family Homalopteridae of China III. Phyletic studies of the Homalopterid fishes. *Acta Zootaxonomica Sinica*, *5*, 200–211.

Chen, Y. Y. & Tang, W. Q. (2000). Homalopteridae. In P. Q. Yue (Ed.) *Fauna sinica. Osteichthyes. Cypriniformes* (pp. 438–567). Beijing: Sciences Press.

Dai, D. Y. (1985). A new species of the genus *Lepturichthys* from China. *Acta Zotaxonomica Sinica*, *10*, 221–223.

Demarais, B. D. & Minckley, W. L. (1992). Hybridization in native cyprinid fishes, *Gila ditaenia* and *Gila* sp in Northwestern Mexico. *Copeia*, *69*, 7–703.

Demarais, B. D., Dowling, T. E., Douglas, M. E., Minckley, W. L. & Marsh, P. C. (1992). Origin of *Gilaseminuda* (Teleostei:Cyprinidae) through introgressive hybridization: implications for evolution and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *89*, 2747–2751.

Dowling, T. E. & DeMarais, B. D. (1993). Evolutionary significance of introgressive hybridization in cyprinid fishes. *Nature*, *362*, 444–446.

Dowling, T. E. & Secor, C. L. (1997). The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics*, *28*, 593–619.

Drummond, A. J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214.

Durand, J. D., Tsigenopoulos, C. S., Ünlü, E. & Berrebi, P. (2002). Phylogeny and biogeography of the family Cyprinidae in the Middle East inferred from Cytochrome b DNA–evolutionary significance of this region. *Molecular Phylogenetics and Evolution*, *22*, 91–100.

Egger, B., Koblmueller, S., Sturmbauer, C. & Sefc, K. M. (2007). Nuclear and mitochondrial data reveal different evolutionary processes in the Lake Tanganyika cichlid genus *Tropheus*. *BMC Evolutionary Biology*, 7, 137.

Excoffier, L., Laval, G. & Schneider, S. (2005). Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, *1*, 47–50.

Ferris, S. D., Sage, R. D., Huang, C. M., Nielsen, J. T., Ritte, U. & Wilson, A. C. (1983). Flow of mitochondrial DNA across a species boundary. *Proceedings of the National Academy of Sciences of the United States of America*, *80*, 2290–2294.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial Cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, *3*, 294–299.

Freyhof, J., Lieckfeldt, D., Pitra, C. & Ludwig, A. (2005). Molecules and morphology: evidence for introgression of mitochondrial DNA in Dalmatian cyprinids. *Molecular Phylogenetics and Evolution*, *37*, 347–354.

Froese, R. & Pauly, D. (2011). FishBase. World Wide Web electronic publication. Available via http://www.fishbase.org, version (06/2011).

Funk, D. J. & Omland, K. E. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics*, *34*, 397–423.

Galtier, N. & Daubin, V. (2008). Dealing with incongruence in phylogenomic analyses. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *363*, 4023–4029.

Galtier, N., Gouy, M. & Gautier, C. (1996). SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Bioinformatics*, *12*, 543–548.

Genner, M. J. & Turner, G. F. (2011). Ancient hybridization and phenotypic novelty within lake Malawi's cichlid fish radiation. *Molecular Biology and Evolution*, *29*, 195–206.

Gilles, A., Costedoat, C., Barascud, B., Voisin, A., Banarescu, P., Bianco, P. G., Economidis, P. S., Marić, D. & Chappaz, R. (2010). Speciation pattern of *Telestes souffia* complex (Teleostei, Cyprinidae) in Europe using morphological and molecular markers. *Zoologica Scripta*, *39*, 225–242.

Gompert, Z., Forister, M. L., Fordyce, J. A. & Nice, C. C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in Lycaeides. *Molecular Ecology*, *17*, 5231–5244.

Good, J. M., Hird, S., Reid, N., Demboski, J. R., Steppan, S. J., Martin-Nims, T. R. & Sullivan, J. (2008). Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, *17*, 1313–1327.

Grant, P. R. & Grant, B. R. (1992). Hybridization of bird species. *Science*, *256*, 193–197.

Grant, P. R. & Grant, B. R. (1994). Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution*, *48*, 297–316.

Heckman, K. L., Mariani, C. L., Rasoloarison, R. & Yoder, A. D. (2007). Multiple nuclear loci reveal patterns of incomplete lineage sorting and complex species history within western mouse lemurs (*Microcebus*). *Molecular Phylogenetics and Evolution*, *43*, 353–367.

Kane, N. C., King, M. G., Barker, M. S., Raduski, A., Karrenberg, S., Yatabe, Y., Knapp, S. J. & Rieseberg, L. H. (2009). Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. *Evolution*, *63*, 2061–2075.

Ketmaier, V., Bianco, P. G., Cobolli, M., Krivokapic, M., Caniglia, R. & De Matthaeis, E. (2004). Molecular phylogeny of two lineages of Leuciscinae cyprinids (*Telestes* and *Scardinius*) from the peri-Mediterranean area based on Cytochrome b data. *Molecular Phylogenetics and Evolution*, *32*, 1061–1071.

Koblmüller, S., Nord, M., Wayne, R. K. & Leonard, J. A. (2009). Origin and status of the Great Lakes wolf. *Molecular Ecology*, *18*, 2313–2326.

Koblmüller, S., Egger, B., Sturmbauer, C. & Sefc, K. M. (2010). Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Molecular Phylogenetics and Evolution*, *55*, 318–334.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X. & Wilson, A. C. (1989). Dynamics of mitochondrial-DNA evolution in animals – amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America*, *86*, 6196–6200.

Kottelat, M. & Chu, X. L. (1988). A synopsis of Chinese balitorine loaches (Osteichthyes: Homalopteridae) with comments on their phylogeny and description of a new genus. *Revue Suisse de Zoologie*, *95*, 181–201.

Lewontin, R. C. & Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution*, *20*, 315–336.

Librado, P. & Rozas, J. (2009). Dnasp v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*, 1451–1452.

Linnen, C. R. & Farrell, B. D. (2007). Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (Hymenoptera: Diprionidae) sawflies. *Evolution*, *61*, 1417–1438.

López, J. A., Chen, W. J. & Orti, G. (2004). Esociform phylogeny. *Copeia*, *3*, 449–464.

Meyer, A. (1993). Evolution of mitochondrial DNA in fishes. In P. W. Hochachka & T. P. Mommsen (Eds) *Biochemistry and Molecular Biology of Fishes* (pp. 1–38). Amsterdam: Elsevier Science Publishers.

Nelson, J. S. (2006). *Fishes of the World*. New York: Wiley.

Nevado, B., Koblmüller, S., Sturmbauer, C., Snoeks, J., Usano-Alemany, J. & Verheyen, E. (2009). Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. *Molecular Ecology*, *18*, 4240–4255.

Palumbi, S. R., Cipriano, F. & Hare, M. P. (2001). Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution*, *55*, 859–868.

Peters, J. L., Zhuravlev, Y., Fefelov, I., Logie, A. & Omland, K. E. (2007). Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution*, *61*, 1992–2006.

Posada, D. & Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics*, *14*, 817–818.

Rabosky, D. L., Talaba, A. L., Donnellan, S. C. & Lovette, I. J. (2009). Molecular evidence for hybridization between two Australian desert skinks, *Ctenotus leonhardii* and *Ctenotus quattuordecimlineatus* (Scincidae: Squamata). *Molecular Phylogenetics and Evolution*, *53*, 368–377.

Rambaut, A. (2010). FigTree. Available via http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut, A. & Drummond, A. J. (2007). Tracer v1.4. Available via http://beast.bio.ed.ac.uk/Tracer.

Redenbach, Z. & Taylor, E. B. (2002). Evidence for historical introgression along a contact zone between two species of char (Pisces : Salmonidae) in northwestern North America. *Evolution*, *56*, 1021–1035.

Regan, C. T. (1911). Classification of Teleostean fishes of the Order Ostariophysi. I. Cyprinoidea. *Annals and Magazine of Natural History*, *8*, 31–32.

Rheindt, F. E., Christidis, L. & Norman, J. A. (2009). Genetic introgression, incomplete lineage sorting and faulty taxonomy create multiple cases of polyphyly in a montane clade of tyrant-flycatchers (*Elaenia*, Eyrannidae). *Zoologica Scripta*, *38*, 143–153.

Ronquist, F. & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, *19*, 1572–1574.

Saitoh, K., Kim, I. S. & Lee, E. H. (2004). Mitochondrial gene introgression between spined loaches via hybridogenesis. *Zoological Science*, *21*, 795–798.

Saitoh, K., Chen, W. J. & Mayden, R. L. (2010). Extensive hybridization and tetrapolyploidy in spined loach fish. *Molecular Phylogenetics and Evolution*, *56*, 1001–1010.

Šlechtová, V., Bohlen, J. & Perdices, A. (2008). Molecular phylogeny of the freshwater fish family Cobitidae (Cypriniformes: Teleostei): delimitation of genera, mitochondrial introgression and evolution of sexual dimorphism. *Molecular Phylogenetics and Evolution*, *47*, 812–831.

Swofford, D. L. (2002). *PAUP*: Phylogenetic Analysis Using Parsimony (* and Other Methods)*, *Version 4*. Sunderland, MA: Sinauer Associates.

Takezaki, N., Rzhetsky, A. & Nei, M. (1995). Phylogenetic tests of the molecular clock and linearized trees. *Molecular Biology and Evolution*, *12*, 823–833.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, *24*, 1596–1599.

Tang, Q. Y., Liu, H. Z., Mayden, R. & Xiong, B. X. (2006). Comparison of evolutionary rates in the mitochondrial DNA Cytochrome b gene and control region and their implications for phylogeny of the Cobitoidea (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution*, *39*, 347–357.

Tang, Q. Y., Freyhof, J., Xiong, B. X. & Liu, H. Z. (2008). Multiple invasions of Europe by east Asian Cobitid loaches (Teleostei: cobitidae). *Hydrobiologia*, *605*, 17–28.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, *25*, 4876–4882.

Willis, B. L., van Oppen, M. J. H., Miller, D. J., Vollmer, S. V. & Ayre, D. J. (2006). The role of hybridization in the evolution of reef corals. *Annual Review of Ecology Evolution and Systematics*, *37*, 489–517.

Wirtz, P. (1999). Mother species-father species: unidirectional hybridization in animals with female choice. *Animal Behaviour*, *58*, 1–12.

Xiao, W. & Zhang, Y. (2000). Genetics and evolution of mitochondrial DNA in fish. *Acta Hydrobiologica Sinica*, *24*, 384–391.

Xiao, W., Zhang, Y. & Liu, H. (2001). Molecular systematics of Xenocyprinae (Teleostei: Cyprinidae): taxonomy, biogeography, and coevolution of a special group restricted in East Asia. *Molecular Phylogenetics and Evolution*, *18*, 163–173.

Zhang, L., Tang, Q. Y. & Liu, H. Z. (2008). Phylogeny and speciation of the eastern Asian cyprinid genus *Sarcocheilichthys*. *Journal of Fish Biology*, *72*, 1122–1137.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Neighbor-joining tree estimated by MEGA using the mitochondrial Cyt*b* haplotype sequences of all of the samples.

Fig. S2. Neighbor-joining tree estimated by MEGA using the mitochondrial COI haplotype sequences of all of the samples.

Fig. S3. Neighbor-joining tree estimated by MEGA using RAG1 gene sequences of all of the samples.

Table S1. All samples used for the present study with locality, sample code, specimen voucher number, and GenBank accession number are provided.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.