Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 39 (2006) 529–541
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Molecular phylogeny of the Southeast Asian freshwater fish family Botiidae (Teleostei: Cobitoidea) and the origin of polyploidy in their evolution

Vendula Šlechtová [a,b,*], Jörg Bohlen [a], Jörg Freyhof [c], Petr Ráb [a,d]

[a] *Institute of Animal Physiology and Genetics of the Academy of Sciences of the Czech Republic, Rumburská 89, 277 21 Liběchov, Czech Republic*
[b] *Department of Zoology, Faculty of Biological Sciences, University of South Bohemia, Branišovská 31, 370 05 České Budějovice, Czech Republic*
[c] *Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany*
[d] *Joint Laboratory of Genetics, Physiology and Reproduction of Fishes of IAPG CAS, Liběchov, and RIFH USB, 389 25 Vodňany, Czech Republic*

Received 1 September 2005; accepted 12 September 2005
Available online 6 December 2005

Abstract

The freshwater fish family Botiidae is represented by seven genera on the Indian subcontinent and in East and Southeast Asia and includes diploid as well as evolutionary tetraploid species. We present a phylogeny of Botiidae including 33 species representing all described genera using the mitochondrial cytochrome *b* and 12s rRNA genes to reconstruct the phylogenetic relationships among the genera and to estimate the number of polyploidisation events during their evolution. Our results show two major lineages, the subfamilies Leptobotiinae with the genera *Leptobotia* and *Parabotia* and Botiinae with the genera *Botia*, *Chromobotia*, *Sinibotia*, *Syncrossus*, and *Yasuhikotakia*. Our results suggest that two species that were traditionally placed into the genus *Yasuhikotakia* form a monophyletic lineage with the species of *Sinibotia*. A review of the data on the ploidy level of the included species shows all diploid species to belong to Leptobotiinae and all tetraploid species to Botiinae. A single polyploidisation event can therefore be hypothesised to have occurred in the ancestral lineage leading to the Botiinae.

Keywords: Loach fishes; Tetraploidy; Polyploidisation; Cytochrome *b*; 12s rRNA

1. Introduction

Botiid loaches represent an interesting model to study the role of polyploidisation in vertebrate evolution since they include diploid (with $2n = 50$ chromosomes) as well as evolutionary tetraploid species ($2n = 98$–100 chromosomes). From the 26 species surveyed by Suzuki and Taki (1996) in a review of ploidy level in botiid fishes, 11 were diploid and 15 tetraploid. Polyploidisation is well known as an important evolutionary force in plants and indications for its importance in the evolution of animals are constantly accumulating (Le Comber and Smith, 2004; Soltis and Soltis, 1995, 1999). Polyploidisation events seem to be more common than they had been until recently believed (Leggatt and Iwama, 2003; Soltis and Soltis, 1999) and recurrent formations of polyploid taxa were already called the norm rather than the exception (Soltis and Soltis, 1999). Examples of fish groups, in which changes in ploidy level have been already identified as key events in their evolution include Acipenseridae (Ludwig et al., 2001), Cyprinidae (Alves et al., 2001; David et al., 2003), Catostomidae (Ueno et al., 1988), and Salmonidae (Crespi and Fulton, 2004; Phillips and Ráb, 2001). Multiple origins of polyploidy were demonstrated in African barbs (Tsigenopoulous et al., 2002) and Cobitidae (Janko et al., 2003). However, to evaluate the origin of polyploidisation within Botiidae and its contribution to its evolution, it is necessary to understand the phylogenetic relationships between the diploid and tetraploid species.

* Corresponding author. Fax: +420 315 639 510.
E-mail address: v.slechtova@iapg.cas.cz (V. Šlechtová).

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.09.018

The relationships within Botiidae and the natural lineages inside this family have been formerly discussed mainly on the basis of morphological data. For a long time, they were considered a subfamily of the family Cobitidae and believed to include only two genera, *Botia* and *Leptobotia*, the former one with three subgenera *Botia*, *Sinibotia* and *Hymenophysa* (today called *Syncrossus*) (Fang, 1936; Nalbant, 1963; Sawada, 1982; Taki, 1972). Nalbant (2002) and Kottelat (2004) used family rank for botiid loaches, genus rank for the former subgenera and described the genera *Yasuhikotakia* and *Chromobotia*. Nalbant's (2002) suggestion to include the genus *Vaillantella* as subfamily *Vaillantellinae* into the Botiidae was in disagreement with the opinions of Kottelat (1994), Roberts (1989), and Sawada (1982) that *Vaillantella* is a nemacheilid loach and was strongly refuted by Kottelat (2004). Nalbant (2002) divided his subfamily Botiinae (Botiidae without *Vaillantella*) into two tribes, Leptobotiini (containing the genera *Leptobotia*, *Parabotia*, and *Sinibotia*) and Botiini (with the genera *Botia*, *Hymenophysa*, and *Yasuhikotakia*). Recently, Tang et al. (2005) demonstrated with mtDNA data that Chinese Botiidae form a monophyletic group distinct from other loach fishes and that the genus *Sinibotia* represents the sister lineage to a lineage formed by the sister genera *Leptobotia* and *Parabotia*, but no other genera were included into their study. Since the taxonomic studies of Nalbant (2002) and Kottelat (2004) have considered *Sinibotia* to be much closer related to *Leptobotia* and *Parabotia* that other genera of Botiidae, the diversity of the family was by far not covered by the study of Tang et al. (2005). Especially the tetraploid species were strongly underrepresented in the study of Tang et al. (2005), since only three species (all *Sinibotia*) were included. To understand the phylogenetic relationships within the Botiidae and the relationship of diploid and tetraploid species, a comparison of all described genera and most species is needed.

The aim of the present study was to reconstruct the phylogenetic relationships among the genera and main lineages of the family Botiidae on the base of a wide spectrum of representatives of all described genera using sequences of mitochondrial DNA. The results of our DNA analyses are compared with the outline of the recently recognised genera that were based on morphological data. Agreement between genetic and morphological data indicates reliable groupings that most likely reflect true phylogenetic relationships. Another goal was to evaluate the phylogenetic relationships of diploid and tetraploid species in order to estimate the origin of polyploidy in the evolution of Botiidae.

2. Materials and methods

2.1. Taxon sampling

According to the most recent review (Kottelat, 2004), Botiidae include seven genera with 47 species. The present study comprises 96 individuals of 33 species, covering the diversity of the family (3 out of 5 species of *Sinibotia*, 6 out of 8 of *Botia*, 4 out of 7 of *Parabotia*, 7 out of 13 of *Leptobotia*, 7 out of 9 of *Yasuhikotakia*, all species of *Syncrossus* (5) and the monotypic *Chromobotia*) (Table 1). Due to the extensive sampling, our study covers all valid genera of Botiidae and all generic type species. Initial analyses showed that *Vaillantella* is not closely related to botiids; therefore *Vaillantella* was not included into this study. Regarding Botiidae, 24 sequences of cytochrome *b* were obtained from GenBank, two sequences were provided by K.-E. Witte; the remaining 70 sequences of cytochrome *b* and all 48 sequences of 12S rRNA are original data. As outgroup taxa of the family Cobitidae we included cytochrome *b* sequences of *Cobitis bilineata* and *Sabanejewia larvata* and 12S rRNA sequence of *Cobitis sinensis* from GenBank and two original sequences of *Sabanejewia balcanica*. Specimens of Botiidae were obtained from commercial imports. Identification of the examined fishes followed Kottelat (2001, 2004), Rainboth (1996), Roberts (1989), Taki (1972), and Yang and Chen (1992). M. Kottelat kindly checked our identification of the specimens of the genera *Syncrossus* and of *Y. sidthimunki* and *Y. nigrolineata*. Our specimens of *B. rostrata* and *B. almorhae* were identified by comparison with the syntypes of *B. rostrata* (BMNH 1855.12.26.694 and 1860.3.19.114) and *B. almorhae* (BMNH 2002.9.18.1). Voucher specimens are available in the collection of IAPG.

2.2. DNA isolation, PCR, and sequencing

Genomic DNA was isolated from fin or muscle tissue following the standard phenol–chloroform method (Sambrook et al., 1989) or with DNeasy Tissue Kit (QIAGEN).

PCR amplification was performed in 50 μl reaction volumes containing 10 mM Tris–HCl, 50 mM $(HN_4)_2SO_4$, 0.1% Triton X-100, 1.2–1.8 mM $MgCl_2$, 2 mM TMA oxalate (PCR enhancer), 10 nmol of each nucleotide, 2.5 U of *Taq* polymerase (all chemicals by Top–Bio) and 25 pmol of each primer. The PCR profile (carried out on MJ Research thermocycler) started with 2 min period of initial denaturation at 95 °C followed by 35 cycles each consisting of a denaturation step at 94 °C for 30 a primer annealing step from 52 to 57 °C for 30 s and elongation at 72 °C for 45 s. The PCR was completed by a final elongation step of 5 min at 72 °C. Problematic samples were PCR amplified in two overlapping fragments using internal primers under the same PCR conditions. PCR products were purified by ethanol precipitation or using Microcon PCR Filter Units (Millipore) and subdued to cycle sequencing employing BigDye Terminator Cycle Sequencing Kit v.3.1 (PE Applied Biosystems) according to manufacturer's instructions. Sequencing products cleaned by ethanol precipitation or with DyeEx 2.0 Spin Kit (QIAGEN) were resolved on ABI Prism 310 Genetic Analyser (Perkin Elmer). Each sample was sequenced with the same primers as used for double strand PCR amplification. Primers used in this study are listed in Table 2.

Table 1
List of samples, individual numbers, and GenBank accession numbers

Genus	Species	Indiv. No.	Cytochrome *b* Accession No.	12S rDNA Accession No.
Botia	*almorhae*[*]	A103	AY887790	AY887735
		A102	AY887791	AY887736
		A104	AY887792	AY887738
	dario	A322	AY887798	—
		A258	AY887799	—
	histrionica	A041	AY887794	AY887733
		A045	AY887795	AY887734
	kubotai	A163	AY887796	AY887731
		A164	AY887797	AY887732
	lohachata	A055	AY887789	AY887737
		A054	AY887793	AY887739
	rostrata	A244	AY887785[a]	—
		A087	AY887786	AY887728
		A088	AY887787	AY887729
		A086	AY887788	AY887730
	striata	A011	AY887783	AY887740
		A010	AY887784	AY887741
Chromobotia	*macracanthus*[*]	A178	AY887840	AY887771
		A179	AY887841	AY887772
Leptobotia	*elongata*[*]	A215	AY887778	—
		A214	AY887779	AY887773
		H1	AY625714[b]	—
		h2	AY625715[b]	—
	guilinensis	A124	AY887780	AY887774
		A205	AY887781	—
	pellegrini	h1	AY625723[b]	—
	rubrilabris	h1	AY625716[b]	—
		h2	AY625717[b]	—
	taeniops	h1	AY625718[b]	—
	tchangi	h1	AY625719[b]	—
		h2	AY625720[b]	—
		h3	AY625722[b]	—
		h4	AY625721[b]	—
	tientaiensis	h1	AY625725[b]	—
		h2	AY625724[b]	—
Parabotia	*banarescui*	h1	AY625711[b]	—
	fasciata[*]	A217	AY887782	AY887775
		h1	AY625709	—
		h2	AY625710[b]	—
	kiangsiensis[c]	h1	AY625712[b]	—
	lijiangensis	h1	AY625713[b]	—
Sinibotia	*pulchra*	A015	AY887800	AY887769
		A016	AY887801	AY887770
		A212	AY887802	—
		A396	AY887803	—
		A397	AY887804	—
		h1	AY625705[b]	—
		h2	AY625706[b]	—
	robusta	A024	AY88780	AY887765
		A242	AY887806[a]	AY887764
		A071	AY887807	AY887763
		h1	AY625707[b]	—
		h2	AY625708[b]	—
	superciliaris[*]	h1	AY625704[b]	—
		h2	AY625702[b]	—
		h3	AY625703[b]	—
Syncrossus	*beauforti*	A098	AY887814	AY887753
		A099	AY887815	AY887755
		A059	AY887816	AY887754
		A060	AY887817	—
		A061	AY887818	—

(*continued on next page*)

Table 1 (*continued*)

Genus	Species	Indiv. No.	Cytochrome *b* Accession No.	12S rDNA Accession No.
		A100	AY887822	—
	berdmorei[*]	A277	AY887823	AY887757
	sp. aff. berdmorei	A278	AY887812	—
		A279	AY887813	AY887756
		A318	AY887821	—
	helodes	A169	AY887819	AY887758
		A170	AY887820	AY887759
	reversa	A049	AY887808	AY887760
		A052	AY887809	AY887761
	hymenophysa	A435	AY887810	AY887762
		A436	AY887811	—
Yasuhikotakia	*caudipunctata*	A209	AY887837	AY887750
	eos	A093	AY887824	AY887743
		A095	AY887825	AY887742
		A166	AY887826	AY887746
		A165	AY887827	AY887747
		A063	AY887828	AY887745
		A062	AY887829	AY887744
	lecontei	A173	AY887838	—
		A172	AY887839	AY887752
	modesta[*]	A428	AY887830	—
		A429	AY887831	—
		A021	AY887832	—
		A200	AY887833	—
		A020	AY887834	AY887751
	morleti	A067	AY887835	AY887748
		A068	AY887836	AY887749
	nigrolineata	A031	AY887845	—
		A032	AY887846	—
		A033	AY887847	AY887766
		A034	AY887848	—
	sidthimunki	A183	AY887842	AY887768
		A186	AY887843	—
		A423	AY887844	—
		A185	AY887849	AY887767
Cobitis	*bilineata*		AF263092[b]	—
	sinensis			AY009147[b]
Sabanejewia	*balcanica*	A457	AY887850	AY887776
		A458	—	AY887777
	larvata	1	AY059334[b]	—

The ype species for each genus is marked with asterisk.

[a] Sequence obtained from K.-E. Witte.

[b] Obtained from GenBank.

[c] Probably misspelling of *P. kwangsiensis*.

2.3. Molecular data analyses

2.3.1. Sequence alignment

Chromatograms were assembled and checked by eye for potential mistakes using the SeqMan II module of the DNA-Star software package (http://www.dnastar.com). Edited sequences were aligned using the default settings in ClustalX software (Thompson et al., 1997). The alignment was manually revised in BioEdit (Biological sequence alignment editor v5.0.9, http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The 12S rRNA sequences were preliminarily aligned with default settings of gap opening/extension penalties and consequently carefully adjusted by eye with help of SSU rRNA Secondary Structure Prediction and Alignment (http://www.cse.ucsc.edu/research/compbio/ssurrna.html), a web service which allows modelling of secondary structure of given sequences. In following phylogenetic analyses the gaps within the alignment of 12S rRNA were treated as indels. All sequences are deposited in GenBank under Accession Nos. AY887730–AY887851 (for details see Table 1).

Plots of the number of absolute sequence difference (transitions, Ti, and transversions, Tv) against the uncorrected sequence divergence (*p*-distance) indicated a moderate tendency for saturation for Ti at third codon position of cytochrome *b*. For the 12S rRNA dataset and for the first and second codon position of cytochrome *b*, both Ti and Tv increased linearly with distance (data not shown).

2.3.2. Phylogenetic analyses

Statistical information on the sequences were obtained in PAUP* version 4.0b10 (Swofford, 2002) and sequence divergences were calculated in MEGA 2.1 software

Table 2
List of primers used in this study

Primer name	Sequence 5′–3′	Source
Cytochrome b		
Glu-L.Ca14337–14359	GAA GAA CCA CCG TTG TTA TTC AA	
CB-L.Ca14975–14994	CAC GAR ACR GGR TCN AAY AA	
CB-H.Ca15057–15035	TCT TTR TAT GAG AAR TAN GGG TG	Witte, pers. com.
Thr-H.Ca15568–15548	ACC TCC RAT CTY CGG ATT ACA	
GluDG.L	TGA CTT GAA RAA CCA YCG TTG	Palumbi et al. (1996)
H16460	CGA YCT TCG GAT TAA CAA GAC CG	Perdices and Doadrio (2001)
12S rRNA		
Phe-L.Ca19–38	AAA GCA TAG CAC TGA AGA TG	
SSU-L.Ca70–87	YAA AGG CWT GGT CCC GAC	
SSU-H.Ca631–615	AGA ACA GGC TCC TCT AG	Witte, pers. com.
Val-H.Ca1065–1046	CTC GGT GTA AGT GAG ATG CT	
L1091	AAA AAG CTT CAA ACT GGG ATT AGA TAC CCC ACT AT	Kocher et al. (1989)
H1478	TGA CTG CAG AGG GTG ACG GGC GGT GTG T	

(Kumar et al., 2001). To estimate the phylogenetic relationships among the botiid taxa we performed Bayesian analyses with use of MrBayes ver. 3.0 (Huelsenbeck and Ronquist, 2001). The analyses were performed on each dataset separately as well as on a combined cytochrome *b* and 12S rRNA dataset. The cytochrome *b* dataset consisted of 99 sequences, the 12S rRNA dataset consisted of 51 sequences and the combined dataset contained 49 sequences of those individuals, which were sequenced for both fragments. For the analyses of the combined datasets the congruence between the both gene partitions was tested using the incongruence length difference test with 100 replications in PAUP* (Farris et al., 1994; Mickevich and Farris, 1981).

Modeltest 3.06 (Posada and Crandall, 1998) was used in order to determine the best fitting models of nucleotide substitution for each dataset. The Akaike information criterion indicated that the general time reversible model with gamma distribution of rate heterogenity and proportion of invariable sites (GTR + Γ + I) is the most appropriate for each of the datasets (data available from authors on request).

For the Bayesian analyses, the cytochrome *b* dataset was partitioned into three partitions corresponding to the first, second and third codon position. The analyses of the combined dataset were conducted with two approaches. First, different evolutionary models were allowed for two partitions corresponding to the both genes (cytochrome *b* and 12S rRNA). The second analysis was performed with four partitions by further dividing cytochrome *b* gene into codon positions. For each analysis, six Markov Chains Monte Carlo were run simultaneously for 1,000,000 generations with trees and likelihood scores sampled each 100 generations. The analyses were based on the estimated models. The burn-in period was evaluated graphically by plotting the likelihood scores against the generation times. The values reached the stability after about 30,000 for cytochrome *b*, 20,000 generations in case of 12S rRNA and after 15,000 generations for the combined datasets. Nevertheless, to ensure the stationarity of log-likelihoods, the first 500 trees were discarded as burnin. The remaining trees were used to build 50% majority-rule consensus trees and the posterior probabilities were used to indicate branch supports. Posterior probabilities of 90% and larger were considered as significant.

We did not calculated the times of divergence between the lineages and genera since the molecular clock hypothesis was rejected by the likelihood ratio test, which was carried out on the cytochrome *b* dataset by comparing the two likelihood scores of the maximum likelihood (ML) trees calculated with molecular clock enforced and with molecular clock not enforced in PAUP* ($\chi^2 = 148.314$, df = 97, $p < 0.001$).

3. Results

Altogether, 2079 nucleotide characters were analysed. The final matrices of botiid cytochrome *b* and 12S rRNA sequences included 1111 and 968 characters, respectively. The sequence lengths ranged from 918 to 959 bp. For detailed information on base composition, absolute numbers of variable and informative characters and Ti/Tv ratios see Table 3. The test of incongruence between the two datasets did not show any indication for significant differences ($P = 0.81$), thus both could be combined in simultaneous analyses.

The overall mean sequence diversity within Botiidae was 12.96% (SE 0.54) for the cytochrome *b* dataset and 4.98% (SE 0.37) for 12S rRNA. The mean sequence divergence among genera of Botiidae was 14.5% (SE 1.7) and 5.84% (SE 0.63) for cytochrome *b* and 12S rRNA, respectively. Details on the mean intra- and interspecific sequence divergences for each genus are listed in Table 4.

3.1. Phylogenetic analyses

The phylogenetic analyses of the cytochrome *b* (Fig. 1), 12S rRNA (Fig. 2) and combined (Fig. 3) datasets show that Botiidae split into two main lineages with very high statistical supports (in all cases the Bayesian posterior probabilities were 100%). One lineage included the genera

Table 3
Number of constant, variable, parsimony informative sites, and base composition for the 1111 bp of cytochrome *b* and 973 bp 12S rRNA datasets analysed in this study

Codon position	Const.	Var.	Inf.	A	C	G	T	Ti/Tv
Cyt b								
All	614	497	458	29.2(0.52)	28.7(1.12)	14.1(0.26)	28 (1.00)	2.5
First	256	115	99	24.2(0.57)	25.1(1.02)	26.2(0.61)	24.4(1.10)	9.4
Second	341	29	19	20(0.05)	25(0.25)	13 (0.23)	41.9(0.38)	2.1
Third	17	353	340	43.2(1.65)	35.9(2.5)	3.1(0.76)	17.8(2.2)	2.1
12S rRNA	703	267	217	31 (0.46)	26.4 (0.51)	23.8 (0.37)	19.9 (0.47)	2.3

The values in the parentheses indicate standard deviations for the base frequencies.

Table 4
Intra- and interspecific sequence divergences within the genera of botiid loaches for cytochrome *b* and 12S rRNA datasets

Genus	N (sp/ind)	% of sequence divergence for cytochrome *b* (mean; SE)			% of sequence divergence for 12S rRNA (mean; SE)		
		Intra	Inter	Overall	Intra	Inter	Overall
Leptobotia	7/16	0.62 (0.16)	7.24 (0.71)	6.24 (0.38)	—	1.65 (0.38)	1.65 (0.38)
Parabotia	4/6	1.86 (0.33)	8.31 (0.80)	7.09 (0.49)	—	—	—
Botia	6/17	0.26 (0.07)	7.64 (0.45)	7.89 (0.44)	0.04 (0.03)	1.01 (0.29)	0.85 (0.17)
Sinibotia	3/15	0.45 (0.09)	6.89 (0.49)	7.33 (0.48)	0.09 (0.05)	2.03 (0.32)	1.82 (0.31)
Syncrossus	6/16	0.56 (0.08)	9.73 (0.87)	7.61 (0.55)	0.07 (0.04)	4.91 (0.71)	4.26 (0.39)
Yasuhikotakia	7/24	1.58 (0.14)	11.55 (0.83)	10.02 (0.48)	0.51 (0.18)	4.28 (0.63)	3.77 (0.38)
Yasuhikotakia[a]	5/16	2.09 (1.45)	10.71 (0.89)	8.56 (0.55)	0.71 (0.23)	4.17 (0.89)	3.01 (0.34)
Chromobotia	1/2	0.36 (0.18)	—	0.36 (0.18)	0.00 (0.00)	—	0.00 (0.00)

Sequence divergence between Leptobotiinae and Botiinae was 15% (SE 0.67) for cytochrome *b* and 7.4% (SE 0.6) for 12S rRNA.
[a] Without the questionable species *Y. sidthimunki* and *Y. nigrolineata*.

Leptobotia and *Parabotia* and the second one the genera *Botia*, *Chromobotia*, *Sinibotia*, *Syncrossus*, and *Yasuhikotakia*. The sequence divergence between the two major lineages was 15.0% (SE 0.7) for cytochrome *b* and 7.4% (SE 0.6) for 12S rRNA data. The second clade was divided into two subclades, with the species of *Botia* and *Chromobotia* representing the sister lineage to the species of *Syncrossus*, *Sinibotia*, and *Yasuhikotakia*. Despite their well-supported sister relationship (98% posterior probability in the analysis of the cytochrome *b* dataset; 97% in the analysis of the combined dataset), the genera *Botia* and *Chromobotia* formed two distinct lineages (sequence divergence 15.07% for the cytochrome *b* and 5.68% for the 12S rRNA dataset, SE 0.92 and 0.67, respectively). The analyses of 12S rRNA did not resolve their relationships.

Within the second subclade, the results from the combined and the 12S rRNA dataset analyses suggested a sister-relationship of *Syncrossus* and *Yasuhikotakia* (statistical support of 100% in both cases), while the analyses of the cytochrome *b* dataset identified *Yasuhikotakia* as sister lineage to the remaining species (posterior probabilities 72%).

Our analyses did not confirm the monophyly of the genus *Yasuhikotakia*. Two of the species, *Y. sidthimunki* and *Y. nigrolineata*, appeared as sister lineage to the genus *Sinibotia* (posterior probabilities 96–100%) and were not closely related to the other species of *Yasuhikotakia*.

The analyses of the combined dataset with two and four defined partitions resulted in topologically congruent trees with only small variations in the values of Bayesian posterior probabilities (generally higher values in the dataset with four partitions). Nevertheless, both analyses evaluated the same nodes as significant. The tree depicted in Fig. 3 bases on the analyses with the dataset divided into four partitions.

4. Discussion

4.1. Main lineages

Within the family Botiidae, our phylogenetic analyses revealed two major well-supported monophyletic lineages. As mentioned above, the genus *Vaillantella* does not belong to Botiidae, therefore we will onwards refer to the two main clades as subfamilies Leptobotiinae and Botiinae. The subfamily Leptobotiinae includes the genera *Leptobotia* and *Parabotia*, while the subfamily Botiinae comprises the genera *Botia*, *Chromobotia*, *Sinibotia*, *Syncrossus*, and *Yasuhikotakia*. The subfamilies correspond to the genera *Leptobotia* and *Botia* sensu Nalbant (1963) and Taki (1972). From the tribes Leptobotiini and Botiini of Nalbant (2002) they differ in the placement of the genus *Sinibotia*, which Nalbant included into Leptobotiinae (his tribus Leptobotiini), but which according to our data belongs to the Botiinae as sister genus of the genera *Yasuhikotakia* and *Syncrossus*.

In a recent study on the phylogeny of Chinese Botiids, Tang et al. (2005) observed that the sister genera *Leptobotia*

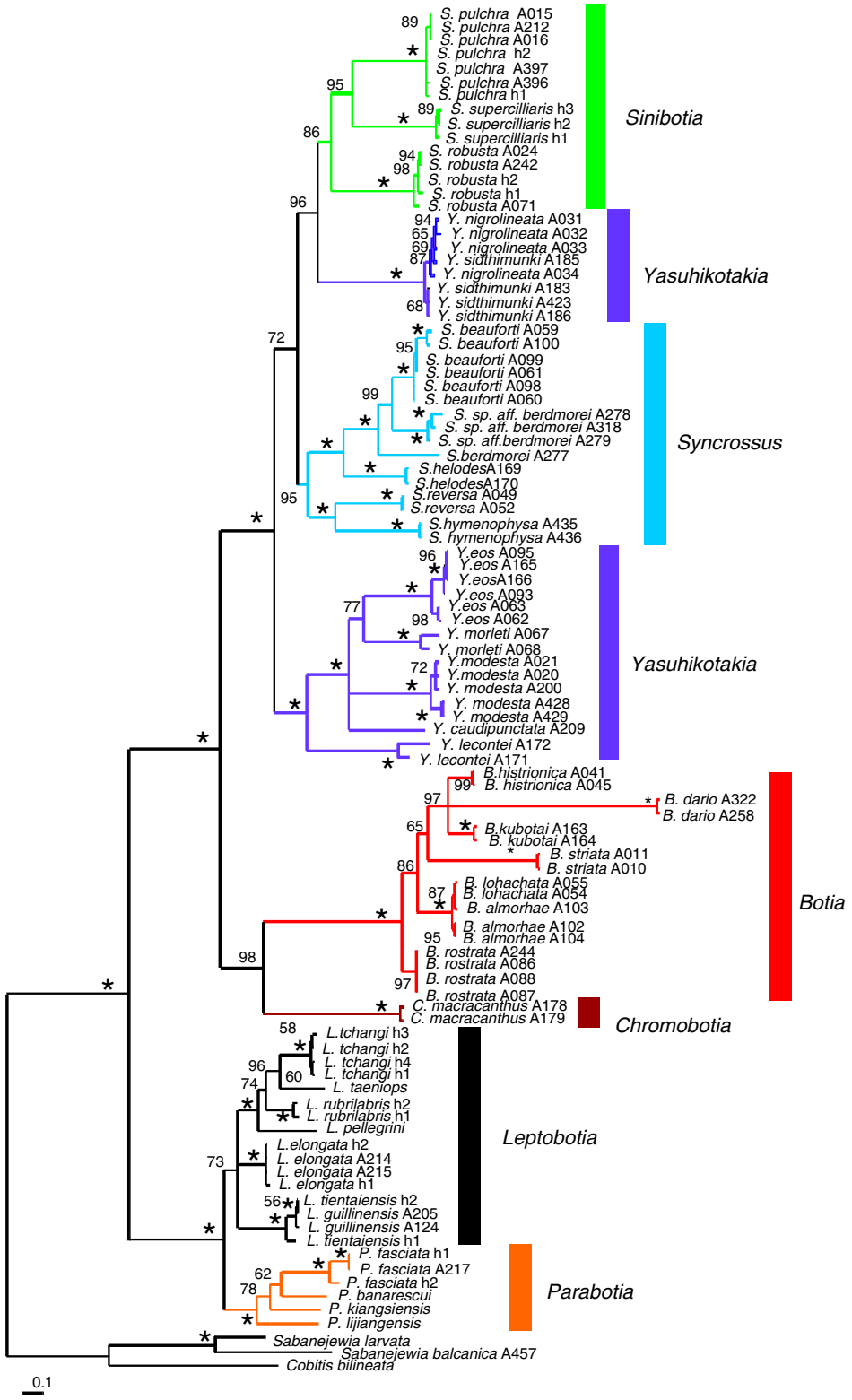

Fig. 1. Bayesian tree for the cytochrome *b* dataset. The numbers at the nodes indicate the Bayesian posterior probability percentages; only values higher than 50% are shown. Where the statistical support was 100%, a star replaces the values. We do not show the bootstrap supports on the intraspecific level.

and *Parabotia* (our Leptobotiinae) form the monophyletic sister lineage to *Sinibotia* (named *Botia* in the work of Tang et al., 2005). However, in recent studies (Kottelat, 2004; Nalbant, 2002) the genus *Sinibotia* was considered to be more closely related to *Leptobotia* and *Parabotia* than to the other genera, therefore *Sinibotia* was not the best choice as representative of the Botiinae. Our study includes all genera of Botiidae and shows that Sinibotia belongs to Botiinae and is not closer related to Leptobotiinae than other genera of Botiinae.

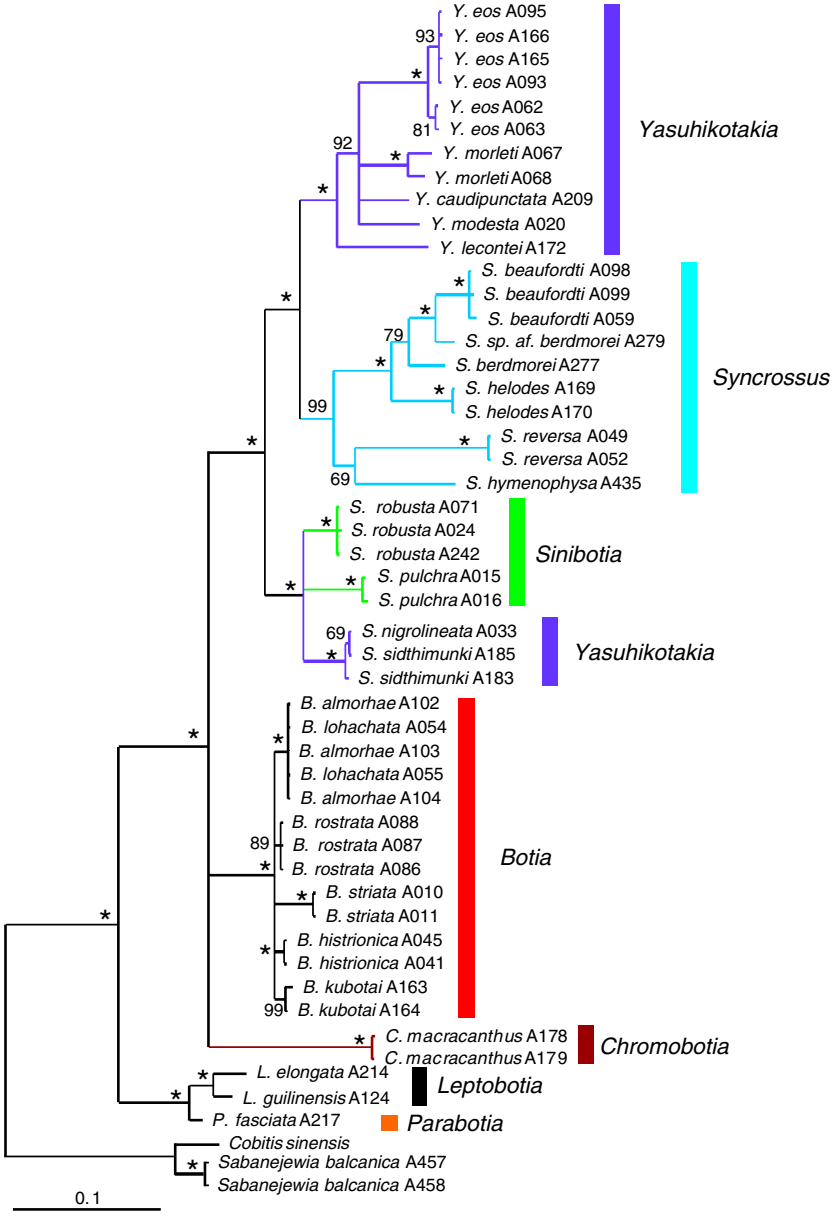

Fig. 2. Bayesian tree for 12S rDNA dataset. The numbers at the nodes indicate Bayesian posterior probability percentages; only values higher than 50% are shown. Where the statistical support was 100%, a star replaces the values.

4.2. Genera and species

Our phylogenetic analyses highly supported the monophyly of most of the genera of Botiidae as recognised by Kottelat (2004). The only exception represents the genus *Yasuhikotakia*, which is polyphyletic with two species (*Y. sidthimunkii*, *Y. nigrolineata*) being more closely related to *Sinibotia* than to the remaining samples of *Yasuhikotakia*. This may be caused by unsuited morphologic definitions of the two genera or by mitochondrial introgression, therefore a careful taxonomic investigation with morphologic and nuclear molecular markers should be carried out to clarify the generic placement of *Y. sidthimunki* and *Y. nigrolineata*.

In three cases, our analyses revealed unrecognised diversity within species. First, the samples of *Y. modesta* form two statistically well-supported groups; one of them (A020, A021, A200) comprises the specimens with red fins, while the other (A428, A429) includes the specimens with yellow fins. Our data show the difference in colouration to correspond to a genetic differentiation between the two types. The second species with two statistically well-supported internal lineages is *Y. eos*, but in this case the colouration of the fins (red versus transparent) was not congruent with the observed genetic division. Third, the specimens identified as *S. berdmorei* formed two distinct groups, making the species paraphyletic. The relatively high sequence divergences between these intraspecific lineages (1.7% in *Y. modesta*, 1.89% in *Y. eos*, 7.62 in *S. berdmorei*) suggest that a careful taxonomic revision of these groups should be carried out.

In general our data support the taxonomic sorting as suggested by morphologic studies (Kottelat, 2004; Nalbant, 2002; Taki, 1972), which means that morphological and

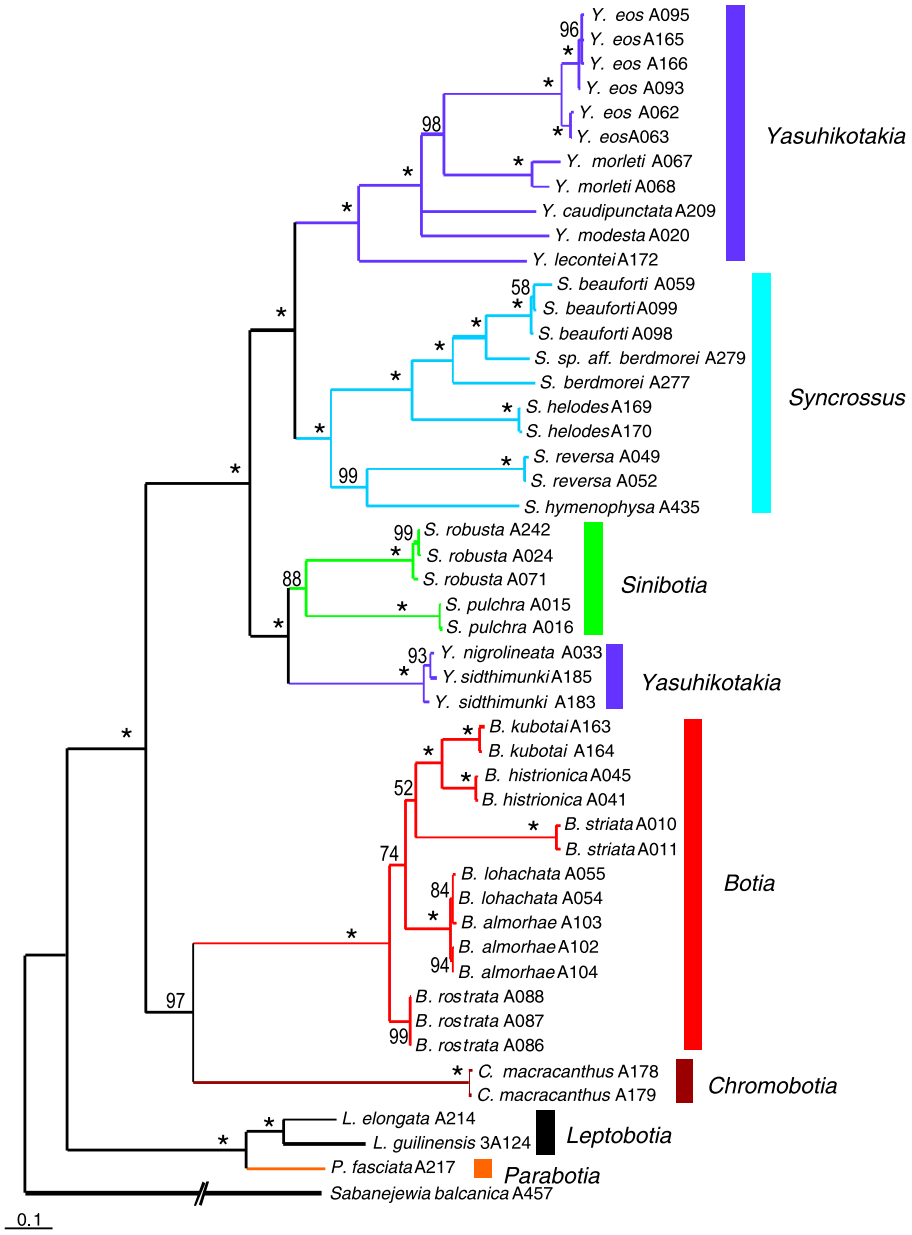

Fig. 3. Bayesian tree for complete dataset. The numbers at the nodes indicate Bayesian posterior probability percentages; only values higher than 50% are shown. Where the statistical support was 100%, a star replaces the values.

molecular data in most cases support each other. This lets us be optimistic that the presented picture reflects the natural system of the family Botiidae.

4.3. Zoogeographic patterns

Fig. 4 shows schematically the recent distribution of the botiid subfamilies and genera. The distribution areas of Leptobotiinae and Botiinae are nearly exclusive, a fact that may have promoted the genetic differentiation between them. The only exception is the genus *Sinibotia* which ranges from the Mekong basin through southern China to northern Vietnam. Since *Sinibotia* is much more closely related to the Indochinese genera of Botiinae than to the East Asian Leptobotiinae, this overlap of range has to be considered the result of a secondary range extension of *Sinibotia* from Indochina into China.

The genus *Botia* occurs on the Indian subcontinent and represents the sister lineage to *Chromobotia* which occurs on Sumatra and Borneo. The gap in distribution area is nowadays inhabited by representatives of *Syncrossus* and *Yasuhikotakia*, only the relation of *Botia* and *Chromobotia* points on a former continuous distribution.

Within the genus *Syncrossus*, the species pair *S. hymenophysa* and *S. reversa* is distinct from all the remaining species. These two species occur on Sumatra, Borneo, and the southern tip of the Malayan Peninsula, while the other species are restricted to the Malayan and Indochinese

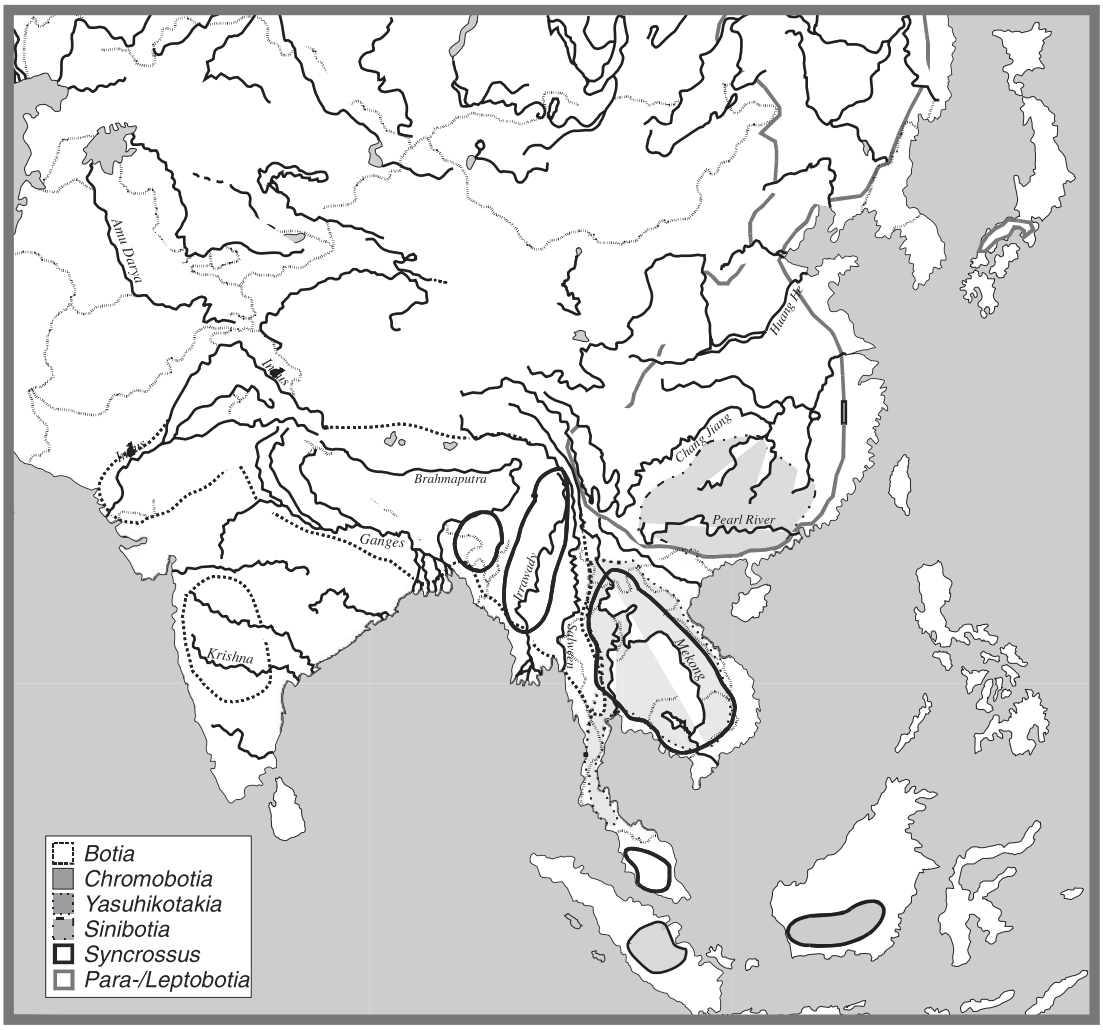

Fig. 4. Geographical distribution of the genera of the subfamily Botiinae. Leptobotiinae and Botiinae have a nearly exclusive occurrence; in the gap between the distribution areas of *Botia* and *Chromobotia* occur *Syncrossus* and *Yasuhikotakia*.

peninsulas. Therefore we take the geographic separation of *S. hymenophysa* and *S. reversa* as explanation for their genetic differentiation.

Distribution patterns also reflect phylogenetic relations in the genus *Botia*. The three closely related species *B. dario*, *B. kubotai*, and *B. histrionica* occur close to each other in the rivers of northeast India, Bangladesh, and Myanmar. Therefore we explain their close relationship by their close geographic distribution. The more distantly related species *B. striata*, *B. almorhae*, and *B. rostrata* are distributed in southern India (*B. striata*) and the western foot of the Himalayan mountain ridge (*B. almorhae* and *B. rostrata*).

4.4. Polyploidisation within the family

As pointed out before, the family Botiidae includes diploid as well as evolutionary tetraploid species. The available information on the ploidy level of botiid loaches (Ráb, unpubl. Data; Suzuki and Taki, 1996) in combination with our phylogenetic data show that Leptobotiinae comprises all diploid species, while Botiinae includes exclusively tetraploid species (Fig. 5). The tetraploid species of the genus *Sinibotia*, formerly considered as an exception within diploid Leptobotiinae, were shown by our analyses to belong to the subfamily Botiinea. This distribution of polyploidy within the group supports the separation of the two major lineages within Botiidae, the subfamilies Leptobotiinae and Botiinae. One of the important outcomes of our study is the finding that the tetraploid taxa form a monophyletic group and hence we assume a single polyploidisation event.

In general, diploidy with a chromosome number of $2n = 50$ is the plesiomorphic stage in loach fishes (Leggatt and Iwama, 2003; Suzuki, 1996), a state still present in Leptobotiinae. We therefore consider the karyotypes of Botiinae ($2n = 98$–100) to have arose from such a plesiomorph karyotype via polyploidisation, although it is unclear at present whether their evolutionarily tetraploidy is of auto- or allopolyploid origin. On the base of the absence of cheek scales and the wider distribution area of Botiinae, Tang et al. (2005) considered *Sinibotia* (meaning Botiinae) to be basal in respect to the Leptobotiinae.

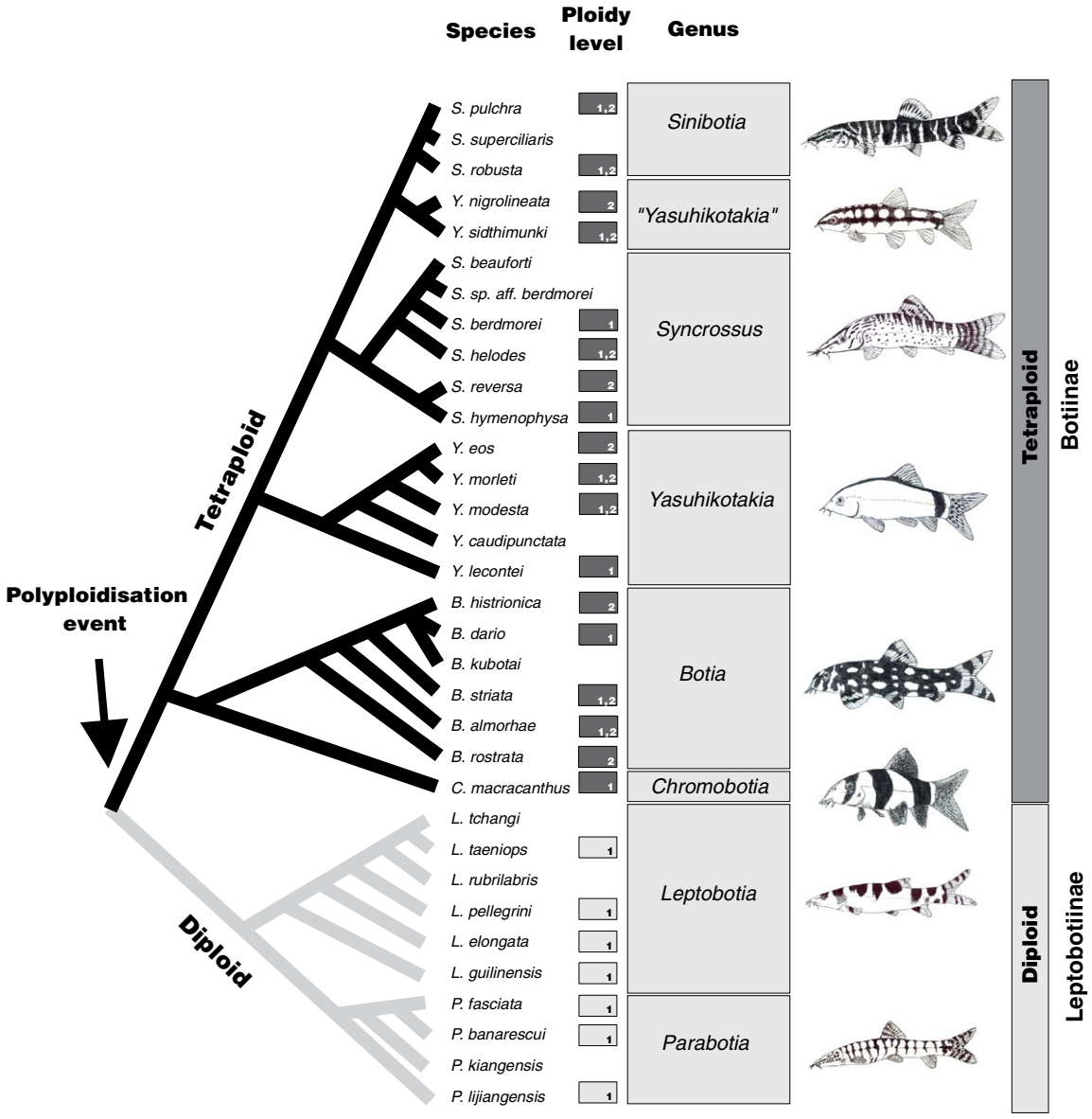

Fig. 5. Schematic overview over the phylogenetic relationships of Botiidae and the distribution of ploidy level. Diploid species are marked by a light grey box behind species name, tetraploid species with a dark grey box, species with lack of ploidy data have no box behind name. Numbers in the ploidy-box indicate source: 1, Suzuki and Taki (1996), 2, P. Ráb, unpubl. data. The arrow points on the hypothesised polyploidisation event.

Despite the conceptional problems involved in basality of recent taxa, the karyotype is considered as rather conservative and slowly developing character in comparison to characters of outer morphology and is more suited to estimate plesiomorph character states.

Besides some autopolyploid species (e.g. *Cobitis biwae* 'big race,' Kitagawa et al., 2003), the salmonids (family Salmonidae) are the only fish group where an autotetraploid origin after a single polyploidisation event was unambiguously shown (Crespi and Fulton, 2004; Phillips and Ráb, 2001), while in several cyprinid (e.g. Chenuil et al., 1999; David et al., 2003) and cobitid (Janko et al., 2003) lineages the data-sets indicate allopolyploid, i.e., hybrid, and also polyphyletic (Janko et al., 2003; Tsigenopoulous et al., 2002) origins of polyploidy.

Nevertheless, polyploid animals often experience a change in growth pattern and a broadening of physiologic tolerances and potential in habitat use due to increased heterozygosity effects. Such increase of ecological potential can allow the animal to colonise new habitats and consequently can promote a diversification due to multiple niche separations (Otto and Whitton, 2000; Uyeno and Smith, 1972).

5. Conclusions

Our phylogenetic reconstruction of the Botiidae has revealed a major split into the monophyletic subfamilies Leptobotiinae and the Botiinae. Leptobotiinae include the genera *Leptobotia and Parabotia*, Botiinae include five lineages, which in general are congruent with the genera *Botia*,

Sinibotia, *Yasuhikotakia*, *Chromobotia*, and *Syncrossus*. The generic placement of some species has to be checked by additional morphological and molecular analyses. A division in the ploidy level of the fishes accompanies the split between Leptobotiinae and Botiinae: Leptobotiinae are diploid while Botiinae are evolutionarily tetraploid. The polyploidisation can be accounted on a single event.

Acknowledgments

For their help with the obtaining of the samples we would like to thank G. Ott, I. Seidel and J. Kula. M. Kottelat kindly checked our identification of the species of the genus *Syncrossus* and of *Y. sidthimunki* and *Y. nigrolineata*. R. Britz helped us by providing photographs of the syntypes of *B. rostrata* and literature. Further we would like to express our thanks to A. Perdices and L. Rüber, who gave valuable advices for the analyses of the data and to K.E. Witte for providing his unpublished sequences and primers. Earlier versions of the manuscript increased due to the comments of M. Kottelat, A. Perdices, O. Říčan and K.E. Witte. The study was supported by Grant No. 206/05/2556 and 2006/00/0668 of the Czech Science Foundation, Grants No. A600450508 and S5045111 of the Grant Agency of the Academy of Sciences of the Czech Republic, the IRP IAPG No. AV0Z50450515 and the grant of Ministry of Education, Youth and Sports FRVŠ No. 27788.

References

Alves, M.J., Coelho, M.M., Collares-Pereira, M.J., 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. Genetica 111, 375–385.

Chenuil, A., Galtier, N., Berrebi, P., 1999. A test of the hypothesis of an autopolyploid vs. allopolyploid origin for a tetraploid lineage: application to the genus *Barbus* (Cyprinidae). Heredity 82, 73–380.

Crespi, B.J., Fulton, N.J., 2004. Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. Mol. Phylogenet. Evol. 31, 658–679.

David, L., Blum, S., Feldman, M.W., Lavi, U., Hillel, J., 2003. Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci. Mol. Biol. Evol. 20, 1425–1434.

Fang, P.W., 1936. Study on the botiid fishes of China. Sinensia 7, 1–49.

Farris, J.S., Källerjö, M., Kluge, A.g., Butt, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Janko, K., Kotlík, P., Ráb, P., 2003. Evolutionary history of asexual hybrid loaches (*Cobitis*: Teleostei) inferred from phylogenetic analysis of mitochondrial DNA variation. J. Evol. Biol. 16, 1280–1287.

Kitagawa, T., Watanabe, M., Kitagawa, E., Yoshioka, M., Kashiwagi, M., Okazaki, T., 2003. Phylogeography and the maternal origin of the tetraploid form of the japanese spined loach, *Cobitis biwae*, revealed by mitochondrial DNA analysis. Ichthyol. Res. 50, 318–325.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. Proc. Nat. Acad. Sci. USA 86, 6196-6200.

Kottelat, M., 1994. *Vaillantella cinnamomea*, a new species of balitorid loach from eastern Borneo. Jpn. J. Ichthyol. 40, 427–431.

Kottelat, M., 2001. Fishes of Laos. WHT, Colombo.

Kottelat, M., 2004. *Botia kubotai*, a new species of loach (Teleostei: Cobitidae) from the Ataran River basin (Myanmar), with comments on botiine nomenclature and diagnoses of two new genera. Zootaxa 401, 1–18.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Arizona State University, Tempe, Arizona, USA.

Le Comber, S.C., Smith, C., 2004. Polyploidy in fishes: patterns and processes. Biol. J. Linn. Soc. 82, 431–442.

Leggatt, R.A., Iwama, G.K., 2003. Occurrence of polyploidy in the fishes. Rev. Fish Biol. Fisheries 13, 237–246.

Ludwig, A., Belfiore, N.M., Pitra, C., 2001. Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics 158, 1203–1215.

Mickevich, M.F., Farris, J.S., 1981. The implications of congruence in *Menidia*. Systematic Zool. 30, 351–369.

Nalbant, T.T., 1963. A study of the genera of Botiinae and Cobitinae (Pisces, Ostariophysi, Cobitidae). Trav. Mus. Hist. Nat. "Grigore Antipa" 4, 343–379.

Nalbant, T.T., 2002. Sixty million years of evolution. Part one: Family Botiidae (Pisces: Ostariophysi: Cobitoidea). Trav. Mus. Hist. nat."Grigore Antipa" 44, 309–344.

Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Ann. Rev. Genet. 34, 401–437.

Palumbi, S.R., Martin, A., Romano, S., McMillan, W.O., Stice, L., Grabowski, G., 1996. The Simple Fool's Guide to PCR. Kewalo Marine Laboratory and Univ. of Hawaii.

Perdices, A., Doadrio, I., 2001. The molecular systematics and biogeography of the European Cobitids based on mitochondrial DNA sequences. Mol. Phylogenet. Evol. 19, 468–478.

Phillips, R.B., Ráb, P., 2001. Chromosome evolution in the Salmonidae (Pisces): an update. Biol. Rev. (Cambridge) 76 (1), 1–25.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitutions. Bioinformatics 14, 817–818.

Rainboth, W.J., 1996. FAO Species Identification Field Guide for Fisheries Purposes. Fishes of the Cambodian Mekong. FAO, Rome.

Roberts, T., 1989. The freshwater fishes of western Borneo (Kalimantan Barat, Indonesia). Mem. Calif. Acad. Sci. 14, 1–210.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

Sawada, Y., 1982. Phylogeny and zoogeography of the superfamily Cobitoidea (Cyprinoidei, Cypriniformes). Mem. Fac. Fish. Hokkaido University 28, 65–223.

Soltis, D.E., Soltis, P.S., 1995. The dynamic nature of polyploid genomes. Proc. Natl. Acad. Sci. USA 92, 8089–8091.

Soltis, D.E., Soltis, P.S., 1999. Polyploidy: recurrent formation and genome evolution. Trends Ecol. Evol. 14, 348–352.

Suzuki, A., 1996. Chromosomal evolution of the suborder Cyprinoidei, with special reference to topical problems of phylogenetic relationships of their families. Cytobios 87, 169–179.

Suzuki, A., Taki, Y., 1996. Tetraploidisation in the cobitid subfamily Botinae (Pisces, Cypriniformes). Cytobios 85, 229–245.

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

Taki, Y., 1972. *Botia eos*, a new spiny loach from Thailand and Laos, with notes on some related forms in Asia. Jpn. J. Ichthyol. 19, 63–81.

Tang, Q., Xiong, B., Yang, X., Liu, H., 2005. Phylogeny of the East Asian botiine loaches (Cypriniformes, Botiidae) inferred from mitochondrial cytochrome *b* sequences. Hydrobiologia 544, 249–258.

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. 24, 4876–4882.

Tsigenopoulous, C.S., Ráb, P., Naran, D., Berrebi, P., 2002. Multiple origins of polyploidy in the phylogeny of southern African barbs (Cyprinidae) as inferred from mtDNA markers. Heredity 88, 466–473.

Ueno, K., Nagase, A., Ye, Y.J., 1988. Tetraploid Origin of the Karyotype of the Asian Sucker, *Myxocyprinus asiaticus*. Jpn. J. Ichthyol. 34, 512–514.

Uyeno, T., Smith, G.R., 1972. Tetraploid origin of karyotype of catostomid fishes. Science 175, 644–646.

Yang, J.X., Chen, Y.R., 1992. Revision of the subgenus *Botia* (*Sinibotia*) with description of a new species (Cypriniformes: Cobitidae). Ichthyol. Explor. Freshwaters 2, 341–349.