ZOOLOGICAL SCIENCE 23: 893–901 (2006)

Phylogenetic Analysis of *Hampala* Fishes (Subfamily Cyprininae) in Malaysia Inferred from Partial Mitochondrial Cytochrome *b* DNA Sequences

Jeffrine R. J. Ryan[1]* and Yuzine B. Esa[2, 3]

[1]*Institute for Biological Diversity, Bukit Rengit, Department of Wildlife and National Parks, 28500 Lanchang, Pahang, Malaysia*
[2]*Molecular Ecology Laboratory, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia*
[3]*Genetics Laboratory, Biology Department, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia*

This study examined 396 base pairs of the mitochondrial cytochrome *b* gene from 110 individuals belonging to the genus *Hampala*, a group of freshwater cyprinids that inhabit Southeast Asia. The samples were taken from various locations throughout Sarawak, Sabah, and peninsular Malaysia. The nucleotide sequences were subjected to phylogenetic analyses by using the neighbor-joining, maximum parsimony, and maximum likelihood methods. All three methods revealed the reciprocally monophyletic relationship of *Hampala macrolepidota* to the other *Hampala* forms, thus strongly supporting its status as a distinct species. Phylogenetic analysis also discovered the existence of two *H. bimaculata* lineages endemic to Borneo: (1) a newly identified species from the southern and central part of Sarawak assigned as *H. bimaculata* Type A and (2) the previously described *H. bimaculata* from northern Sarawak and the west coast of Sabah assigned as *H. bimaculata* Type B. However, the status of *H. sabana* and an intermediate form were not elucidated. The results suggest that the intermediate form from the Tawau population is actually a subpopulation of *H. sabana*, while the highly divergent intermediate form from Kalabakan could represent a cryptic species. The sharing of *H. macrolepidota* haplotypes in the southern peninsular Malaysia and southern and central Sarawak samples (Hm1 and Hm2) reflected the recent disconnection of the two regions, during the late Pleistocene. Overall, the partial sequence of the mitochondrial cytochrome *b* gene was useful for resolving the phylogenetic relationships among *Hampala* fishes in Malaysia.

Key words: Cyprinidae, *Hampala*, molecular systematics, cytochrome *b*, DNA sequencing

INTRODUCTION

Hampala is one of the primary freshwater cyprinid fish genera in Malaysia. It is widely distributed in Southeast Asia (peninsular Malaysia, Sumatra, Java, Borneo, and Indo-China). *Hampala* species are usually abundant, and individuals are large enough (maximum total length ~700 mm) to be of interest as food fishes. One characteristic that attracts our interest to this genus is that it shows much geographical variation in coloration and morphological traits. There are currently seven described species of *Hampala* from Southeast Asia: *H. ampalong* (Bleeker) from Brunei and Indonesia (Sumatera) (Kottelat *et al.*, 1993); *H. bimaculata* (Popta) from Borneo (Kottelat *et al.*, 1993); *H. dispar* (Smith) from Cambodia (Rainboth, 1996); *H. lopezi* (Herre) from the Philippines (Herre, 1924); *H. macrolepidota* (Valenciennes), the most widespread *Hampala* form, from throughout the region (Roberts, 1989); *H. sabana* (Inger and Chin, 1962) from North Borneo (Lim and Wong, 1994); and *H. salweenensis* (Doi and Taki) from Thailand (Doi and Taki, 1994). However, only three *Hampala* species are currently described from Malaysia. A single species, *H. macrolepidota*, has been recorded from peninsular Malaysia (Mohsin and Ambak, 1983; Roberts, 1989), and at least another two species have been recognized in Borneo: *H. bimaculata* (Kottelat *et al.*, 1993) in central and northern Sarawak and the west coast of north Borneo (Sabah), and *H. sabana* (Lim and Wong, 1994) from the Kinabatangan basin and the streams draining into Sandakan Harbor (Inger and Chin, 1962). Surprisingly, *H. ampalong* described from neighboring Brunei (Kottelat *et al.*, 1993) has never been collected from the other parts of Borneo, reflecting its rarity and limited distribution.

Fish systematics has historically depended on morphological techniques for identifying groups of evolutionarily related species (Stepien and Kocher, 1997). However, recently the development of molecular techniques has helped to invigorate the study of fish systematics (Stepien

* Corresponding author. Phone: +6-9-276-2009;
Fax : +6-9-276-1757;
E-mail: jeffrine@wildlife.gov.my
doi:10.2108/zsj.23.893

and Kocher, 1997). Systematics includes the discipline of taxonomy, as well as phylogenetic analysis. DNA sequencing is one of the techniques that has been applied widely in phylogenetic studies. A few studies have incorporated molecular phylogenetic analysis in determining the relationships among freshwater fishes (*e.g.*, Dodson *et al.*, 1995; Chen *et al.*, 1998; Wang *et al.*, 2000). The mitochondrial cytochrome *b* (cyt *b*) gene is perhaps the best studied DNA segment in fishes, and several studies have shown its usefulness in elucidating evolutionary patterns in fishes (*e.g.*, Kocher *et al.*, 1989; Meyer *et al.*, 1990; Carr and Marshall, 1991; Block *et al.*, 1993; Zhu *et al.*, 1994; Carr *et al.*, 1995). It has been well demonstrated that this gene retains a history of past isolation (Avise, 1989; Bilington and Hebert, 1991).

Very little taxonomic work had been done to systematically sort out Malaysian freshwater fishes for proper reference. The systematics of *Hampala*, for example, has not been completely resolved. Although Inger and Chin (1962) elaborated much concerning the diet, habitat, morphological description, and distribution of species of *Hampala*, these authors focused only on the Sabah region. Thus, confusion remains in the classification of this genus, especially in dealing with specimens from other regions.

In Malaysia, there is a lack of molecular studies to verify classical work on fish taxonomy. Our study represents a fresh attempt to assess the phylogenetic relationships among different forms of *Hampala* by sequence analysis of partial cyt *b* sequences. Our objectives were to (1) investigate the phylogenetic relationships among different forms of *Hampala* in Malaysia, and (2) clarify of the taxonomic status of Malaysian *Hampala* species.

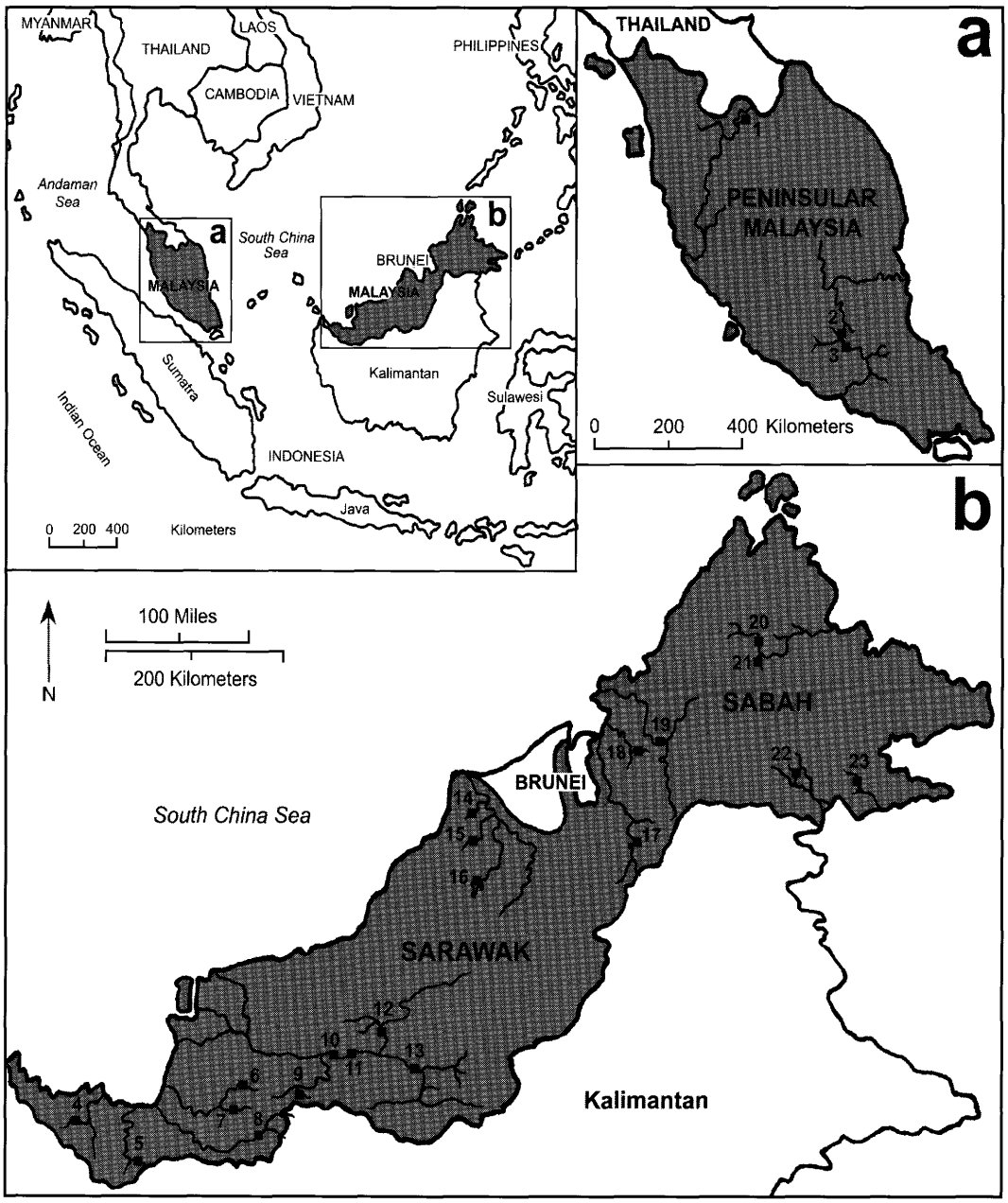

Fig. 1. Sampling localities of *Hampala* species throughout Malaysia (see Table 1 for detailed information on each locality).

MATERIALS AND METHODS

Sample collection

A total of 110 fish of various forms of *Hampala (H. bimaculata, H. macrolepidota*, and *H. sabana)* were caught from 23 locations throughout Sarawak, Sabah, and peninsular Malaysia (Fig. 1; Table 1) and were analyzed. Whole fish were collected and either preserved in 95% ethanol or frozen at –80°C. In some cases, only a small portion of tissue was available during sampling, and most identification was done prior to preservation. Out of 110 individuals

Table 1. List of Species ID, No. of Individual (n), Map point, Location (river, district, state), Habitat, Abbreviation (Abbr.), No. of Haplotypes (Nhap), Haplotype Diversity (*h*), Nucleotide Diversity (Pi, %) and GenBank Accession No. for each population in the present study.

Species ID	n	Map point	Location: River, District, State	Location: Geographical Region	Habitat	Abbr.	NHap	*h*	Pi	GenBank Accession No.
H. macrolepidota	6	1	Temenggor Lake, Perak	Northern P. Malaysia	Man-made lake; clear water; slow moving current; mixed dipterocarp forest vegetation	HmPB	2	0.533	0.135	AY697351 – AY697356
H. macrolepidota	5	2	Serting River, Negeri Sembilan	Southern P. Malaysia	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmS	2	0.4	0.101	AY697346 – AY697350
H. macrolepidota	3	3	Jempol River, Negeri Sembilan	Southern P. Malaysia	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmJ	1	–	–	AY697343 – AY697345
H. macrolepidota	1	4	Siniawan River, Bau, Kuching	Southern Sarawak	Middle stream; slow moving current; agricultural vegetation such as palm and rubber plantation	HmB	1	–	–	AY697301
H. macrolepidota	5	5	Krang River, Balai Ringin, Serian	Southern Sarawak	Brackish water, slow moving current; swamp vegetation	HmBR	2	0.4	0.101	AY697302 – AY697306
H. macrolepidota	2	6	Layar River, Betong	Southern Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmBL	1	–	–	AY697319 – AY697320
H. macrolepidota	2	7	Spak River, Betong	Southern Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmSP	1	–	–	AY697321 – AY697322
H. macrolepidota	12	8	Batang Ai River, Sri Aman	Southern Sarawak	Man-made lake; clear water; mixed dipterocarp forest vegetation	HmBA	3	0.591	0.302	AY697307 – AY697318
H. macrolepidota	2	10	Lan River, Kapit	Central Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmL	2	1	0.253	AY697325 – AY697326
H. macrolepidota	5	11	Menuan River, Kapit	Central Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmM	1	–	–	AY697327 – AY697331
H. macrolepidota	2	12	Pelagus River, Kapit	Central Sarawak	Upper streams water with sandy, gravel and rocky bottom; rapids	HmP	2	1	0.505	AY697332 – AY697333
H. macrolepidota	2	13	Ulu Baleh River, Kapit	Central Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HmUB	1	–	–	AY697323 – AY697324
H. macrolepidota	2	14	Bakong River, Baram	Northern Sarawak	Middle stream; slow moving current; flood plain	HmBK	2	1	0.253	AY697334 – AY697335
H. macrolepidota	2	15	Tinjar River, Baram	Northern Sarawak	Middle stream; slow moving current; flood plain	HmT	1	–	–	AY697336 – AY697337
H. macrolepidota	5	16	Loagan Bunut Oxbow Lake, Baram	Northern Sarawak	Natural lake; brackish water; slow moving current	HmLB	2	0.4	0.707	AY697338 – AY697342
H. macrolepidota	3	19	Ulu Padas River, Beaufort	West coast Sabah	Middle streams water with sandy, gravel and rocky bottom; moderately fast current.	HmBE	1	–	–	AY697357 – AY697359
H. bimaculata Type A	4	8	Batang Ai River, Sri Aman	Southern Sarawak	Man-made lake; clear water; slow moving current; mixed dipterocarp forest vegetation	HbABA	3	0.833	0.253	AY697360 – AY697363
H. bimaculata Type A	5	9	Bloh River, Lanjak Entimau	Central Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HbALE	2	0.4	0.404	AY697364 – AY697368
H. bimaculata Type A	7	13	Ulu Baleh River, Kapit	Central Sarawak	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HbAUB	2	0.286	0.144	AY697369 – AY697375
H. bimaculata Type B	3	18	Iti River, Sipitang	West coast Sabah	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest vegetation	HbBS	1	–	–	AY697383 – AY697385
H. bimaculata Type B	7	17	Ba' Kelalan, Lawas	Northern Sarawak	Moderately clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; sub-montane vegetation	HbBB	1	–	–	AY697376 – AY697382
H. sabana	1	20	Liwagu River, Ranau	Central eastern Sabah	Moderately clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; sub-montane vegetation	HsL	1	–	–	AY697405
H. sabana	5	21	Susuban River, Ranau	Central eastern Sabah	Clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; sub-montane vegetation	HsS	1	–	–	AY697406 – AY697410
Hampala intermediate	16	22	Uyun River, Kalabakan	Southeastern Sabah	Moderately clear upper streams water with sandy, gravel and rocky bottom; relatively fast current; mixed dipterocarp forest	HiU	2	0.125	0.063	AY697389 – AY697404
Hampala intermediate	3	23	Balung River, Tawau	Southeastern Sabah	Middle stream; slow moving current; agricultural vegetation such as palm and rubber plantation	HiB	2	0.667	0.168	AY697386 – AY697388

caught, only 80 individuals were analyzed by measurement and examination of selected characters, as was done by Inger and Chin (1962). Whole fish were identified with the keys of Inger and Chin (1962), Mohsin and Ambak (1983), Roberts (1989), and Kottelat *et al.* (1993). Voucher specimens of each species identified were deposited at the Faculty of Resource Science and Technology, Universiti Malaysia Sarawak Zoological Museum.

DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from about 1–2 mm^3 of fresh and ethanol-preserved tissue samples following a modified CTAB (cetyltrimethylammonium bromide) procedure as described by Grewe *et al.* (1993). Amplifications were carried out in a thermal cycler (Biometra). Optimization of PCR was done using 25 µl reaction volumes containing 2 µl DNA (~15–20ng), 1×PCR buffer, 0.2mM dNTPs (Promega), 1.5mM $MgCl_2$, 1.25 pmol of each primer, and 0.05 U *Taq* polymerase (Promega). Two cyt *b* primers were used: GluDG-L, 5'-TGACTTGAARAACCAYCGTT G- 3' (Palumbi *et al.*, 1991) and CB2-H, 5'-CCCTCAGAATGATATTTGTCCTCA- 3'. Amplification was done using the following PCR profile: a preliminary denaturation at 96°C for 5 min followed by 25 cycles of 95°C for 45 sec, 47°C for 45 sec, and 72°C for 45 sec. This was followed by a final extension period at 72°C for 7 min before the samples were cooled to 4°C. Each PCR product was run on 1% agarose gel for confirmation of equal length against an appropriate size marker. The PCR products obtained were later purified using DNA purification kits (Fermentas and Promega) according to the manufacturers' instructions. Cycle sequencing was performed for 35 cycles of 96°C for 10 sec, 55°C for 5 sec, 60°C for 4 min, followed by cooling to 4°C. Sequencing reaction products were purified by precipitation with ethanol and sodium acetate. Sequencing for each sample was carried out only on the forward strand with an ABI PRISM®377 DNA Sequencer.

Sequence alignment and phylogenetic analyses

Multiple alignments of the nucleotide sequences were done with the program ClustalX 1.81 (Thompson *et al.*, 1997) and subsequently adjusted by eye. To estimate levels of inter- and intra-population genetic diversity, standard genetic diversity indices including the number of haplotypes, haplotype diversity (*h*) (Nei, 1987), and nucleotide diversity (Pi) (Nei, 1987) were calculated using the DNA polymorphism option implemented in DNA SP version 3.50 (Rozas and Rozas, 1999). Two analyses of patterns of geographical subdivision, F_{ST} (Hudson *et al.*, 1992) and pair-wise distance using the Kimura two-parameter model (Kimura, 1980), were done to estimate genetic distances among the different forms of *Hampala*. Calculations of F_{ST} were performed using DNA SP version 3.50 (Rozas and Rozas, 1999), while pair-wise distance analyses were performed using *MEGA2* (Kumar *et al.*, 2001).

Phylogenetic trees were constructed (analysis done using haplotypes) by using the unweighted maximum parsimony (MP) and neighbor-joining (NJ) methods implemented in *MEGA2* (Kumar *et al.*, 2001) and the maximum likelihood (ML) method implemented in PAUP* version 4.0b 10 (Swofford, 1999). The NJ clustering was performed using the Kimura 2-parameter substitution model (Kimura, 1980) with the complete deletion option, and the MP analysis was done using the close-neighbor-interchange (CNI) option. The ML analysis was conducted by using the puzzle method available in PAUP*. All trees inferred from the partial cyt *b* sequences were rooted with *Puntius binotatus* (Valenciennes) (Cyprinidae: spotted barb) and *Helostoma temmincki* (Cuvier) (Helostomatidae: kissing goramy) as outgroup taxa. Phylogenetic confidence was estimated by bootstrapping (Felsenstien, 1985) with 1,000 replicate data sets.

The nucleotide sequences were later translated into amino acid sequences by using the genetic code for vertebrate mtDNA. The *Barbus canis* sequence obtained from GenBank (Accession No. AF288486) was also aligned as the reference sequence. Finally, all the sequences were registered with GenBank under accession numbers AY697301–AY697412.

RESULTS

Morphological identification of samples

The samples caught from the 23 different locations showed much morphological variation in coloration and external-character counts (Table 2). Morphological measurements identified two forms of *H. bimaculata*, which differed in coloration and gill-raker counts (Table 2). One form collected from northern Sarawak and the west coast of Sabah corresponded to the *H. bimaculata* previously described by Inger and Chin (1962), while another newly identified form was collected from southern and central Sarawak. Thus, prior to phylogenetic analyses, we treated these as two different forms of *H. bimaculata*: (1) *H. bimaculata* Type A (southern and central Sarawak samples) and (2) *H. bimaculata* Type B (northern Sarawak and west coast Sabah samples). In contrast, *H. macrolepidota* was easily identified, comprising a single form distinct from the two *H.*

Table 2. Morphological description based on certain characters for each form of *Hampala* caught (adapted from Inger and Chin, 1962).

	Species/Form	Description	n	LLS	CS	BPR	GR (total)
1	*H. macrolepidota*	Having a single dark band below the dorsal fin	33	26–29	12	15–16	9–11
2	*H. sabana*	Having a single dark band below dorsal fin	6	30–32	13–15	13–14	9–10
3	*H. bimaculata* Type A (Southern and central Sarawak populations)	Having two clear dark spots/blotches; one below dorsal fin and one in front of caudal fin	10	26–28	12	14–15	12–13
4	*H. bimaculata* Type B (Northern Sarawak and west coast Sabah populations)	Having two moderately dark bands/spots; one below dorsal fin and one in front of caudal fin	8	26–28	12	14–15	10–11
5	Intermediate form (Tawau population)	Having one dark band below dorsal fin. A very faint spot in front of caudal fin is absent when fish length is above 45mm	3	27–28	12	14–15	10–11
6	Intermediate form (Kalabakan population)	Having one dark band below dorsal fin. A very faint spot in front of caudal fin is absent when fish length is above 45mm	20	27–29	12	14–15	12–13

(n= samples analyzed; LLS = Lateral Line Scales; CS = Circumpeduncular Scales; BPR = Branched Pectoral Rays; GR = Gill Raker)

bimaculata forms in coloration and gill-raker counts. The Sabah endemic, *H. sabana*, exhibited striking morphological differences from the other forms of *Hampala* by having higher counts of circumpeduncular scales (30–32) and lateral line scales (13–15), in keeping with Lim and Wong's (1994) taxonomic classification. Tawau and Kalabakan samples (southeastern Sabah), as described by Inger and Chin (1962), represent a form intermediate between *H. sabana* and *H. bimaculata* Type B. Morphological measurements showed differences in gill-raker counts between the intermediate-form samples from Kalabakan (gill-raker counts: 12–13) and Tawau (gill-raker counts: 10–11). Our morphological study did not identify any fish as *H. ampalong*; thus, this species was not included in the phylogenetic study. Overall, morphological measurements in this study generally correspond with the previously described *Hampala* species in Borneo (Inger and Chin, 1962; Kottelat *et al.*, 1993), but with the additional important findings of the newly identified form of *H. bimaculata* (Type A) in central and southern Sarawak and possibly a cryptic species in Kalabakan, Sabah.

Sequence analysis and genetic diversity

Aligned partial sequences of 396 bp of the mitochondrial cyt *b* gene from 110 *Hampala* individuals represented only 34.7% of the total length of the cyt *b* gene in fishes (*e.g.*, Briolay *et al.*, 1998; Perdices *et al.*, 2002). The base composition showed an anti-G bias (Table 3), which is characteristic for this mitochondrial gene (Cantatore *et al.*, 1994;

Table 3. Comparison of the nucleotide composition, haplotype diversity and nucleotide diversity (Pi; %) among the five forms of *Hampala* studied.

Forms	Percentage (%)				n	NHap	Haplotype Diversity, *h*	Nucleotide Diversity, Pi (%)
	A	T	C	G				
Hm	29.7	29.5	26.1	14.8	59	11	0.686	0.93
Hs	30.6	27.5	28	13.9	6	1	–	–
HbA	30.4	29.5	26.1	14	16	6	0.617	1.05
HbB	29.8	28.8	26.8	14.6	10	1	–	–
Hi	30.5	29.3	26.8	13.5	19	4	0.38	1.9
Overall	30	29.2	26.4	14.4	110	23	0.875	7.1

n= number of samples; Nhap= number of haplotypes; Hm= *Hampala macrolepidota*; Hs= *Hampala sabana*; HbA= *Hampala bimaculata* Type A; HbB= *Hampala bimaculata* Type B; Hi= *Hampala* intermediate form

Table 4. Pairwise distances (%, below the diagonal) and FST estimates (above the diagonal) among the geographical populations of *Hampala* analyzed.

		Hm						Hb				Hs	Hi	
		NPM	**SPM**	**SS**	**CS**	**NS**	**WCS**	**SS***	**CS***	**NS****	**WCS****	**CES**	**TW**	**KB**
Hm	**NPM**		0.89	0.87	0.85	0.75	0.97							
	SPM	0.9		0.01	0	0.73	0.99							
	SS	1	0.1		0.08	0.7	0.96							
	CS	0.9	0.1	0.2		0.71	0.97							
	NS	2.2	1.9	1.9	1.9		0.75							
	WCS	2.7	2.4	2.4	2.4	1.9								
Hb	**SS***	10.7	11.6	11.7	11.6	10.8	11.9		0.89	0.99	0.99			
	CS*	10.6	11.5	11.6	11.5	10.8	11.8	2.3		0.99	0.99			
	NS**	12.2	13.1	13.2	13.1	12.9	13.3	10.2	8.9		0			
	WCS**	12.2	13.1	13.2	13.1	12.9	13.3	10.2	8.9	0				
Hs	**CES**	13.4	14.3	14.4	14.3	13.5	14.6	8.3	7.7	8.2	8.2		0.75	0.99
Hi	**TW**	13.2	14.1	14.2	14.1	13.3	14.3	8.7	8.1	8.6	8.6	0.3		0.98
	KB	12.1	13	13.1	13	12.4	13.3	7.2	6.6	8.2	8.2	6.7	7.1	
Overall		n=59 NHap=11 Pi=0.930% *h*=0.686						n=16 NHap=6 Pi=1.054% *h*=0.617		n=10 NHap=1		n=6 NHap=1	n=19 Nhap=4 Pi=1.923% *h*=0.380	
								N=26 Nhap=7 Pi=4.554% *h*=0.686						

NPM=Northern Peninsular Malaysia; SPM= Southern Peninsular Malaysia; SS=Southern Sarawak; CS=Central Sarawak; NS=Northern Sarawak; WCS=West coast Sabah; CES=Central eastern Sabah; TW=Tawau; KB=Kalabakan

*=*H. sabana*; **=*Hampala* intermediate

Briolay *et al.*, 1998). The low G content (mean: 14.4%) and the almost identical A, C, and T contents (mean: 30.0, 26.4, and 29.2%, respectively) were similar to those previously reported for fish cyt *b* sequences (*e.g.*, Martin and Bermingham, 1998; Briolay *et al.*, 1998; Perdices *et al.*, 2004). Haplotype diversity (*h*) and nucleotide diversity (Pi) were also calculated for each population under study (Table 1) and for each of the *Hampala* forms (Table 3).

From the 396-bp sequence, 144 (36.4%) variable or polymorphic sites were observed, with most variation (68.1%) occurring at third-codon positions. In addition, among the 144 variable sites, 95 (66.0%) were parsimony-informative sites. The total number of mutations was 186, with most of the mutation events being transitions (60.8%). The pattern of nucleotide substitutions in *Hampala* species is very similar to that observed for other fishes (Martin, 1995). The rate of transition substitutions is at least an order of magnitude higher than the rate of transversion substitutions (data not shown). Protein translation of the 396-bp fragment revealed 132 acid amino residues, among which 31 (23.5%) were variable sites and 12 (38.7%) were parsimony-informative.

In total, 23 haplotypes were distinguished in the nucleotide sequence data set: 11 haplotypes for *H. macrolepidota*, six for *H. bimaculata* Type A, four for the intermediate form, and one each for *H. bimaculata* Type B and *H. sabana*. The average nucleotide diversity was relatively low for *H. macrolepidota* (0.9%), *H. sabana*, and the intermediate form (both 0%), but was high for *H. bimaculata* (4.5%). Nucleotide translations retained only 16 haplotypes in the amino acid sequences. Six haplotypes belonged to *H. macrolepidota* and five to *H. bimaculata*, while the *H. sabana* haplotype clustered together with the intermediate form haplotype (Hi2).

Pairwise genetic distances (number of nucleotide substitutions per site) among the geographical populations are shown in Table 4, and the summarized genetic distances among the five *Hampala* forms in Tables 5. The inter-population analysis revealed only small genetic differences between *H. macrolepidota* from southern peninsular Malaysia and southern and central Sarawak (0.1–1.0%; Table 4). The distance value of 9.2% (Table 5) indicated a high level of genetic divergence between *H. bimaculata* Type B and Type A, congruent with their high level of morphological differences. The pairwise genetic distances also revealed interesting findings within the Sabah samples. The intermediate form (described as an intermediate between *H. sabana* and *H. bimaculata* Type B) from Tawau exhibited high genetic divergence from the Kalabakan (7.1%; Table 4) and *H. bimaculata* Type B samples (8.6%; Table 4), but was very closely genetically similar to the *H. sabana* samples (0.3%; Table 4). In contrast, the intermediate-form samples from Kalabakan were genetically divergent from both *H. sabana* (6.7%; Table 4) and *H. bimaculata* Type B (8.2%; Table 4).

Table 5. A summarized of mean pairwise distance (%) among the five forms of *Hampala* analyzed.

	Hm	Hs	HbA	HbB	Hi	Pb	Ht
Hm							
Hs	14.1						
HbA	11.4	7.8					
HbB	13	8.2	9.2				
Hi	13	5.7	7	8.3			
Pb	15.5	18.1	17.9	17.5	17.8		
Ht	28	26.5	27.9	27.7	26.8	26.9	

Hm=*Hampala macrolepidota*; Hs=*Hampala sabana*; HbA=*Hampala bimaculata* Type A; HbBB=*Hampala bimaculata* Type B; Hi=*Hampala* intermediate form; Pb: *Puntius binotatus*; Ht: *Helostoma temmincki*.

Phylogenetic relationships and population structure

The phylogenetic relationships of *Hampala* constructed from the 23 haplotypes are summarized in Figs. 2 and 3. The NJ analysis produced a tree topology identical to that from the ML analysis (-ln L, unconstrained=1255.92874). Thus we combined the trees from both methods into a single tree represented by the NJ tree (Fig. 2). The tree from the unweighted MP analysis (tree length=244; CI= 0.762295; RI= 0.876858) revealed a similar topology (Fig. 3) with that from the NJ and ML methods, but with a few differences. The phylogenetic trees (nucleotide data set) obtained from all three methods divided the samples into two clades supported by high bootstrap values (>95% for all methods). Group 1 consisted of all the widespread *H. macrolepidota* haplotypes, while group 2 contained the Borneo endemic *Hampala*-form haplotypes: *H. sabana*, the intermediate form, and both *H. bimaculata* forms (Type A and Type B). In Group 1, there was a close relationship among the *H.*

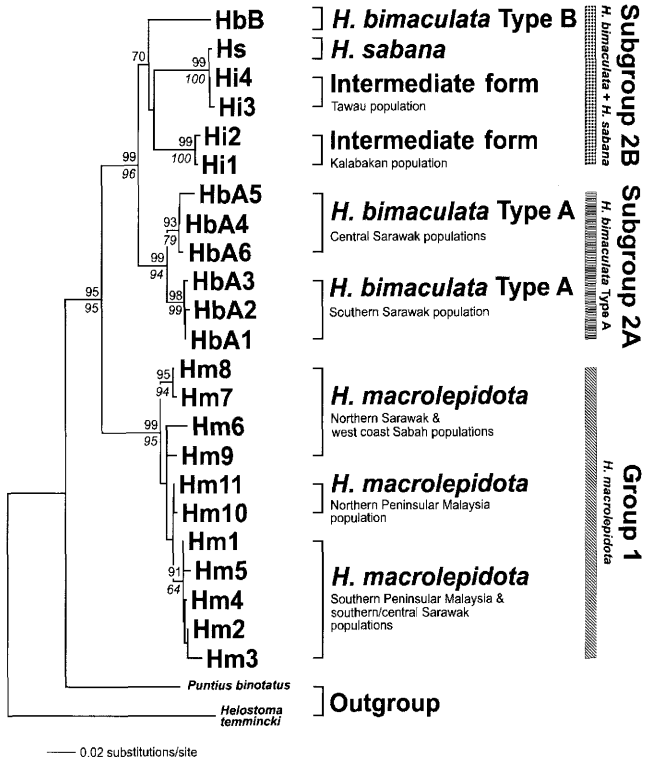

Fig. 2. *Hampala* phylogenetic relationships based on nucleotide sequences of part of the mitochondrial cyt *b* gene. The phylogeny shown is a single tree recovered using NJ analysis. The ML tree (-ln L= 1255.92874) was highly congruent and identified the same topology. Values above branches are NJ bootstrap estimates; those below branches are ML estimates, based on 1,000 replicates in both cases. Only bootstrap values>50% are shown.

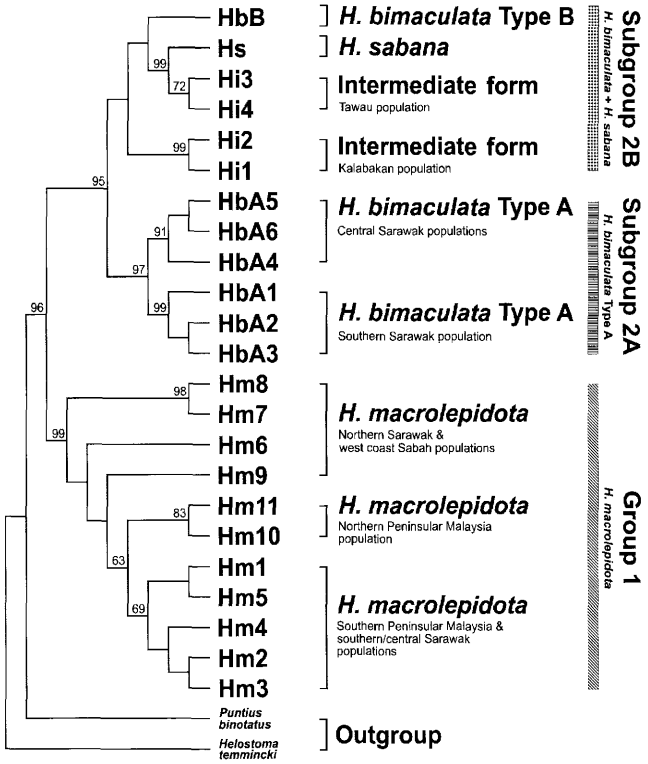

Fig. 3. Unweighted MP tree based on nucleotide sequences of part of the mitochondrial cyt *b* gene. Only bootstrap values>50% are shown.

macrolepidota haplotypes, with haplotypes from northern Sarawak and the west coast of Sabah forming the basal clade. The Group 2 clade was further divided into two subgroups supported by high bootstrap values (>95% for all methods). Subgroup 2A consisted of all the *H. bimaculata* Type A haplotypes, while subgroup 2B contained haplotypes of *H. bimaculata* Type B, *H. sabana*, and the intermediate form. Within subgroup 2B, both the NJ and ML methods placed *H. bimaculata* Type B as basal, whereas the MP analysis placed the intermediate-form haplotypes from Kalabakan as basal. Interestingly, all methods supported the grouping of *H. sabana* with the intermediate-form haplotypes from Tawau, with strong (99%) bootstrap values. Furthermore, the NJ and ML analyses showed that the intermediate-form haplotypes from Kalabakan comprised a clade separate from that of the Tawau haplotypes (99% bootstrap value), suggesting that these populations comprise distinct lineages.

The analysis of population structure indicated low genetic differentiation among *H. macrolepidota* samples from southern peninsular Malaysia, and southern and central Sarawak (F_{ST} values ranging from 0.000–0.079) (Table 4), while the other regions (northern Peninsular Malaysia, northern Sarawak, and the west coast of Sabah) showed apparent population subdivisions (F_{ST} ranging from 0.702–0.986). High geographical structuring was also found between the *H. bimaculata* Type A and Type B samples (F_{ST} values ranging from 0.985–0.986). Although no population structuring was found for *H. bimaculata* Type B between northern Sarawak and the west coast of Sabah, high population structuring was observed for *H. bimaculata* Type A between southern and central Sarawak (F_{ST}=0.889).

DISCUSSION

Taxonomy within the genus *Hampala*

Phylogenetic analysis of a partial cyt *b* fragment supported the reciprocally monophyletic relationship between the widely distributed *H. macrolepidota* and the other Bornean endemic *Hampala* forms (*H. bimaculata*, *H. sabana*, and the intermediate form). High genetic divergence separating *H. macrolepidota* from the other *Hampala* lineages (10.7–14.1%) confirmed its taxonomic status as a distinct species. The phylogenetic analysis also supported the existence of two *H. bimaculata* lineages in Malaysian Borneo. The newly identified lineage from southern and central Sarawak (*H. bimaculata* Type A) comprised a sister clade to the northern Sarawak lineage (*H. bimaculata* Type B). High genetic divergence between these two lineages (9.2%) and their morphological distinction in gill-raker counts support their status as distinct taxa. Therefore, a revision of the taxonomic status of *H. bimaculata* in Sarawak is suggested, with recognition of the newly identified lineage (*H. bimaculata* Type A) as a new species. We believe that the exclusion of *H. ampalong* sequences did not affect the significant findings of this study, particularly the identification of the new *H. bimaculata* mtDNA lineage, since the former species was morphologically distinct from the latter as well as from the other *Hampala* forms currently found in Borneo. Briefly, *H. ampalong* was easily identified by the presence of a series of dark spots anterior to the caudal base, while both *H. bimaculata* lineages were identified by the presence of a vertical black bar on each side below the dorsal midline and a second vertical bar or spot on the caudal peduncle (Kottelat *et al.*, 1993).

The phylogenetic relationship between *Hampala* from northern Sarawak and Sabah (subgroup 2B consisting of *H. bimaculata* type B, *H. sabana*, and the intermediate form) remains unclear, particularly regarding the basal group. All three methods supported a sister-group relationship between the undescribed intermediate form from Kalabakan and the Tawau population (genetic distance of 7.1%), suggesting their status as separate lineages. The genetic divergence was congruent with morphological differences in gill-raker counts (12–13 for Kalabakan and 10–11 for Tawau). Furthermore, the high genetic divergences between the Kalabakan lineage and the *H. macrolepidota* (12.8%), *H. bimaculata* Type B (8.2%), and *H. sabana* (6.7%) lineages clearly indicate its genetic distinctiveness from the other *Hampala* forms. Thus, the Kalabakan population could represent a distinct evolutionary unit for conservation (Moritz, 1994; Nielsen, 1995), possibly a cryptic species.

The close genetic relationship found between *H. sabana* and the intermediate-form population from Tawau (only 0.3% divergence) suggested a different scenario from that of the Kalabakan population mentioned above. It may be that the Tawau population is actually a subpopulation of *H. sabana*. Alternatively, the results may indicate recent secondary contact of the *H. sabana* lineage with the undescribed Tawau lineage due to breakdown in a geographical barrier to gene flow. This contact could have happened during Pleistocene glacial intervals, as observed by Dodson *et al.* (1995) in explaining the close relationship between cat-

fish species from Borneo and the Asian mainland. Another explanation could be that the intermediate-form population from Tawau represents a hybrid between *H. sabana* and *H. bimaculata* Type B, as was suggested by Inger and Chin (1962), exhibiting mtDNA characteristics very close to the former but morphologically similar to the latter. However, the mtDNA used in this study was insufficient to detect a hybridization event, and data from a nuclear gene are needed. Additionally, the length of the fragment used in this study was very short (only ~27% of the complete cyt *b* gene), and perhaps polymorphic sites that could differentiate the two species lie outside the fragment studied. Finally, plasticity in behavioral or morphological characters without any major genetic divergence (mutation) within a population could also explain the condition above. In order to adapt to various ecological habitats (sometimes very extreme conditions), morphological adaptation perhaps occurred without any changes in the genetic structure, as seen by Greenberg *et al.* (1998) in the coastal plain swamp sparrow. However, choosing the most parsimonious among these various hypotheses is difficult based on the current data alone. Further studies using more fish samples from other localities throughout the Tawau and Labuk-Segama regions and highly variable co-dominant nuclear markers such as microsatellites are needed to fully elucidate the taxonomic status, the occurrence of hybridization, and genetic boundaries between *Hampala* lineages in North Borneo.

In this study, the phylogenetic relationships among *Hampala* fishes were clarified, particularly between *H. macrolepidota* and *H. bimaculata*. Our results also clarified the taxonomic status of the two *H. bimaculata* forms in Sarawak (Type A from the southern and central regions and Type B from the northern region). We recommend a revision of the taxonomy of *H. bimaculata*, with the recognition of *H. bimaculata* Type A as a new species. The high levels of genetic and morphological differences of the Kalabakan population from the southeastern Sabah region suggest it is a newly identified cryptic species. However, a larger sample size (about 30 individuals per river system) from the watersheds of central and southeastern Sabah should be analyzed to give better insight into the phylogeny and phylogeography of *Hampala* species in these regions. The use of other, rapidly evolving mitochondrial genes (*e.g.*, control region) or the complete cyt *b* gene may also facilitate understanding the phylogenetic relationship between *H sabana* and the intermediate form from southeastern Sabah. Finally, the inclusion of the other four *Hampala* species (*H. ampalong*, *H. dispar*, *H. lopezi*, and *H. salweenensis*) in a future study will provide a better picture of the evolutionary history and radiation of fishes of the genus *Hampala* in Southeast Asia.

ACKNOWLEDGMENTS

We thank Dr. Robert F. Inger from the Field Museum of Natural History, Chicago, USA; Professor Tan Soon Guan from Universiti Putra Malaysia; and Associate Professor Dr. Mohd Tajuddin Abdullah from Universiti Malaysia Sarawak for valuable comments on the manuscript. We also thank the Fisheries Department of Malaysia; the staff of the Indigenous Fisheries Research Centre, Tarat, Sarawak; all the Unimas staff involved in the project; and all the local field guides and fishermen for their tremendous work and assistance during the field studies. This study was carried out with financial support from a Unimas Fundamental Research Grant (Project No. 245/2001(4)) and an ASEAN Regional Centre for Biodiversity Conservation Research Grant (Project No. RE-MYS002).

REFERENCES

Avise JT (1989) Gene trees and organismal histories: A phylogenetic approach to population biology. Evolution 43: 1192–1208

Billington N, Hebert PDN (1991) Mitochondrial DNA diversity in fishes and its implications for introductions. Can J Fish Aquat Sci 48 (Suppl 1): 80–94

Block BB, Finnerty JR, Stewart AFR, Kidd J (1993) Evolution of endothermy in fish: mapping physiological traits on a molecular phylogeny. Science 260: 210–214

Briolay J, Galtier N, Brito RM, Bouvet Y (1998) Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. Mol Phylogenet Evol 9: 100–108

Cantatore P, Roberti M, Pesole G, Ludovico A, Milella F, Gadaleta MN, Saccone C (1994) Evolutionary analysis of cytochrome *b* sequences in some perciformes: evidence for a slower rate of evolution than in mammals. J Mol Evol 39: 589–597

Carr SM, Marshall HD (1991) Detection of intraspecific DNA sequence variation in the mitochondrial cytochrome *b* gene of Atlantic cod (*Gadus morhua*) by the polymerase chain reaction. Can J Fish Aquat Sci 48: 48–52

Carr SM, Snellen AJ, Howse KA, Wroblewski JS (1995) Mitochondrial DNA sequence variation and genetic stock structure of Atlantic cod (*Gadus morhua*) from bay and offshore locations on the Newfoundland continental shelf. Mol Ecol 4: 79–88

Chen IS, Hsu CH, Hui CF, Shao KT, Miller PJ, Fang LS (1998) Sequence length and variation in the mitochondrial control region of two freshwater gobiid fishes belonging to *Rhinogobius* (Teleostei: Gobioidei). J Fish Biol 53: 179–191

Dodson JJ, Colombani F, Ng PKL (1995) Phylogeographic structure in mitochondrial DNA of a Southeast Asian freshwater fish, *Hemibagrus nemurus* (Siluroidei; Bagridae) and Pleistocene sea-level changes on the Sunda shelf. Mol Ecol 4: 331–346

Doi A, Taki Y (1994) A new cyprinid fish, *Hampala salweenensis*, from the Mae Pai River system, Salween Basin, Thailand. Jpn J Ichthyol 40: 405–412

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791

Greenberg R, Cordero PJ, Droege S, Fleischer RC (1998) Morphological adaptation with no mitochondrial DNA differentiation in the coastal plain swamp sparrow. Auk 115: 706–712

Grewe PM, Krueger CC, Aquadro CF, Bermingham E, Kincaid HL, May B (1993) Mitochondrial variation among lake trout (*Salvelinus namaycush*) strains stocked into Lake Ontorio. Can J Fish Aquat Sci 50: 2397–2403

Herre AWCT (1924) Distribution of the true freshwater fishes in the Philippines. I. The Philippine Cyprinidae. Philippines J Sci 24: 249–307

Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. Genetics 132: 583–589

Inger RF, Chin PK (1962) The Freshwater Fishes of North Borneo. Natural History Museum, Chicago

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120

Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Villablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in mammals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA 86: 6196–6200

Kottelat M, Whitten AJ, Kartokasari SN, Wirjoratmodjo S (1993) Freshwater Fishes of Western Indonesia and Sulawesi. Berkeley Book Pte Ltd, Singapore

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) *MEGA2*: Molecular

Evolutionary Genetics Analysis Software. Arizona State University, Tempe

Lim KKP, Wong A (1994) Fishes of Kinabatangan Basin, Sandakan District, Sabah, East Malaysia. Sabah Mus J 1: 39–71

Martin AP, Bermingham E (1998) Systematics and evolution of lower Central American cichlids inferred from analysis of cytochrome *b* gene sequences. Mol Phylogenet Evol 9: 192–203

Martin AP (1995) Mitochondrial DNA sequence evolution in sharks: rates, patterns, and phylogenetic inferences. Mol Biol Evol 12: 1114–1123

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature 347: 550–553

Mohsin AKM, Ambak MA (1983) Freshwater Fishes of Peninsular Malaysia. Universiti Pertanian Malaysia, Selangor

Moritz C (1994) Defining "evolutionary significant units" for conservation. Trends Ecol Evol 9: 373–375

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York

Nielsen JL (1995) Evolution and the Aquatic Ecosystem: Defining Unique Units in Populations Conservation. American Fisheries Society, Symposium 17. American Fisheries Society, Bethesda

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G (1991) The Simple Fool's Guide To PCR. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu

Perdices A, Cunha C, Coelho MM (2004) Phylogenetic structure of *Zacco platypus* (Teleostei, Cyprinidae) populations on the upper and middle Chang Jiang (=Yangtze) drainage inferred from cytochrome *b* sequences. Mol Phylogenet Evol 31: 192–203

Perdices A, Doadrio I, Economidis PS, Bohlen J, Bănărescu P (2002) Pleistocene effects on the European freshwater fish fauna: double origin of the cobitid genus *Sabanejewia* in the Danube basin (Osteichthyes: Cobitidae). Mol Phylogenet Evol 26: 289–299

Rainboth WJ (1996) Fishes of the Cambodian Mekong. FAO Species Identification Field Guide for Fishery Purposes, FAO, Rome

Roberts TR (1989) The Freshwater Fishes of Western Borneo (Kalimantan Barat, Indonesia). California Academy of Sciences, San Francisco

Rozas J, Rozas R (1999) DNASP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15: 174–175

Stepien CA, Kocher TD (1997) Molecules and morphology in studies of fish evolution. In "Molecular Systematics of Fishes" Ed by TD Kocher, CA Stepien, Academic Press, San Diego, pp 1–11

Swofford DL (1999) PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4.0b 10. Sinaeur Associates, Sunderland, Massachusetts

Thompson JD, Gibson TJ, Plewniak F, Jeanmourgin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by the quality analysis tools. Nucleic Acids Res 24: 4876–4882

Wang JP, Hsu KC, Chiang TY (2000) Mitochondrial DNA phylogeography of *Acrossocheilus paradoxus* (Cyprinidae) in Taiwan. Mol Ecol 9: 1483–1494

Zhu D, Jamieson BGM, Hugall A, Moritz C (1994) Sequence evolution and phylogenetic signal in control region and cytochrome *b* sequences of rainbow fishes (Melanotaeniidae). Mol Biol Evol 11: 672–676

(Received October 16, 2005 / Accepted May 28, 2006)