

Research Article

A DNA Barcode-Based Evaluation of the Southeast Asian Catfish Genus *Hemibagrus* Bleeker, 1862 (Teleostei: Siluriformes; Bagridae)

Julian J. Dodson¹ and Frédéric Lecomte²

¹Département de Biologie, Université Laval, 1045 Avenue de la Médecine, Québec, QC, Canada G1V 0A6

²Direction Générale de l'Expertise sur la Faune et ses Habitats-Secteur de la Faune, Ministère des Forêts, de la Faune et des Parcs, 880 chemin Ste-Foy, 2e étage, Québec, QC, Canada G1S 4X4

Correspondence should be addressed to Julian J. Dodson; julian.dodson@bio.ulaval.ca

Received 25 July 2014; Revised 1 December 2014; Accepted 22 December 2014

Academic Editor: Bernd Schierwater

Copyright © 2015 J. J. Dodson and F. Lecomte. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Species of the genus *Hemibagrus* are large river catfishes found throughout South-east Asia. The complexity of the region's biogeographical history and the lack of well-defined morphological characters render the taxonomy and phylogenetic reconstruction of *Hemibagrus* problematical. Early molecular studies of the *H. nemurus* species group revealed extensive genetic subdivisions, the taxonomic status of which remained unclear. A recent, morphologically-based, revision of the genus provides an opportunity to clarify the taxonomic status of these lineages. We employ a DNA barcode derived from the mitochondrial cytochrome *b* gene to expand our genetic analyses of the genus and to test the congruence of morphologically and genetically based taxonomies. Secondly, we evaluate phylogenetic relationships among taxa. Thirdly, we describe the phylogeography of *Hemibagrus* in South-east Asia. The species groups and nominal species proposed in the morphology-based revision generally reflect a hierarchy of monophyletic groups based on phenetic and maximum likelihood reconstructions of mtDNA phylogenies. The most notable exception involves the definition of a morphologically cryptic group from North Borneo. *H. nemurus* from West Java appears to be a regional population of *H. capitulum*. The phylogeography of the genus has been principally influenced by the formation of North Borneo and the emergence of the Sunda Islands.

1. Introduction

Southeast Asia represents a unique biogeographical situation, combining complex geological dynamics with extraordinary species richness [1]. The Sunda continental shelf unites mainland Asia and the Thai-Malay Peninsula with the Greater Sunda Islands (Sumatra, Java, and Borneo). The shelf is no more than 100 meters in depth and most of it is much shallower. During the early and middle Miocene, from 24 to 13 Ma ago, sea levels were typically 100–150 m above present day sea levels [2]. During these 11 Ma of generally high sea levels, there were 3 significant marine regressions, 2 of about 100 ka and one of 1.0 Ma duration. In the early Pliocene, sea levels rose 90–100 m and stayed there for nearly 1.4 Ma, from 5.5 to 4.2 Ma ago [2]. These early, periodic marine transgressions would have been sufficient to isolate Borneo

from the Thai-Malay Peninsula and isolate the peninsula from the mainland, providing many opportunities for cladogenesis and allopatric speciation. As these periods of marine transgressions were followed by periods of regressions, there were also opportunities for dispersal and secondary contact among species and lineages. During the more recent Pleistocene glaciations, it is generally accepted that sea level was between 100 and 200 m lower than today [3]. Great areas of the Sunda Shelf were thus converted to land traversed by rivers during periods of glacial maxima, thus facilitating relatively recent exchanges of the freshwater fish fauna [4, 5]. The contemporary distribution patterns of species have thus been shaped largely by complex pre-Pleistocene dispersal and vicariance events, whereas more recent changes in the connectivity of islands and the mainland have influenced the partitioning of intraspecific variation [1].

Species of the genus *Hemibagrus* [6] are large catfishes found in rivers east from India's Godavari River and south from the Yangtze in China. They occupy a wide range of habitats ranging from near estuarine to high altitude fast-flowing headstreams. The genus reaches its greatest diversity in Southeast Asia [7]. Throughout its geographical range, and particularly in Southeast Asia, species of *Hemibagrus* are valuable food fishes and species identification is of paramount importance to sustainably exploit this species complex. However, the taxonomy of *Hemibagrus* is confusing and the validity of many nominal species is unclear. Mo's [6] rehabilitation of the genus *Hemibagrus* was not accompanied by a comprehensive examination of types of all the nominal species. The complexity of the region's biogeographical history coupled with a lack of well-defined morphological characters and considerable plasticity in phenotypic traits commonly used for diagnosing species in other catfish groups renders the identification and phylogenetic reconstruction of *Hemibagrus* problematical [7]. P. K. L. Ng and H. H. Ng [8] divided the Sundaic *Hemibagrus* into four species groups, but subsequent work by Ng and coworkers revealed that the morphological characters they used may not always distinguish the species reliably, nor do they allow easy assignment of many of the species to any of the groups [7]. This situation led to the publication of a revision of the genus by Ng and Kottelat [7]. The revision is based on morphological criteria and recognizes eight species groups: *H. baramensis*, *H. guttatus*, *H. menoda*, *H. nemurus*, *H. olyroides*, *H. planiceps*, *H. pluriradiatus*, and *H. wyckii* species groups. A species group, as defined in Tan and Kottelat [9], is "an assemblage of species sharing a set of diagnostic characters, which may or may not be a monophyletic lineage." Ng and Kottelat [7] employed various morphometric and meristic characters to define species and species groups, including various body dimensions, fin lengths, head dimensions, eye diameter, and various fin ray counts (see [7, Figure 1] for a complete list of morphological variables used by Ng and Kottelat). The taxonomic revision of Ng and Kottelat [7] was based on the diagnosis of unique combinations of characters rather than discrete autapomorphies. Thirty-two valid species are recognized in the revision, three of which are new [7].

Early molecular phylogenetic and phylogeographic studies of the putative *H. nemurus* species group in Southeast Asia revealed extensive genetic subdivision of the group [10, 11]. The occurrence of genetically divergent groups in sympatry in widely separated locations supported the proposition that low sea levels promoted the dispersal and mingling of some genetic groups [10]. These studies also revealed very divergent lineages within *Hemibagrus*, predating the Pleistocene [10, 11]. However, in the absence of a comprehensive examination of types of all the nominal species in the genus, the taxonomic status of these genetic lineages remained unclear. The revision of the genus proposed by Ng and Kottelat [7] provides a basis by which we may clarify the taxonomic status and the biogeography of independent genetic lineages within *Hemibagrus*.

The analysis of single-locus data from genomic regions, such as short sequences of mitochondrial genes used as DNA barcodes [12], represents a useful tool for species

identification and evaluating taxonomic diagnoses based on morphological characters (e.g., [13, 14]). Given the phenotypic plasticity inherent in many of the morphological variables used by Ng and Kottelat in their revision [7] and the possibility that species groups so defined may not represent true monophyletic lineages, we here employ a DNA barcode derived from the mitochondrial cytochrome *b* gene (Cyt *b*) to expand our genetic analyses of the genus in Southeast Asia and to evaluate the revision of the genus *Hemibagrus* proposed by Ng and Kottelat [7] by testing the congruence of the morphologically and genetically based taxonomies. We chose this marker to be consistent with earlier work done on the genus in Southeast Asia [10, 11]. Our principal objective was to test the hypothesis that species groups and nominal species defined with morphological criteria by Ng and Kottelat [7] reflect a hierarchical arrangement of monophyletic groups based on the similarity of mtDNA sequences. Our second objective was to provide a tentative evaluation of phylogenetic relationships among taxa. Our third objective was to describe the geographical distribution of taxa and to relate major cladogenetic events with major geological events that may have contributed to the radiation of *Hemibagrus* in Southeast Asia. We pay particular attention to the phylogeography of the widespread *H. nemurus* species group.

2. Materials and Methods

A total of 452 hemibagrid catfish, particularly those of the presumed *H. nemurus* species group, were sampled throughout their distribution range in Southeast Asia and successfully sequenced, encompassing 67 sampling sites located on 37 river systems (Figure 1, Table 1). Sampling of the other species groups defined by Ng and Kottelat [7] was sparser; thus not all species groups were represented in the genetic analysis. Fish were principally obtained from local fishermen. Many fish were placed in the Zoological Reference Collection (ZRC) of the Raffles Museum of Biodiversity Research, University of Singapore, where they received a ZRC accession number (see Table S1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2015/490158>). Ethanol-preserved tissue samples of many of the specimens identified in Table S1 are freely available upon request from JJD.

Genetic Analysis. Total genomic DNA was extracted from 100 μg of tissue (fresh or 95% ethanol-preserved) with a standard phenol-chloroform-isoamyl protocol [15], precipitated with cold 95% ethanol and suspended in 200 μL TE buffer (10 nM Tris-HCl, pH 8; 1 mM EDTA). A 605 bp portion from the 5'-end of cytochrome *b* was amplified (including 8-bp located before the start codon) using derivatives of the universal primers of Kocher et al. [16]: P1 5'-AAAACC-ACCGTTGTTATTCAACTACA-3' (light strand) and P2 5'-GGGTTGTTTGATCCTGTTTCGTG-3' (heavy strand). Polymerase chain reaction (PCR) conditions were carried out in 50 μL total volumes containing (final concentrations) 200 μM of each of dATP, dGTP, dCTP, and dTTP; 400 nM of each primer; 1–2.5 U of Taq DNA polymerase; 2.5 mM MgCl_2 ; 10 mM Tris-HCl (pH 8.3); 50 mM KCl; and between 50 and 200 ng (3–5 μL) of template (purified mtDNA). DNA was

TABLE 1: Rivers, river code (as presented in Figure 1), sampling sites, latitude/longitude, and distance upstream from river's mouth of sampling sites and the number of specimens of *Hemibagrus* spp. sequenced.

River code	Site code	River	Sample site	Latitude d° mm' ss''	Longitude d° mm' ss''	Distance upstream from river mouth (km)	Number of specimens
1a	1	Mekong	Bungkan	18 21 54 N	103 39 04 E	1285	4
1b	2	Mekong	Bannan	17 58 00 N	104 14 00 E	1200	2
1c	3	Mekong	That Phanom	17 24 00 N	104 48 00 E	1100	2
1d	4	Mekong	Warinchamrap	15 12 00 N	104 53 00 E	885	3
1e	5	Mekong	Mun River	15 17 00 N	105 04 00 E	800	13
1f	6	Mekong	Phibum	15 14 00 N	105 14 00 E	845	4
1g	7	Mekong	Nomh Penh	11 33 00 N	104 55 00 E	250	4
2a	8	Chao Phraya	Market	14 35 00 N	100 27 00 E	100	8
2b	9	Chao Phraya	Nakhom Sawan	15 41 00 N	100 07 00 E	275	5
3a	10	Chanthaburi		12 48 00 N	102 09 00 E	60	6
4a	11	Mea Khlong	Kanchanaburi	14 01 00 N	99 32 00 E	100	5
5a	12	Thai Pen	Phatthalung	7 37 00 N	100 05 00 E	90	1
6a	13	Kelantan		6 00 00 N	102 10 00 E	50	12
7a	14	Terengganu	Kuala Berang	5 04 00 N	103 01 00 E	40	14
7b	15	Terengganu	Sekayu	4 58 00 N	103 01 00 E	80	3
8a	16	Pahang	Kampong Mengkarak	3 19 00 N	102 27 00 E	100	9
9a	17	Endau	Park	2 40 00 N	103 15 00 E	100	6
9b	18	Endau	Kahang	2 18 00 N	103 35 00 E	50	1
9c	19	Endau	Ambat	2 11 00 N	103 51 00 E	70	2
9d	20	Endau	Mersing	2 25 00 N	103 56 00 E	5	10
10a	21	Johor	Kota Tinggi	1 4 00 N	103 54 00 E	45	3
11a	22	Perak		5 12 00 N	101 02 00 E	180	13
11b	23	Perak	Chenderoh	4 58 00 N	100 57 00 E	160	6
11c	24	Perak	Gerik	5 25 00 N	101 08 00 E	200	12
12a	25	Bernam	Sabak	3 46 00 N	100 59 00 E	50	1
13a	26	Muar	Selabu	2 9 00 N	102 36 00 E	20	6
13b	27	Muar	Market	2 14 00 N	102 46 00 E	40	4
13c	28	Muar	Kundang Ulu	2 14 00 N	102 46 00 E	40	2
13d	29	Muar	Panchor	2 10 00 N	102 43 00 E	35	4
14a	30	Batang Hari	Jambi	1 36 00 S	103 37 00 E	100	27
14b	31	Batang Hari	Kerinci	2 05 00 S	101 23 00 E	800	7
15a	32	Musi	Palembang	2 55 00 S	103 37 00 E	70	15
16a	33	Ci Liwong	Sawangan	6 24 00 S	106 46 00 E	45	2
17a	34	Ci Sadane	Bogor	6 35 00 S	106 47 00 E	60	1
18a	35	Ci Tarum	Cirata	6 33 00 S	107 18 00 E	90	18
19a	36	Ci Manuk	Garut	7 13 00 S	107 54 00 E	150	14
20a	37	Kali Progo	Mayudan	7 45 00 S	110 11 00 E	30	2
21a	38	Brantas River	Surabaya	7 28 00 S	112 26 00 E	50	6
22a	39	Mahakam	Samarinda	0 24 00 S	116 58 00 E	50	10
23a	40	Barito	Banjarmassin	3 5 00 S	114 38 00 E	50	9
24a	41	Kapuas	Pontianak	0 17 00 S	109 19 00 E	20	12
24b	42	Kapuas	Sanggau	0 08 00 N	110 36 00 E	120	2
24c	43	Kapuas	Nanga Pinoh	0 20 00 S	111 44 00 E	320	3

TABLE I: Continued.

River code	Site code	River	Sample site	Latitude d° mm' ss''	Longitude d° mm' ss''	Distance upstream from river mouth (km)	Number of specimens
24d	44	Kapuas	Sintang	0 04 00 N	111 30 00 E	270	3
24e	45	Kapuas	Danau Sentarum	0 51 00 N	112 6 00 E	550	2
24f	46	Kapuas	Putussibau	0 50 00 N	112 56 00 E	600	2
24g	47	Kapuas	Sibau	0 55 00 N	112 57 00 E	750	2
24h	48	Kapuas	Mendalam	1 00 00 N	113 16 00 E	800	6
25a	49	Sadong	Serian	1 10 00 N	110 34 00 E	50	25
25b	50	Sadong	Balai Ringin	1 03 00 N	110 45 00 E	70	11
26a	51	Rajang	Kapit	2 01 00 N	112 56 00 E	210	7
26b	52	Rajang	Sibu	2 18 00 N	111 49 00 E	80	10
26c	53	Rajang	Katibas	1 43 00 N	112 33 00 E	170	12
27a	54	Baram	Mulu	3 30 00 N	115 30 00 E	380	3
27b	55	Baram	Bakung	4 17 00 N	114 14 00 E	60	6
27c	56	Baram	Marudi	4 11 00 N	114 19 00 E	80	15
28a	57	Tutong	Merimbun	4 36 00 N	114 40 00 E	40	3
29a	58	Belait	Ingei	4 10 00 N	114 43 00 E	100	3
30a	59	Padas	Tenom	5 08 00 N	115 57 00 E	70	9
31a	60	Segaliud	Sandakan	5 45 00 N	117 52 00 E	50	1
32a	61	Segama	Danum Valley	5 01 00 N	117 47 00 E	170	16
32b	62	Segama	Lahad Datu	5 02 00 N	118 19 00 E	90	6
33a	63	Kinabatangan	Batu Putih	5 25 00 N	117 55 00 E	100	6
34a	64	Umas-Umas	Tawau	4 22 00 N	117 44 00 E	15	2
35	65	Salween	Myanmar	—	—	—	2
36	66	Red	Hanoi	—	—	—	2
37	67	Pearl	Guangdong	—	—	—	8
Total							452

—: exact origin unknown.

amplified in a programmable thermal cycler using the following profile: one single preliminary denaturation at 95°C for 5 min., 40 cycles of amplification with denaturation at 94°C for 60 s, primer annealing at 45°C for 60 s, primer extension at 72°C for 90 s, and a final single extension step performed at 72°C for 5 min. Total 50 µL PCR products were purified on a 2% agarose gel. After electrophoresis at 90 V for 2 hours, the amplified fragments were excised from the gel and extracted and purified using the QIAquick Gel Extraction Kit Protocol (using a microcentrifuge) from QIAGEN to a final volume of 15 to 50 µL (in EB buffer; 10 mM Tris Cl, pH 7.5).

Double stranded purified PCR products were sequenced in both directions (complementary strand was sequenced in all cases) using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems Inc.). Sequencing reactions were carried out in 20 µL total volume containing 3.2 picomoles of the sequencing primer, 30 to 60 ng of purified PCR product, 6 µL of 2.5X dilution buffer (CSA seq. Buffer; Perkin Elmer), and 2 µL of ABI PRISM BigDye Terminator Ready Reaction Mix. Reactions were performed in thermal cycler GeneAmp PCR system 9600 using the following profile: 10 s at 96°C, 5 s at 50°C, and

4 min at 60°C for 25 cycles. Sequences were separated by electrophoresis at 1680 V (150 W, 50.0 mA) for 5 to 7 h on ABI PRISM 377 DNA Sequencer (Applied Biosystems Inc.).

Sequences were extracted from the gel image using ABI PRISM Sequencing Analysis 3.0 program. The sequences of the complementary strands were then imported in Sequence Navigator program for comparative alignment and correction of ambiguities (Applied Biosystems Inc.). Sequences were aligned using the BioEdit sequence alignment editor [17]. Each sequence was identified by a unique code designating the origin of the sample (river name), the cumulative number of the sequences found in that river, and the number of individuals characterized by that particular sequence. For example, the code MEK1502 designates the 15th sequence found in the Mekong River and occurring in two specimens. Following the comparison of morphological and genetic designations, mtDNA codes were preceded by the putative species designation (e.g., SPMEK1502 indicates that specimens bearing sequence MEK1502 were assigned to *Hemibagrus spilopterus*). Sequences found in multiple rivers were identified by river name and the total number of specimens. Sequences were deposited in GenBank [18] with

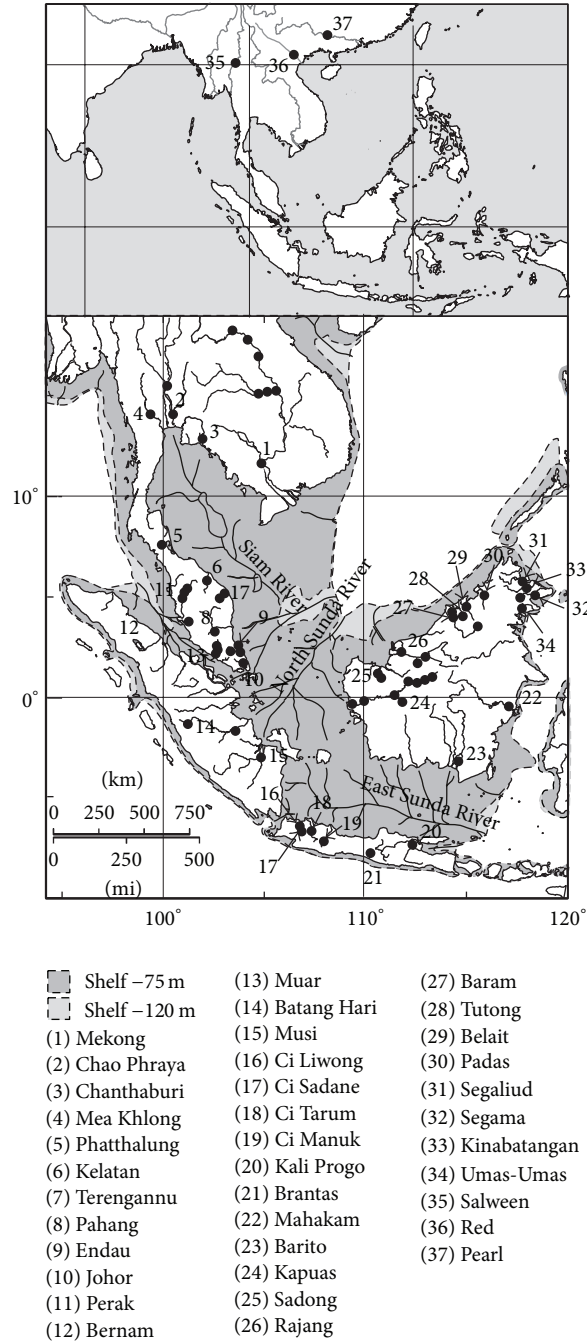


FIGURE 1: Location map of Southeast Asia illustrating the 37 rivers sampled in this study. Black dots along rivers represent the sampling sites as identified in Table 1 and Table S1. Dark grey shading illustrates the extent of the emerging Sunda Shelf at sea levels 75 m below present day levels; light grey denotes emerging shelf at sea levels 120 m below present day levels (as illustrated by Voris 2000). Rivers illustrated on the Sunda Shelf represent drowned Pleistocene river drainage systems, according to Voris (2000).

accession numbers corresponding to sequence codes and reported in Supplementary Material, Table S1.

Data Analysis. We first used the Neighbor-Joining (NJ) algorithm in MEGA 5.2 (Tamura et al., 2011) to develop a taxon identification phenogram of *Hemibagrus*. The Tamura-Nei distance matrix considering invariable site contribution

(+I) was selected according to the model selection algorithm implemented in MEGA 5.2 [19]. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). Recognizing the limitations of single-locus analyses in phylogeny reconstruction and the objections of cladists to distance-based methodology, we nevertheless sought to provide a tentative evaluation of phylogenetic

relationships among taxa. To do so, the Cyt *b* sequence of the Chinese longsnout catfish (*Leiocassis longirostris*) obtained from GenBank (accession number DQ321755.1) was used as an outgroup to root the phenogram. Chinese species currently assigned to *Leiocassis* may be considered members of the genus *Pseudobagrus* [20]. Furthermore, we employed a maximum likelihood (ML) phylogenetic reconstruction based on the comparison of nucleotides within sites to assess the congruence of the taxon identification phenogram and the ML phylogenetic reconstruction. In the latter case, the evolutionary model used was the Hasegawa-Kishino-Yano model (HKY) with gamma distribution (+G = 1.3290). Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining (NJ) method to a matrix of pairwise distances estimated using the maximum composite likelihood (MCL) approach. The rate-variation model allowed for some sites to be evolutionarily invariable (+I, 46.44% of sites). The ML tree was rooted with the Cyt *b* sequence of the Chinese longsnout catfish (*Leiocassis longirostris*) and constructed using MEGA5.2 [19].

The NJ phenogram classified mtDNA sequences based on overall similarity. Ng and Kottelat [7] also defined species groups based on overall similarity, but according to morphological criteria. Of the 8 species groups defined by Ng and Kottelat [7], two groups (*H. olyroides* (1 species) and *H. pluriradiatus* (5 species)) were not included in our samples. We thus evaluated the congruence of the two methods in defining the remaining six species groups (*H. baramensis*, *H. guttatus*, *H. menoda*, *H. nemurus*, *H. planiceps*, and *H. wyckii* species groups). When possible, mtDNA sequences were linked to ZRC accession numbers identifying specimens examined by Ng and Kottelat [7] in their morphological revision of the genus (Table S1).

To provide time estimates for the major phylogenetic events, we present a calibrated linearized tree obtained from our ML reconstruction (MEGA 5.2). Time of divergence between *Hemibagrus* lineages has been approximated by applying a rate of between 1.0 and 1.5% genetic divergence per million years as reported for bagrid fish by Yang and He [21]. However, the justification for this rate is not clear. We thus applied the more widely used rate of 2% sequence divergence per million years [22]. As there is undoubtedly a wide error associated with any such molecular clock calibrations, estimated divergence times are used to discriminate more recent from more ancient cladogenetic events while providing very approximate dates for such events.

Genetic divergence estimates using *p*-distances (*p*) were used in the present study to identify the different taxonomic levels observed among the *Hemibagrus* spp. analyzed. Since such information is readily available for thousands of animal groups [23], we used *p*-distance (*p*) as a benchmark for taxonomic levels. The survey of 20 731 vertebrate and invertebrate species reported by Kartavtsev [23] revealed increasing levels of *p*-distances (%) of the mtDNA Cyt *b* gene with taxonomic level: populations within species, $p = 1.38 \pm 0.30$ (SD); subspecies or sibling species, $p = 5.10 \pm 0.91$; morphologically distinct species within genera, $p = 10.31 \pm 0.93$; genera within a family, $p = 17.86 \pm 1.36$; and families within an order, $p = 26.36 \pm 3.88$. Although heterogeneity in the

gene evolution rate differs among different taxonomic groups, these estimates provide an order of magnitude evaluation of intraspecific versus interspecific divergence and permitted us to discriminate between population- and species-level divergences. However, as *p*-distances underestimate true genetic distance (e.g., site saturation and back-mutation), divergence times based on *p*-distances greater than about 10% are generally underestimated. We calculated *p*-distances in MEGA 5.2 and report them as percentages (standard error).

3. Results

The 452 samples yielded a total of 118 distinct sequences (not including the outgroup taxon). The Neighbor-Joining phenogram clustered the 119 sequences into a large number of hierarchically organized monophyletic clades. To facilitate presentation, we first present a compressed NJ tree and identify 11 major mtDNA clades (Figure 2). One unidentified fish from the Rajang River was characterized by a genotype that clustered with the outgroup taxon (XXRAJ0604, GenBank accession number KP322872, Figures 2 and 3). A second fish from the Rajang was characterized by an isolated genotype (XXRAJ0301, GenBank accession number KP322869) of unknown affiliation (Figures 2 and 3). To assess the congruence of the species groups defined by Ng and Kottelat ([7], hereafter referred to as N&K), and the mtDNA groupings revealed by the present analysis, we provide a direct comparison of the two methods following the order determined by the taxon identification phenogram (Figure 2, Table S1). Table S1 also provides additional information about the geographical distribution of sequences. Gene trees of individual clades, including all sequences, are presented at each step in the analysis along with their geographical distributions. We also present a compressed maximum likelihood phylogenetic tree (Figure 3). The phenogram and the ML tree provided congruent phylogenetic reconstructions. However, both trees failed to resolve the phylogenetic relationships of *H. wyckii*, *H. guttatus*, and *H. menoda* species groups, most likely due to undersampling of these lineages. The relationships among an unspecified species group, *H. spilopterus*, and *H. fortis* are also not well resolved. Their position in the NJ tree receives weak bootstrap support and they appear as a multifurcation along with a clade composed of *H. capitulum* and *H. hoevenii* in the ML phylogenetic tree (Figures 2 and 3). To facilitate comparisons with the taxonomic scheme of N&K, we follow the sequence of taxa as identified in Figure 2.

mtDNA Clade 1: The H. wyckii Species Group. The first clade is clearly consistent with the *H. wyckii* species group (Figure 4). This species group is composed of 4 valid species: *H. maydelli*, *H. microphthalmus*, *H. wyckii*, and *H. wyckioides* (N&K). The latter species is known from the Mekong and Chao Phraya drainage. Only specimens of *H. wyckioides* obtained from the Mekong and the Chao Phraya were available for DNA analysis. Three specimens from the Chao Phraya bearing sequence WYCHP0203 (Table S1) were identified as *H. wyckioides* (H. H. Ng, personal communication), although

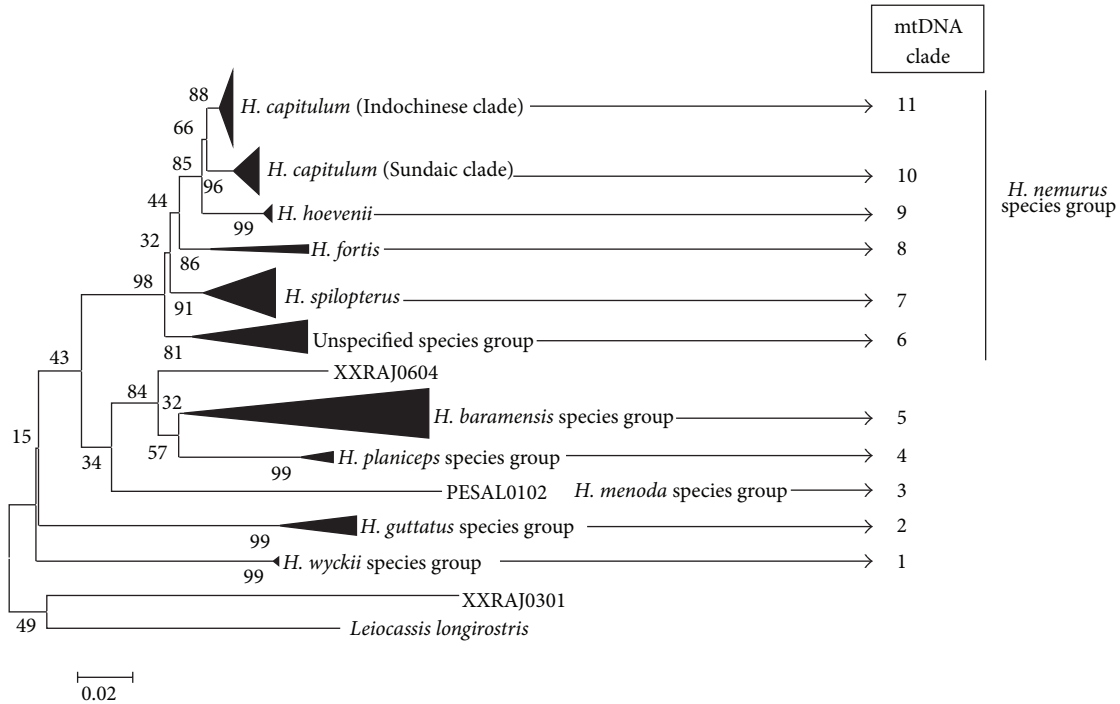


FIGURE 2: The taxon identification phenogram of *Hemibagrus* inferred using the Neighbor-Joining method in MEGA 5.2. The optimal tree with the sum of branch length = 1.86668598 is shown. Bootstrap values are indicated on each node of the tree (500 replicates). Bar indicates the number of substitutions changes estimated using Tamura-Nei distance considering invariable site contribution (+I). The analysis involved 119 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 597 positions in the final dataset.

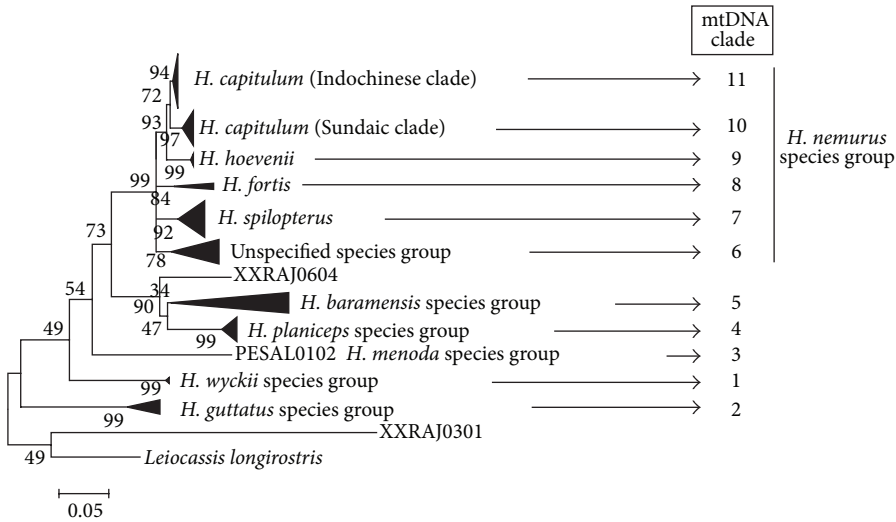


FIGURE 3: Evolutionary history of *Hemibagrus* inferred by using the maximum likelihood method based on the Hasegawa-Kishino-Yano model. The tree with the highest log likelihood (-5677.4825) is shown. Bootstrap values are indicated on each node of the tree (500 replicates). Bar indicates the number of substitution changes estimated using Hasegawa-Kishino-Yano model (HKY) with gamma distribution (+G = 1.3290). The analysis involved 119 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 597 positions in the final dataset.

their ZRC numbers do not appear in the analysis of N&K. As no specimens of the other 3 species comprising the *H. wyckii* species group were available for analysis, we cannot confirm if this species group is monophyletic.

mtDNA Clade 2: The H. guttatus Species Group. The second clade of the *Hemibagrus* gene tree (Figure 2) is composed of two reciprocally monophyletic subclades clearly reflecting the *H. guttatus* species group as defined by N&K (Figure 4).

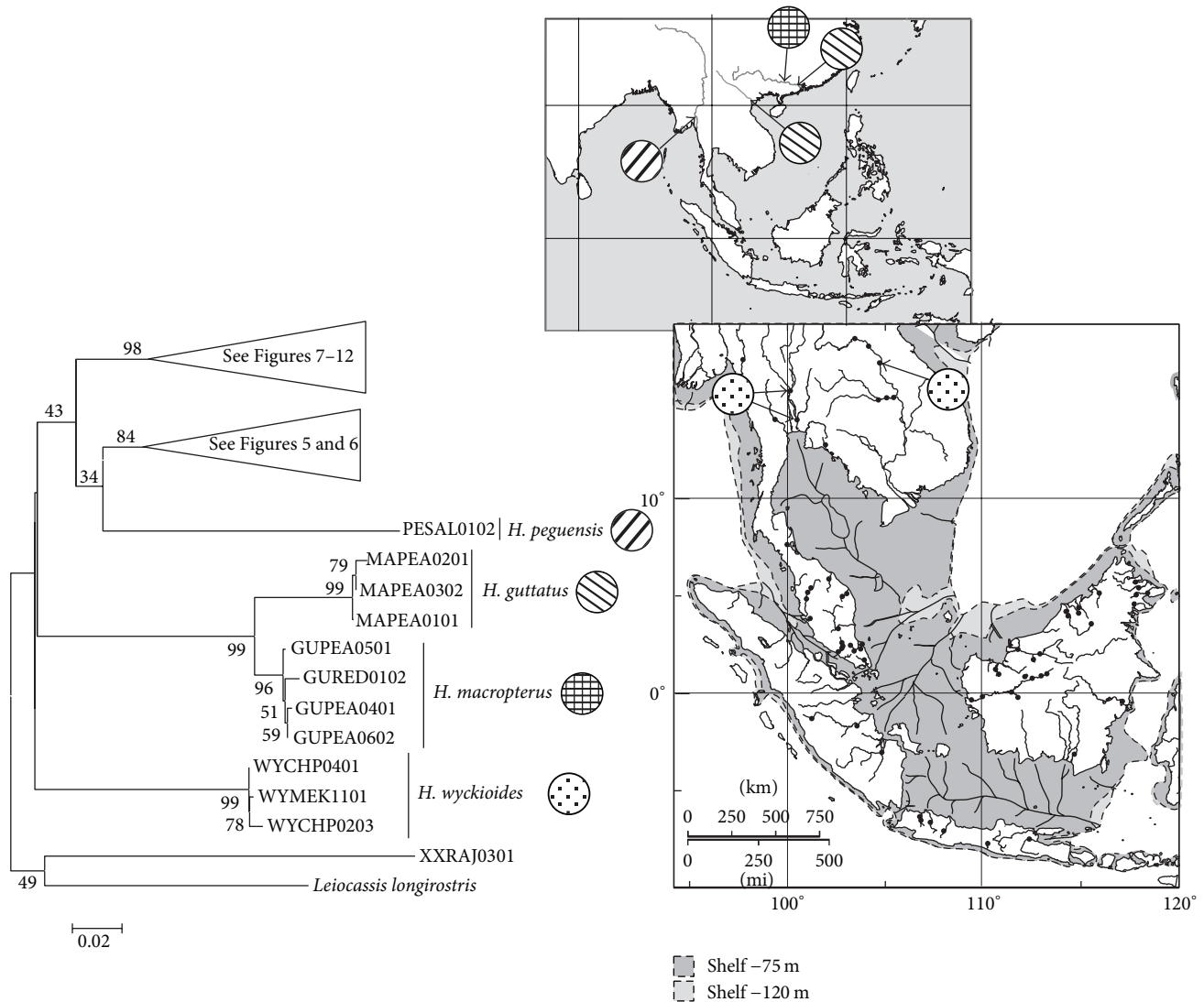


FIGURE 4: Phylogenetic relationships of the individual sequences comprising four species: *H. guttatus*, *H. macropterus* (together forming the *H. guttatus* species group), *H. wyckioides*, and *H. peguensis*. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

The group comprises three valid species: *H. guttatus*, *H. macropterus*, and *H. vietnamicus* (N&K). Specimens available for DNA analysis were obtained from the Pearl River and Red River (11 fish, 7 sequences; Table 1 and Table S1). The two reciprocally monophyletic subclades comprising the *H. guttatus* species group represent *H. guttatus* and *H. macropterus* (Figure 4). Although no ZRC specimens were examined by N&K, species identification was confirmed (H. H. Ng, personal communication). No *H. vietnamicus* specimens were available for DNA analysis. The two species composing the group are closely related ($p = 5.1\%$ (SE = 0.8%)).

mtDNA "Clade" 3: The H. menoda Species Group. This isolated genotype (Figure 4) most likely identifies a clade encompassing the *H. menoda* species group, although the

single sequence (Table S1) is shared by only 2 specimens of *H. peguensis* from the Salween River, Myanmar, also included in the analysis of N&K (where its provenance was identified as the neighboring Irrawaddy River). The other species comprising this group (*H. menoda*, *H. caveatus*, and *H. punctatus*) were not sampled for DNA analysis. The monophyletic nature of this species group thus remains to be demonstrated.

3.1. The H. planiceps + H. baramensis Group. The remainder of the gene tree is split into two large sister groups with marginal bootstrap support (73% in ML and 43% in NJ): the *H. planiceps* + *H. baramensis* group and the *H. nemurus* species group (Figures 2 and 3). The *H. planiceps* + *H. baramensis* group is strongly supported in both NJ and ML analyses (84% and 90% bootstrap support, respectively;

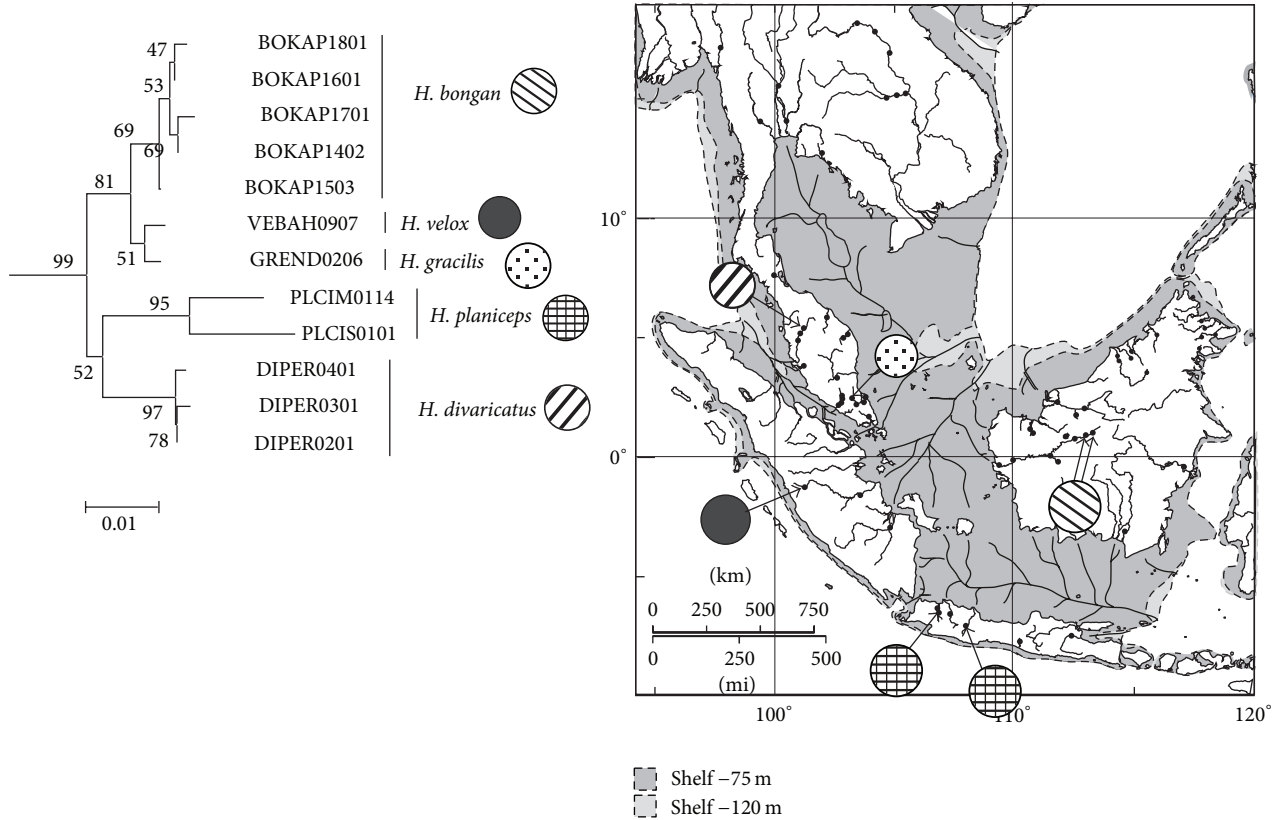


FIGURE 5: Phylogenetic relationships of the individual sequences and species comprising the *H. planiceps* species group: *H. bongan*, *H. velox*, *H. gracilis*, *H. planiceps*, and *H. divaricatus*. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

Figures 2 and 3). The two groups are genetically similar, with a genetic divergence estimate of $p = 12.1\%$ ($SE = 0.9$). Although the *H. planiceps* species group is strongly supported in both trees (99%), bootstrap support of the *H. baramensis* species group is weak (Figure 2: 32%, Figure 3: 34%). The geographical distribution of the two groups differs. Members of the *H. planiceps* species group are found throughout the Greater Sunda Islands and Peninsular Malaysia, but they are isolated in the upper reaches of large rivers (Figure 5). In contrast, members of the *H. baramensis* species group appear to be restricted to Sabah, North Borneo (Figure 6).

mtDNA Clade 4: The H. planiceps Species Group. The *H. planiceps* species group (Figure 5) is comprised of 5 closely related but genetically distinct species (mean within-group divergence, $p = 2.0\%$ ($SD = 0.33$); *H. planiceps* (Java), *H. gracilis* (eastern Peninsular Malaysia), *H. divaricatus* (western Peninsular Malaysia), *H. bongan* (Borneo), and *H. velox* (Sumatra)). *H. lacustrinus*, the sixth species of this group (N&K), was not sampled in the present study. Within the *H. planiceps* species group, *H. planiceps* formed a distinct species (Figure 5) comprising two sequences. One sequence (PLCIM0114, Table S1) characterized 14 specimens obtained from Ci Manuk, West Java, 9 of which were identified as *H. planiceps* by N&K (Table S1). A second sequence (PLCIS0101,

Table S1) was obtained from one specimen sampled at Ci Sadane, West Java, that we classified as *H. planiceps*. *H. bongan* is represented by 5 sequences obtained from 8 specimens sampled in the Kapuas River (Figure 5). Two of these specimens (ZRC 43013 and 43015) were identified as *H. bongan* by N&K (Table S1). *H. divaricatus* is composed of 3 sequences from three specimens obtained at Gerik on the Perak River (western Peninsular Malaysia). One of these specimens (ZRC 41151) is the holotype of this new species (N&K). *H. gracilis* is represented by only one sequence (GREND0206, Table S1) shared by 6 specimens sampled at Ulu Endau, Jasin River, eastern Peninsular Malaysia, four of which were identified by N&K (Table S1). ZRC 21484 comprises the holotype and ZRC 21482, 21486, and 21487 comprise the paratypes (N&K) of this species (Table S1). Finally, *H. velox* is represented by only one sequence (VEBAH0907) from 7 specimens sampled at Kerinci, Batang Hari River, Sumatra. Three of these specimens (ZRC 41503, Table S1) represent 3 paratypes of the species (N&K).

mtDNA Clade 5: The H. baramensis Species Group. The three species comprising the *H. baramensis* species group (Figures 2 and 3) form three monophyletic subclades (mean within-group divergence $p = 10.6\%$ (0.8): Figure 6). *H. semotus* is composed of two sequences (SEPAD0101, SEPAD0202) that

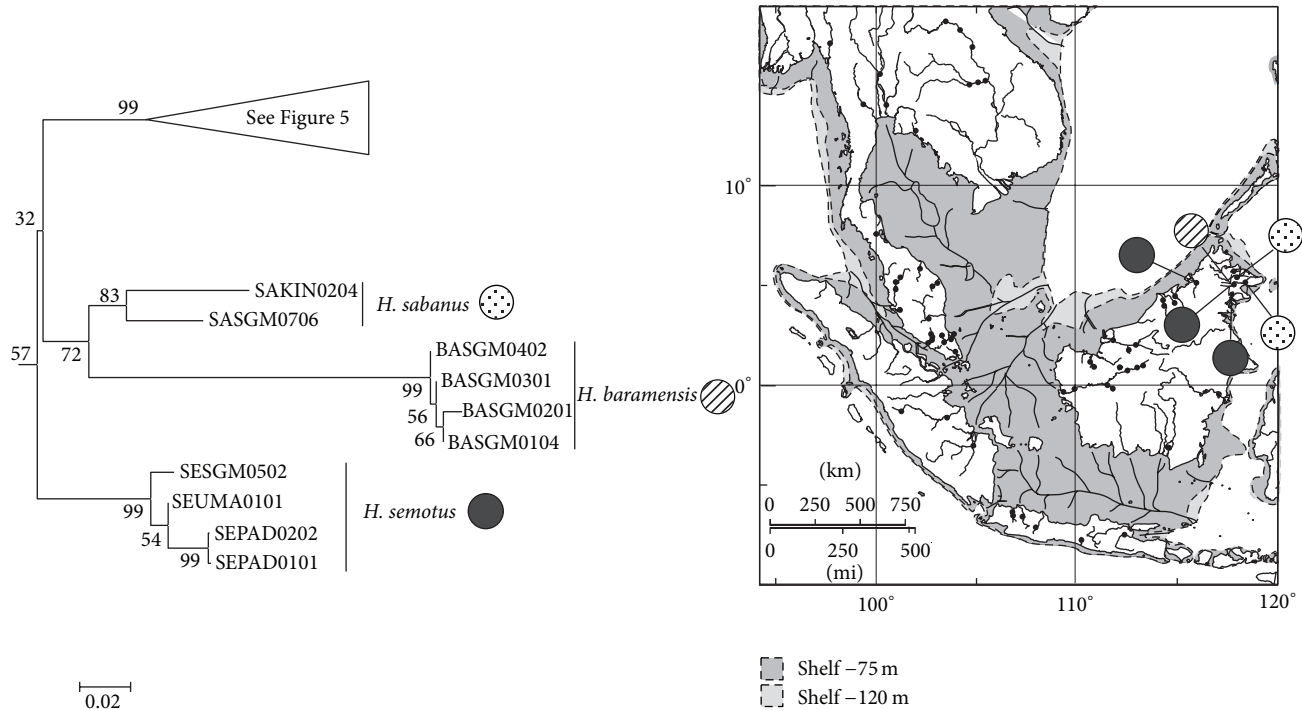


FIGURE 6: Phylogenetic relationships of the individual sequences and three species comprising the *H. baramensis* species group: *H. semotus*, *H. sabanus*, and *H. baramensis* (*furcatus*). The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

characterize fish that are the holotype and paratypes of *H. semotus* (ZRC 46122, Table S1) along with two somewhat more divergent sequences (Figure 6). One of these sequences (SGM0502) was found in the Segama River (Danum Valley) and identified as *H. baramensis* by N&K (ZRC 40525) (Table S1). The second sequence (UMA0101) was from a fish sampled at Tawau, Umas Umas River (ZRC 46123), also identified as *H. baramensis* by N&K. These sequences most likely reflect distinct lineages within *H. semotus* inhabiting different rivers as the mean intraspecific sequence divergence ($p = 1.6\%$) indicates population-level divergence.

H. sabanus and *H. furcatus* (synonymized with *H. baramensis* by N&K) cluster together to form a distinct subclade but clearly diverge at the species level (Figure 6). *H. sabanus* comprises 2 sequences that are shared, respectively, by 6 specimens (ZRC 42741) from the Segama River, Sabah, and 4 specimens (part of ZRC 42743) from the Kinabatangan River, Sabah (Table S1). All of these specimens were identified as *H. sabanus* by N&K (Table S1). Finally, *H. baramensis* forms a distant subclade (14.7% mean sequence divergence between *H. baramensis* and *H. sabanus*) composed of 4 sequences (Figure 6) shared by 8 specimens, all sampled in the Segama River, Danum Valley Conservation Area, Sabah. Six of these specimens (ZRC 40523 (2 specimens), ZRC 40527 (2 specimens), and ZRC 40526 (2 specimens) Table S1) comprise the paratypes of *H. baramensis* (*furcatus*) (N&K).

3.2. The *H. nemurus* Species Group. In both NJ and ML trees (Figures 2 and 3), the *H. nemurus* species group is strongly

supported (bootstrap support 98% in NJ and 99% in ML), indicating that it forms a distinct and widespread genetic entity within the genus *Hemibagrus*. However, relationships within the group are not so clear. In general, the *H. nemurus* species group is composed of 4 clades. The first clade may be considered as a species group nested within the *H. nemurus* species group (85% and 93% bootstrap support in NJ and ML, resp.) encompassing three clades: *H. hoevenii* (clade 9), the *H. capitulum* Sundaic clade (clade 10, including *H. nemurus*), and the *H. capitulum* Indochinese clade (clade 11) (Figures 2 and 3). An additional three clades (Figures 2 and 3) are formed by an unidentified species group (clade 6), *H. spilopterus* (clade 7), and *H. fortis* (clade 8). These three latter clades are distinct at the congeneric, species level (unspecified versus *H. spilopterus*, $p = 7.7\%$ (0.7), unspecified versus *H. fortis*, $p = 8.5\%$ (0.7), and *H. spilopterus* versus *H. fortis*, $p = 7.5\%$ (0.8)). The distribution of these three clades is somewhat disparate; *H. fortis* and the unspecified species group are principally found in Borneo, whereas *H. spilopterus* is widely distributed on the Indochinese mainland.

mtDNA Clade 6: Unspecified Species Group. This genetically diverse group, encompassing at least 3 subclades (*H. spp.* A, B, and C, Figure 7, within-group mean sequence divergence $p = 6.8\%$ (0.6)), is confounded with the *H. baramensis* species group by N&K. It is composed of sequences from fish sampled in Sarawak, Sabah, and eastern Peninsular Malaysia. Subclade *H. spp.* A (Figure 7) is particularly interesting as it contains 2 sequences (CHSARA20, CHRAJ0401) that characterize

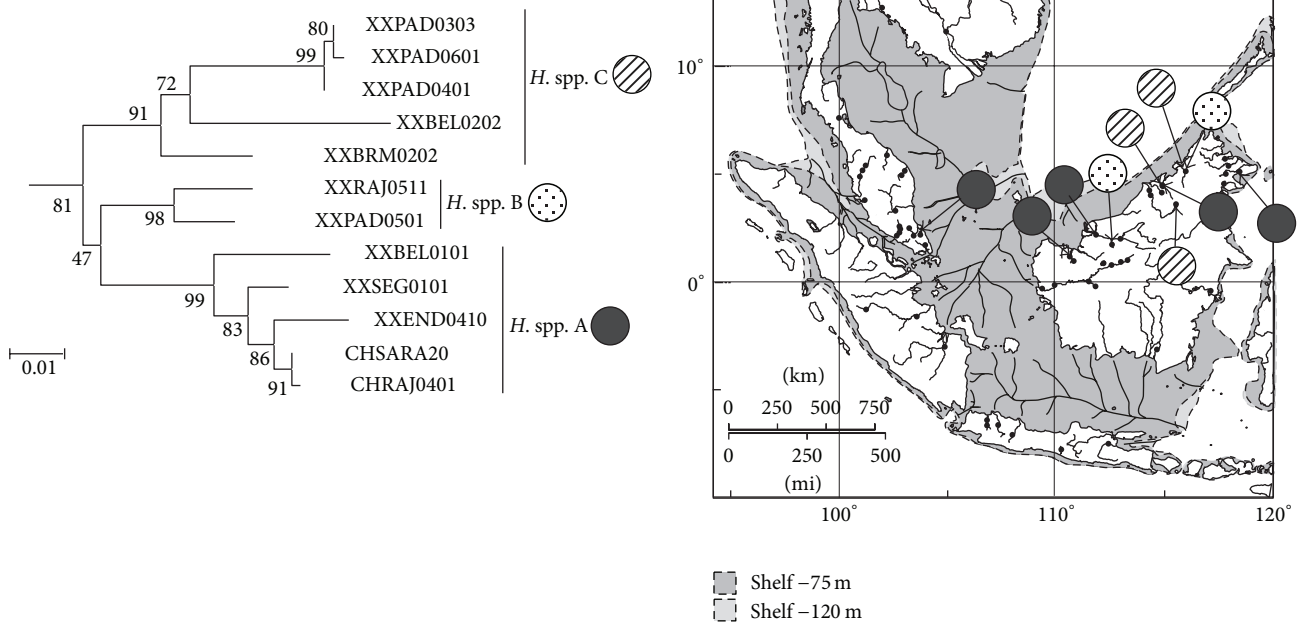


FIGURE 7: Phylogenetic relationships of the individual sequences comprising the unspecified species group, containing three subclades identified as *H. spp. A*, *B*, and *C*. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

H. chrysops, a new species identified by Ng and Dodson [11]. However, specimens of *H. chrysops* from the Rajang River were subsequently synonymized by N&K with *H. capitulum* of the *H. nemurus* species group (ZRC 23008–23014, Table S1). However, an additional 14 fish from the Sadong River at Serian, bearing genotype SARA20 (Table S1), and thus genetically identical to the Rajang *H. chrysops* samples, were identified by N&K as *H. spilopterus* (ZRC 29583–96, Table S1). These specimens, however, were erroneously identified in N&K as originating in the Chao Phraya River (Thailand). Subclade A also contains one sequence obtained from 10 fish sampled on the Endau River (Mersing) in eastern Peninsular Malaysia. These fish were not morphologically examined. Finally, subclade A also contains one sequence from one unidentified fish sampled in the Segaliud River, Sabah, and another unidentified fish sampled in the Belait River, Brunei. Subclade *H. spp. B* encompasses 2 sequences, one of which comes from a fish sampled in the Padas River (XXPAD0501) and part of the ZRC 46122 fish lot identified as *H. semotus* by N&K (Table S1). The second sequence (XXRAJ0511) was shared by 11 unidentified fish sampled at Katibas on the Rajang River, Sarawak (Table S1). Finally, subclade *H. spp. C* encompasses 5 sequences. Three sequences (XXPAD0303, XXPAD0401, and XXPAD0601) were obtained from the Padas River and identified by N&K as *H. semotus* (ZRC 46122, Table S1). The fourth sequence (XXBRM0202) was obtained from 2 fish sampled in the Baram River and was identified

as *H. baramensis* by N&K (ZRC 42742). The 5th sequence composing this group was obtained from 2 unidentified fish sampled in the Belait River, Brunei (Table S1).

mtDNA clade 6 is thus problematical within the context of the revision of the genus *Hemibagrus*. Two highly divergent mtDNA clades (clade 5 and clade 6, Figures 2 and 3) occur in sympatry principally in the Padas River of Sabah, North Borneo (Figures 6 and 7). Specimens of the unspecified species group are morphologically similar to the *H. baramensis* species group, where they are assigned in the analysis of N&K. However, the mean genetic distance between the unspecified species group and the *H. baramensis* species group is $p = 19.3\%$ (2.8), close to genus level divergence. Furthermore, subclade *H. spp. A* (Figure 7) contains not only *H. chrysops* but also an unidentified population from eastern Peninsular Malaysia (XXEND0410, Table S1) and two additional Sabah sequences. On the other hand, the unspecified species group clusters with all other members of the *H. nemurus* species group (mean genetic divergence of unspecified group versus *H. capitulum* (Sundaic clade), $p = 7.6\%$ (0.7) (Figure 2)), with strong bootstrap support. Such genetic clustering may justify the inclusion of *H. chrysops* and associated sequences in the *H. nemurus* species group, but its synonymization with *H. capitulum* (N&K) is difficult to defend on genetic grounds. We tentatively conclude that *H. chrysops* is a valid species based on genetic identity [11], forming part of an unknown

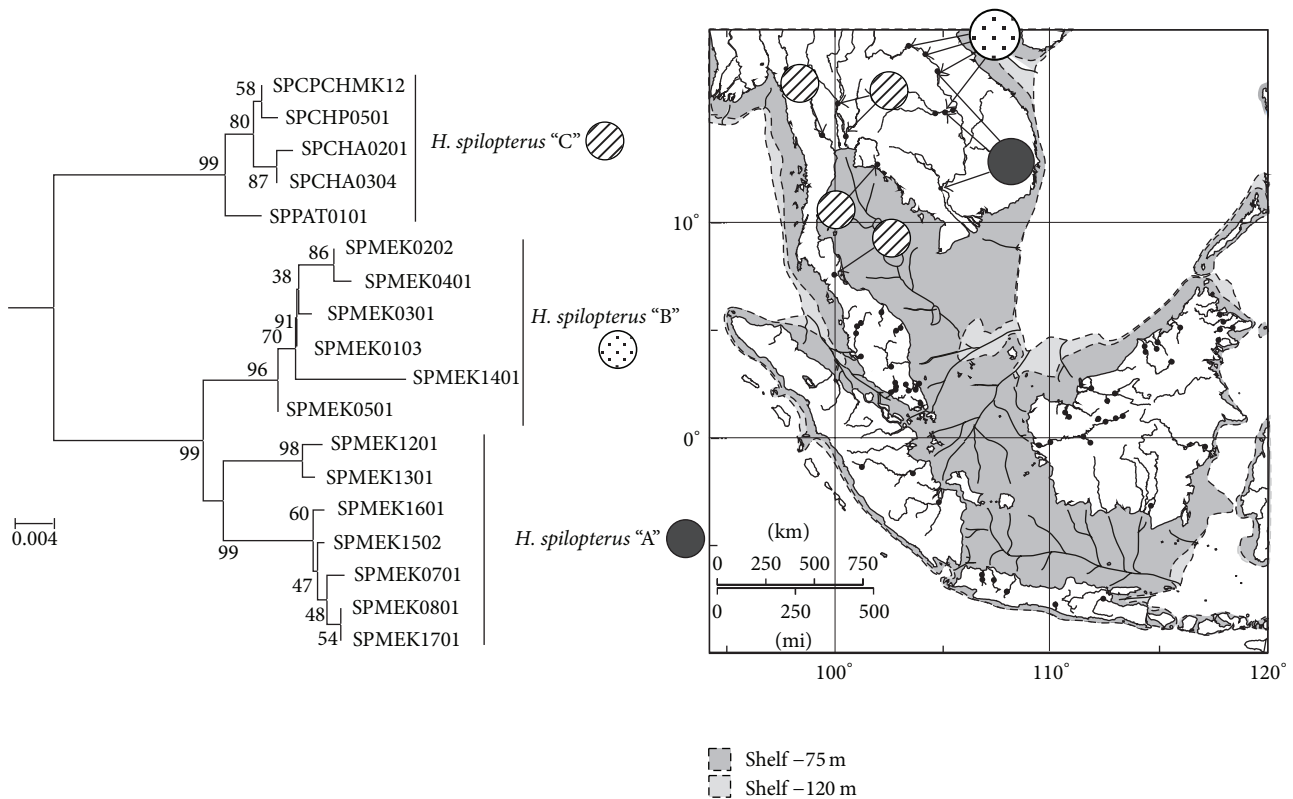


FIGURE 8: Phylogenetic relationships within *H. spilopterus*, comprised of 3 subclades identified as A, B, and C. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

species group found throughout North Borneo (and to a lesser degree in eastern Peninsular Malaysia) that may be considered as part of the *H. nemurus* species group. However, the morphological confusion of the unspecified species group with the sympatric, highly divergent, *H. baramensis* species group suggests the existence of a cryptic species assemblage (see Section 4).

mtDNA Clade 7. Clade 7 (Figures 2 and 3) is composed of at least 3 distinct subclades that appear to reflect 3 genetically divergent populations of *H. spilopterus*. Subclade A (Figure 8) encompasses 7 sequences obtained from 8 fish sampled in the Mekong River. Four fish bearing 3 of these sequences were identified as paratypes (ZRC 43324) of *H. spilopterus* by N&K (Table S1). One additional fish was identified as *H. spilopterus* by N&K (ZRC 46157, Table S1). Subclade B encompasses 6 sequences obtained from 9 fish, all sampled from the Mekong River (Table S1). The fish sampled at That Phanom (ZRC 46012) was identified as *H. spilopterus* in the analysis of H&K (Table S1). Subclade C encompasses 5 sequences obtained from 19 fish sampled in the Chao Phraya, Mae Khlong, and Chanthaburi Rivers, Thailand. One sequence (CPCMK12) was found in all of these drainages (Table S1). A closely allied sequence was found in one fish sampled at Phatthalung just south of the Isthmus of Kra on the Thai-Malay Peninsula. None of these fish were examined

by N&K and thus we are unable to definitively identify the species. However, given the widespread occurrence of *H. spilopterus* in the Chao Phraya and Mae Klong Rivers reported by N&K, we conclude that these fish are part of a widespread *H. spilopterus* population complex. The mean within-group (clade 7) genetic divergence ($p = 2.9\%$ (0.4)) suggests intraspecific, population-level divergence.

mtDNA Clade 8. Clade 8 encompasses 3 sequences obtained from 9 fish sampled in east Sabah, specifically the Kinabatangan, Segama, and Umas Umas Rivers (Figure 9). The Kinabatangan sequence is somewhat divergent within the clade and the specimen was not identified by N&K. One sequence (SGM0606) characterized 6 fish from the Segama River, 2 of which were identified as *H. fortis* (ZEC 40521, 40522, Table S1) by N&K (Table S1). Finally, one fish bearing a closely related genotype from the Umas Umas River was identified as *H. baramensis* by N&K (Table S1).

mtDNA Clade 9. Clade 9 is a sister group to the remaining 2 clades composing the *H. nemurus* species group. This well-defined clade (Figure 10) clusters 6 sequences from 42 fish that are clearly identified as *H. hoevenii* (N&K). Some of these haplotypes are geographically widespread; one sequence (JOMSKA11, Table S1) was found in the Johor River (East Peninsular Malaysia), the Musi River (Sumatra), and

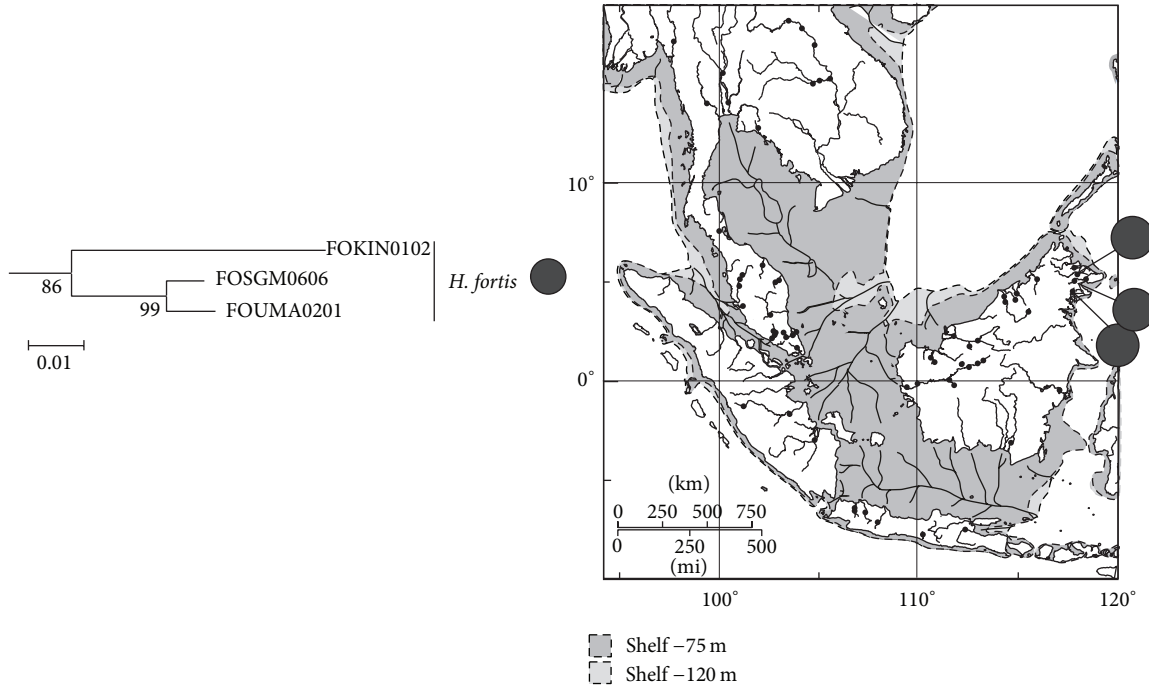


FIGURE 9: Phylogenetic relationships within *H. fortis*, comprised of 3 sequences. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

the Kapuas River (West Borneo) and a second sequence (HMUBHKABM23, Table S1) was found in Muar River (West Peninsular Malaysia) and the Batang Hari (Sumatra), the Kapuas (Sarawak), and the Baram Rivers (Sarawak). Many of these fish were identified as *H. hoevenii* by N&K (Table S1). The existence of two distinct but very closely related subclades (Figure 10, $p = 0.6\%$ (0.3)) that are not geographically segregated suggests some recent isolation event followed by widespread dispersal and secondary contact.

3.3. The *H. capitulum* Group. The final two monophyletic groups of the *Hemibagrus* gene tree are approximately equivalent to the remaining species composing the *H. nemurus* species group, as defined by N&K. Dodson et al. [10] and Ng and Dodson [11] have previously referred to clade 10 as the Sundaic clade, as most sequences were obtained from fish sampled on the Greater Sunda Islands, and clade 11 as the Indochinese clade, as many sequences were obtained from fish sampled on the Indochinese mainland (Figure 2, Table S1). Although most of these sequences are assigned to *H. capitulum* by N&K, there are some significant phylogeographic differences between the 2 clades. In particular, there appears to be an east-west division in the distribution of the two clades on the Thai-Malay Peninsula, with the Sundaic clade (clade 10) found on the west coast and the Indochinese clade found on the east coast, with sympatric occurrence in the Muar River. Nevertheless, mean sequence divergence between the two clades is $p = 2.9\%$ (0.5), a value that is indicative of population-level or subspecies

divergence. The two clades are thus best considered as comprising one species.

mtDNA Clade 10. The Sundaic clade includes a number of closely related sequences that comprise part of *H. capitulum* and *H. nemurus*. Very little genetic structure is apparent and bootstrap support for tree topology within this group is generally very weak (Figure 11). A widely spread genotype (NCLCTBA20) characterized 6 of 17 specimens from Ci Tarum at the Cirata reservoir (West Java) (Table S1) and 5 were identified as *H. nemurus* in the analysis of N&K (ZRC 42564) (Table S1). The remaining fish characterized by this sequence were obtained from Ci Liwong at Sawangan (West Java) and one fish was obtained from the Barito River at Banjarmasin, south Borneo. This latter specimen (ZRC 40553) was identified by N&K as *H. capitulum* (Table S1). Another common genotype (NBHKPBR09) was identified as either *H. capitulum*, *H. hoevenii*, or *H. nemurus* depending on the river of origin (Table S1). Many of the remaining sequences of the Sundaic clade appear to characterize fish that N&K identified as *H. capitulum*. A geographically widespread genotype (CPEBEMUKA36) characterized 9 fish sampled on the Perak River at Gerik that were identified by N&K as *H. capitulum* (ZRC 41150, Table S1). Another common genotype within this clade (CABHMS20) included 2 fish from the Batang Hari at Jambi and identified by N&K as *H. capitulum* (ZRC40534, Table S1). Another 2 sequences (CABAR0203, CABAR0305) characterized 8 fish sampled on the Barito River at Banjarmasin and were identified as

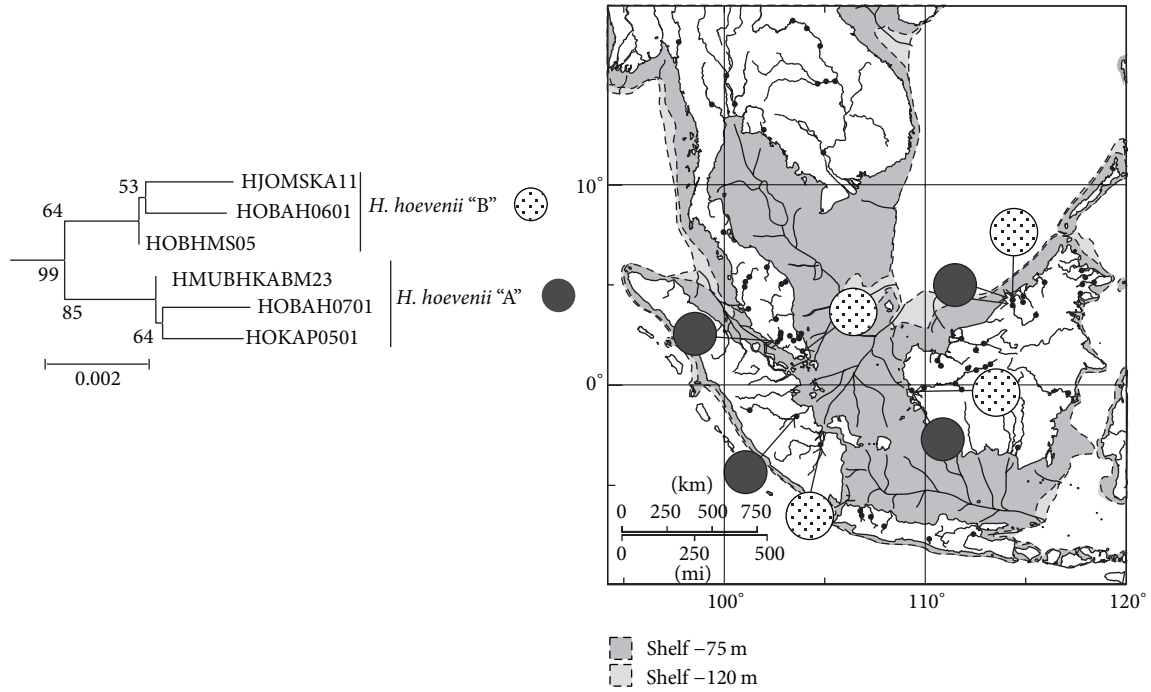


FIGURE 10: Phylogenetic relationships within *H. hoevenii*, comprised of 2 subclades and 6 sequences. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

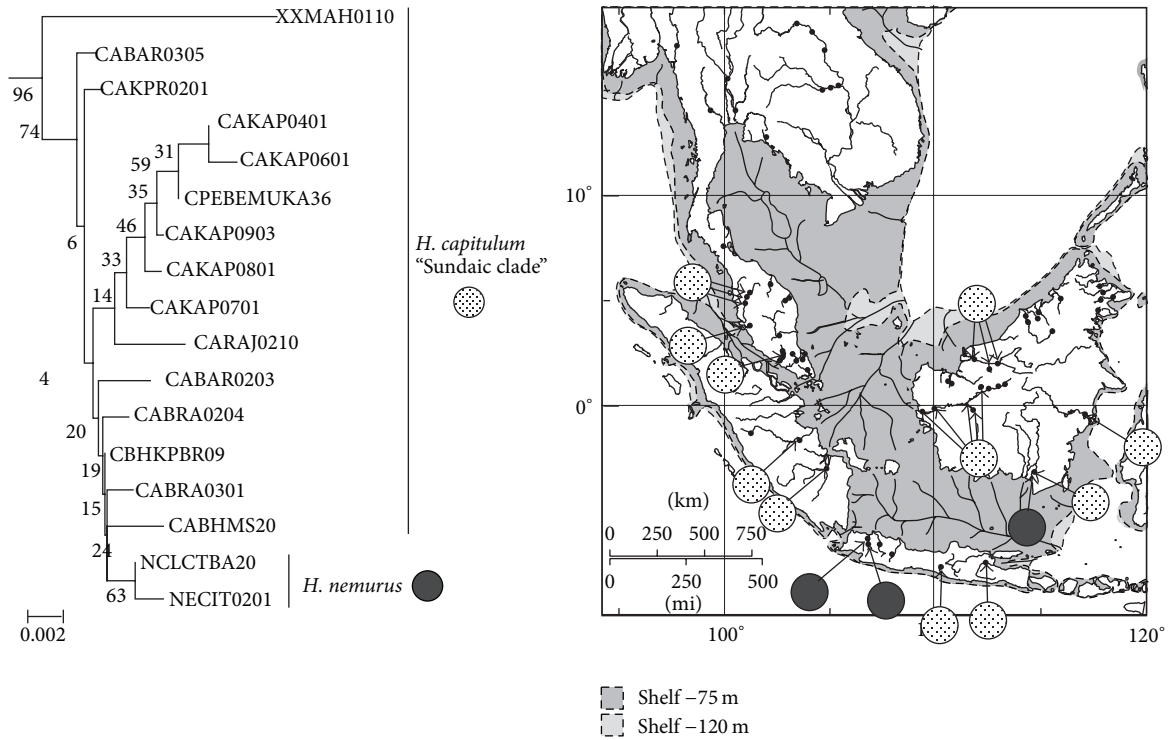


FIGURE 11: Phylogenetic relationships within the *H. capitulum* "Sundaic" clade, including 2 sequences associated with *H. nemurus*. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

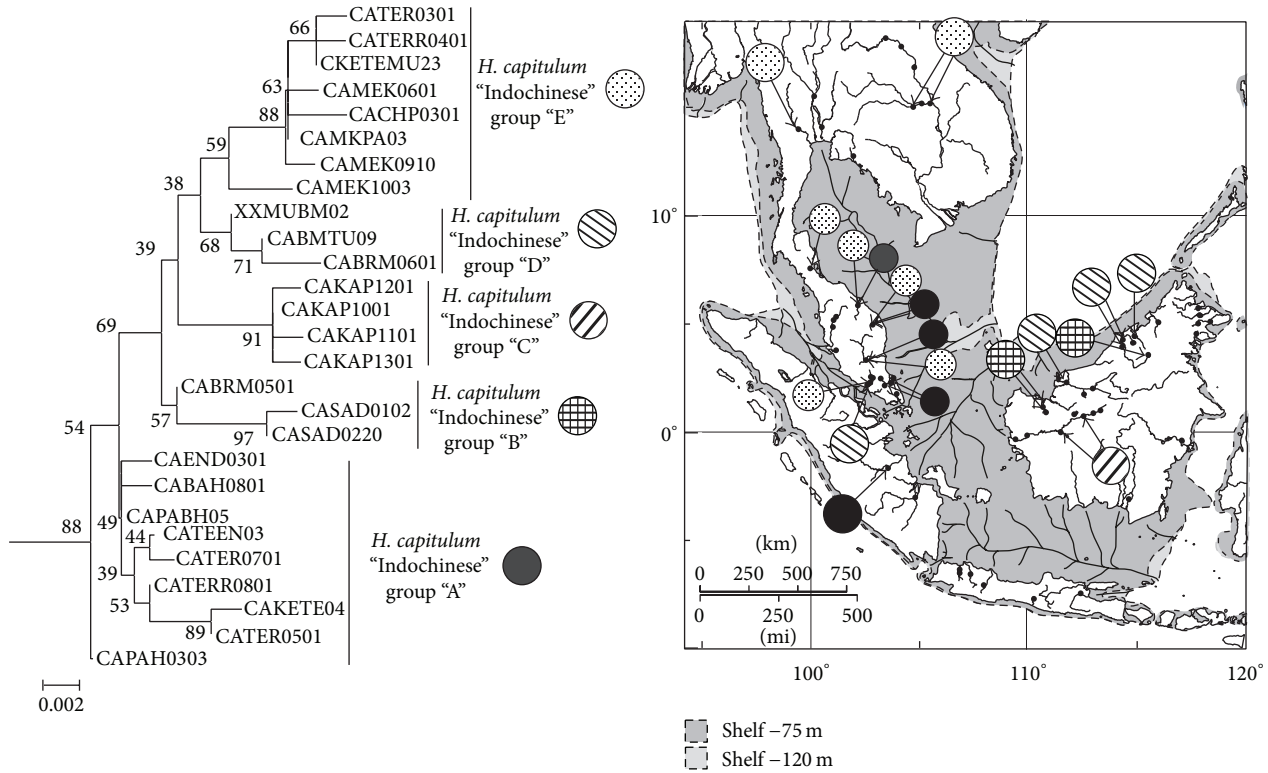


FIGURE 12: Phylogenetic relationships within the *H. capitulum* "Indochinese" clade, encompassing 5 subclades identified as A, B, C, D, and E. The subtree is extracted from the taxon identification phenogram inferred using the Neighbor-Joining method (see Figure 2). The percentage of replicate trees in which the associated sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches. The distribution of these sequences is illustrated in the right panel.

H. capitulum by N&K (ZRC 40553, Table S1). Although not included in the analysis of N&K, samples from the Sundaic clade were also found in the Kapuas River (West Borneo). Finally, 10 fish sampled on the Mahakam River at Samarinda, east Borneo, are all characterized by one relatively divergent genotype (CAMAH0110) but were considered as *H. capitulum* in the analysis of N&K (ZRC 40533, Table S1). The Mahakam population sample is the most genetically distinct of the Sundaic clade ($p = 1.7\%$ (0.4)), but nevertheless indicative of intraspecific, population-level divergence (Figure 12).

The topology of the Sundaic clade does not permit us to draw a clear distinction between *H. nemurus* and the Sundaic *H. capitulum*. The mean within-group divergence of the Sundaic clade is only $p = 0.8\%$ (0.2), well within population-level divergence values. We postulate that *H. nemurus* may be one of several regional variants of a widespread Sundaic *H. capitulum* clade that has radiated relatively recently. This may explain the existence of many shared sequences throughout the Sundaic islands assigned to different species in different regions.

mtDNA Clade 11. Finally, the Indochinese clade of *H. capitulum* represents a geographically widespread and weakly diversified clade (within-group mean sequence divergence $p = 1.1\%$ (0.2); Figure 12). Although few fish from this genetic clade were included in the morphological

analysis of N&K, most may be considered as *H. capitulum* given their geographic distribution and close genetic relationship to the Sundaic *H. capitulum*. Four sequences in this clade were identified as *H. capitulum* in the analysis of N&K: 3 specimens characterized by one sequence (CABMTU09) found in the Tutong River of western Borneo (ZRC 40482), 6 characterized by the same sequence from the Baram River (ZRC 40498), another 2 sequences from the Baram River (ZRC 40498), and 20 specimens characterized by an abundant genotype (CASAD0220) found at 2 sites in the Sadong River, also in western Borneo (ZRC 39526, 5 specimens; 39527, 3 specimens; and 38753, 4 specimens) (Table S1). Unfortunately, none of the specimens collected on the Indochinese mainland for genetic analysis were examined by N&K. Nevertheless, the distribution of specimens analyzed for mtDNA resembles the distribution of specimens examined by N&K. Although weakly structured, we may recognize 5 subclades. One genotype (CAPAH0303) is basal to the phenogram. Subclade A encompasses 8 sequences mainly found in eastern Peninsular Malaysia (Table S1, Figure 12). These rivers furnished many of the specimens N&K examined to define *H. capitulum*. However, some sequences in this group occur in sympatry with sequences identified within the Sundaic *H. capitulum* group. Two sequences from subclade A occurred in the Batang Hari, Sumatra (Table S1). Subclade E also encompasses sequences

found in Peninsular Malaysia as well as in the Mekong, Chao Phraya, and Mae Khlong Rivers (Figure 12). In addition, 3 subclades encompass sequences found principally in western Borneo; subclade B in the Baram and Sadong Rivers, Sarawak; subclade C in the Kapuas River, West Kalimantan; and subclade D in the Baram and Tutong Rivers of Sarawak and Brunei, respectively (Figure 12). This latter subclade also includes one genotype found in the Muar River of western peninsular Malaysia. Finally, subclade E encompasses 5 sequences from the Mekong, Chao Praya, and Mae Khlong Rivers, all within the distribution area of *H. spilopterus*.

Evolutionary History. The evolutionary history of the *Hemibagrus* species groups analyzed here dates from at least the Miocene, 15–20 Ma (Figure 13) when considering the widely used molecular clock of 2% sequence divergence per million years. Bootstrap support for the sister group relationships of *H. wyckioides*, the *H. guttatus* species group, and *H. peguensis* are all weak (clearly due to undersampling) and little can be said of their evolutionary history other than the fact that they occupy basal positions in the rooted trees. The remainder of both the NJ and ML trees is composed of two major sister groups: the *H. planiceps* + *H. baramensis* group and the *H. nemurus* species group (Figures 2, 3, and 13). The *H. planiceps* + *H. baramensis* group is strongly supported in both NJ and ML analyses (84% and 90% bootstrap support, resp., Figures 2 and 3). The *H. nemurus* species group is also strongly supported in both analyses (98% and 99% bootstrap support, Figures 2 and 3). These two major phylogenetic groups are estimated to have diverged at least 8 to 12 Ma ago, during the mid to late Miocene (Figure 13).

Much of the evolutionary history of *Hemibagrus* has occurred during the Quaternary, starting approximately 2.6 Ma ago. This is particularly true of the *H. planiceps* and *H. nemurus* species groups. The radiation of the 5 species included in the *H. planiceps* species group appears to coincide with the middle Pleistocene transition period of approximately 1 Ma ago. Their high-elevation distribution, however, appears to have largely isolated them from the effects of lowered sea levels during glacial maxima and the attendant opportunities for widespread dispersal (Figure 5). The evolutionary histories of the *H. baramensis* species group and the unspecified species group date back to the late Pliocene. In contrast, the low-elevation distribution of the *H. nemurus* species group in Sundaland appears to have favored their widespread distribution during the Quaternary ice age. In particular, *H. hoevenii* appears to be the most recently dispersed of the species under study, with diversification and dispersal occurring approximately 0.3 Ma ago (Figure 13).

4. Discussion

4.1. The Congruence of Morphological and Genetic Analyses. Combining genetic and morphological analyses to elucidate taxonomy allowed us to better understand the diversification and evolutionary relationships within this speciose group of fishes. Given the rapid and relatively constant rate of mtDNA evolution, the deeper the branches within the phylogenetic tree are, the greater our ability to distinguish valid taxa

should be. Conversely, the lack of complete congruence between morphological- and genetic-based identifications is expected to increase among the shallower branches of the phylogenetic tree, resulting in the appearance of polyphyletic species. Although 2 of the 8 species groups defined by Ng and Kottelat [7] were not sampled in this study, the validity of the remaining 6 species groups and their component species defined on the basis of morphological similarity was largely confirmed by the analysis of the mitochondrial Cyt *b* bar code. As expected, congruence is strongest at the base of the NJ and ML trees, where the *H. guttatus* species group formed a well-defined monophyletic group. The *H. menoda* and *H. wyckii* species group were each represented by only one species, but they were clearly distinct relative to sequences characterizing the neighboring species groups. The *H. planiceps*, *H. baramensis*, and *H. nemurus* species groups and their component species were also generally congruent in morphological and genetic analyses. Also, as expected, most cases of discrepancy between morphological-based species identification and genetic-based species identification occurred on relatively shallow branches of the phylogenetic tree. These involve principally *H. spilopterus* and the two mtDNA subclades composing *H. capitulum*. However, a far more unexpected lack of congruence involves the definition of a deeply divergent unspecified genetic group confounding members of the *H. baramensis* species group in the morphological analysis of Ng and Kottelat [7].

Interpreting Polyphyly. The apparent polyphyletic origin of a species may be due to incomplete lineage sorting among recently diverged species, the introgression of mtDNA due to species hybridization, or the erroneous identification of species due to the lumping of morphologically cryptic species. In the absence of nuclear DNA analyses, it is often difficult to distinguish between introgression and incomplete lineage sorting. For example, several sequences characterizing the Indochinese clade of *H. capitulum* (subclade E) are found in the Mekong and Chao Phraya Rivers, well within the distributional area of *H. spilopterus*. *H. spilopterus* and *H. capitulum* first diverged at least 3 Ma. On the one hand, this apparent case of polyphyly may have resulted from incomplete lineage sorting between recently diverged *H. spilopterus* and *H. capitulum*. We would, however, have expected similar cases of polyphyly among populations of the Indochinese clade of *H. capitulum*, but this was not detected. On the other hand, the sympatric occurrence of *H. spilopterus* subclade A and subclade B sequences with *H. capitulum* Indochinese subclade E sequences in the Mekong River may represent a migration event and local introgression of *H. capitulum* with *H. spilopterus*. A similar event may have occurred in the Chao Phraya and Mae Khlong Rivers where *H. spilopterus* sequences cooccur with *H. capitulum* sequences. Computing the divergence of *H. capitulum* sequences (Indochinese subclade E), cooccurring with *H. spilopterus*, and *H. capitulum* sequences occurring in eastern Peninsular Malaysia (subclade A) reveals a mean sequence divergence $p = 1.5\%$ (0.4). Thus, the *H. capitulum* sequences occurring in sympatry with *H. spilopterus* diverged at most 750 ka ago

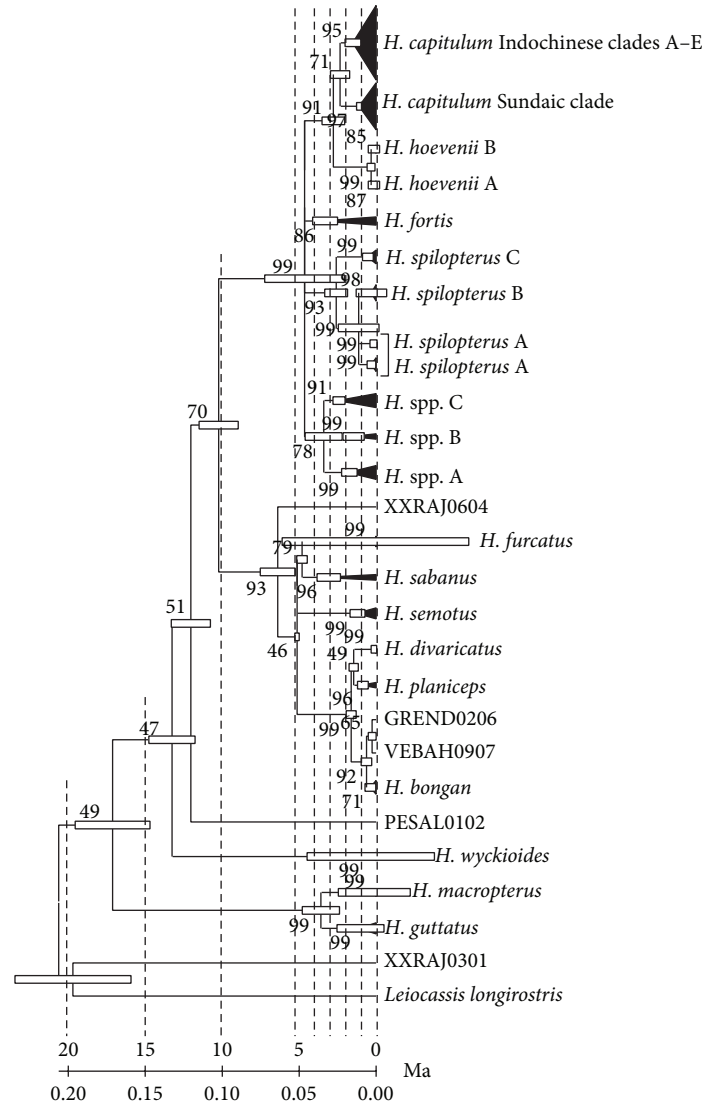


FIGURE 13: Linearized tree with branch lengths estimated with the molecular clock using a mutation rate of 1% per Ma (=divergence rate of 2% per Ma). The tree topology is identical to the ML reconstruction illustrated in Figure 3 (see the specificities of the tree reconstruction in the caption of Figure 3). Boxes represent 95% confidence intervals of divergence time estimates for each node. Vertical dashed lines redirect to dates indicated on the timescale (Ma). Numbers below timescale indicate the estimated number of substitution changes using Hasegawa-Kishino-Yano model (HKY) with gamma distribution (+G = 1.3290) after linearizing the tree. *H. furcatus* is synonymized with *H. baramensis*.

from eastern Peninsular Malaysia populations, suggesting a recent introgression event. Similarly, *H. capitulum* in the Kapuas River is characterized by sequences of both the Sundaic and the Indochinese clades. As both sequences are broadly sympatric, it is not possible to distinguish between incomplete lineage sorting and introgression of these closely related intraspecific lineages.

A Possible Case of Morphological Stasis. Clade 6 of the present analysis defines a well-supported monophyletic clade found in Sabah, Sarawak, and the Endau River of eastern peninsula Malaysia that has no equivalent taxonomic status in the revision of Ng and Kottelat [7]. In Borneo, they occur sympatrically with the *H. baramensis* species group and are confounded as either *H. baramensis* or *H. semotus*. These

two clades are genetically highly divergent, having diverged between 8 and 12 Ma ago. It thus seems highly unlikely that they represent retained polymorphisms due to incomplete lineage sorting. Sequences previously identified as characterizing *H. chrysops* [11] are part of clade 6 and extend to the south (Rajang and Sadong Rivers of Sarawak, West Borneo), with one closely related genotype found in eastern peninsular Malaysia (Endau River) where it may well be confounded with *H. capitulum*. *H. chrysops* was synonymized with *H. capitulum* by Ng and Kottelat [7], based on their analysis of specimens of *H. chrysops* from the Rajang River. Thus, despite a long and apparently independent phylogenetic history, the unspecified species group appears largely cryptic and confounded principally with members of the *H. baramensis* species group.

We suggest that this situation is an example of morphological stasis, involving the coexistence of a number of genetically divergent yet morphologically cryptic species that are lumped together [24]. A striking example of morphological and ecological stasis among fishes is provided by the coastal marine bonefishes (*Albula* spp.). Distinct evolutionary lineages (separated by sequence divergence between 5 and 30% [25]) of bonefish coexist in several sample locations, yet they show little morphological or ecological differentiation in sympatry [25]. Similarly, a proportion of the grey mullet species (Mugilidae) also appears to consist of cryptic species [26]. For example, *Mugil cephalus* in Taiwan is characterized by three divergent lineages that have been shown to characterize 3 reproductively isolated but morphologically cryptic species [26, 27].

Here, we have discovered 2 highly distinct and structured clades (each composed of 3 mtDNA subclades) that are morphologically confounded within different species. For example, *H.* spp. C of the unspecified species group is identified as *H. baramensis* in the Baram River and as *H. semotus* in the Padas River. We speculate that North Borneo harbors a morphologically cryptic species complex, possibly composed of several species pairs inhabiting different river systems. The reconstruction of this phylogenetic history and the taxonomic status of its components remains a major challenge to the revision of the *Hemibagrus* genus.

4.2. Phylogeographic Hypotheses. The phylogeography of the genus appears to have been principally influenced by two geological histories: one encompassing the formation of North Borneo, the second, the emergence of the Sunda Islands and their connectivity with the Indochinese mainland associated with recurrent sea-level changes. North Borneo appears to have acted as a species sink, promoting speciation and locally high biodiversity within the genus but with little evidence of subsequent dispersal out of North Borneo. In contrast, the geological events influencing the Indochinese mainland and the Sunda Islands have permitted both speciation and differential dispersal depending on the biology of the species. For example, the *H. planiceps* species group is found throughout the Greater Sunda Islands and the Thai-Malay Peninsula as is the *H. nemurus* species group. Whereas the *H. nemurus* species group generally occupies the low altitude, lower reaches of large rivers, the species composing the *H. planiceps* group are found in the upper reaches of large rivers. The *H. nemurus* species group is composed of a number of widely distributed closely related species whereas the *H. planiceps* group is composed of a large number of species that are highly localized geographically. It thus appears that sea-level changes have had little impact on the dispersal history of the *H. planiceps* species group but have deeply influenced that of the *H. nemurus* species group.

The Case of North Borneo. From 15 Ma to 5 Ma, greater areas of Borneo became emergent and the central mountains on the Sarawak-Kalimantan border extending into Sabah became wider and higher which may have represented new barriers to dispersal [28, 29]. Borneo gradually developed into the present large island by emergence of land in the

north and east. The *H. baramensis* species group is restricted to North Borneo as is the unspecified species group of the *H. nemurus* species group. *H. fortis* is also distributed in North Borneo. The isolation of the freshwater fish communities of North Borneo (particularly in the Labuk-Segama watershed) has long been recognized based on the number of endemic species and the concentration of relict species [30]. The late Pliocene-early Pleistocene diversification of *H. fortis*, the unspecified species group, and the *H. baramensis* species group may thus be associated with their dispersal to and isolation within the watersheds of North Borneo. Only one sequence of the unspecified species group is found outside of North Borneo, and *H. fortis* appears to be distributed throughout central and east Borneo [7].

The Phylogeography of the *H. nemurus* Species Group. Periodic marine transgressions on the Sundaland platform during the early and middle Miocene [2, 29] may have been sufficient to isolate populations of freshwater organisms and promote cladogenesis [31]. This may have been responsible for the initial mid to late Miocene divergence of the *H. nemurus* species group. The majority of sequences assigned to this species group belong to 3 species (*H. spilopterus*, *H. capitulum* (including *H. nemurus*), and *H. hoevenii*), all of which have been influenced by the more recent geological history of the Sunda Islands and their connectivity with the Indochinese mainland. *H. spilopterus* is known from river drainages in Indochina, including the Mekong, Dong Nai, Ban Pakong, Chao Phraya, and Mae Klong, where it is found in a wide variety of lotic and lentic habitats [7]. *H. spilopterus* diverged in the late Pliocene-early Pleistocene, possibly isolated to the north of the Kra Peninsula during the late Pliocene marine transgression that is purported to have created a major biogeographic transition zone between the Sundaic and Indochinese biotas ([32], but see [33] for a critique of this biogeographic scenario). The existence of 2 subclades (A and B) in the Mekong River, separated by 2% sequence divergence, is consistent with intraspecific divergence. The two subclades possibly represent 2 historical mtDNA lineages that are retained within the species, reflecting a Pleistocene isolation event followed by secondary contact within the contemporary Mekong River. Subclade C from the Chao Phraya and Mae Klong Rivers exhibits a level of genetic divergence consistent with sibling species-level divergence, showing a 5% sequence divergence relative to Mekong River *H. spilopterus*.

H. capitulum may be considered as part of the Sundaic biota, located principally to the south of the Kra marine transgression zone that may have isolated *H. spilopterus* to the north on the Indochinese mainland. The biogeographical event leading to the subsequent formation of two intraspecific groups within *H. capitulum* is not at all clear. The only area where the two clades are clearly geographically separate is on the Thai-Malay Peninsula, with the “Sundaic” clade on the west coast and the “Indochinese” clade on the east coast, separated by the central Titiwangsa mountain range. The genetic distinctiveness of east and west coast populations of *H. capitulum* has also been documented for nuclear DNA [34]. A similar pattern of separation has been

reported for 5 additional freshwater fishes within this region [35]. However, as the divergence of the two *H. capitulum* clades postdates the formation of the mountain range by approximately 200 Ma, the geographic separation of the two clades most likely represents differential dispersal rather than ancient vicariance. The presence of the Sundaic clade on the west coast of the Thai-Malay Peninsula as well as on the Sundaic Islands reflects the historical connections between east Sumatra and west Peninsular Malaysia across the Strait of Malacca. This area encompasses the Northern Central Sumatra-Western Malaysia freshwater ecoregion [36]. The Muar River forms its southern border, coincident with the southernmost distribution of the Sundaic clade on the Thai-Malay Peninsula. It is important to note, however, that most of Sumatra and Java emerged to achieve their present size and elevation only 5 Ma ago, and much of East Java continued to be the site of marine deposition until late Pliocene-early Pleistocene [29]. We may speculate that the Sundaic clade of *H. capitulum* resulted from the relatively recent colonization of these islands and Borneo. The radiation within the clade only dates back to approximately 1 Ma ago. *H. nemurus* from West Java may thus best be considered an isolated, regional population of very recent origin.

The diversification of the Indochinese clade dates from at least 1.2 Ma ago, coincident with the mid-Pleistocene transition at which time low-frequency, high-amplitude, and quasiperiodic (100 ka) glacial variability emerged. It began 1.25 Ma ago and was complete by 0.7 Ma ago [37]. The transition marked the beginning of major changes in sea levels during which the Sunda Shelf periodically emerged and was traversed by enormous fluvial systems [3]. The contribution of these events to the diversification of this clade cannot be satisfactorily reconstructed. One possible explanation is that high sea levels produced barriers to gene flow and the populations of West Borneo (subclades B, C, and D) were isolated from those of east Peninsular Malaysia. This could be due to one or more dispersal events from Borneo via the Great Sunda River (Figure 1) to east Peninsular Malaysia or lineage sorting whereby sequences of subclades A and E have been lost from West Borneo. As in the case of the sympatric occurrence of *H. spilopterus* and *H. capitulum* subclade E (see above) in the Mekong River, we are unable to clearly define the relative roles of incomplete lineage sorting and introgression following migration. Given the recent origins of the Indochinese *H. capitulum* sequences, incomplete lineage sorting most probably blurs patterns of recent isolation.

The last species of the *H. nemurus* species group is *H. hoevenii*. The species is widespread throughout Peninsular Malaysia and the Sundaic Islands. It is found primarily in the lower reaches of rivers and is apparently capable of withstanding low salt concentration [7]. *H. hoevenii* thus appears to be the most liable of all members of the *H. nemurus* species group to be impacted by Pleistocene sea-level changes. This is clearly reflected in the widespread distribution of the sequences of this species. Two lineages are evident, having diverged no more than 0.3 Ma ago, suggesting a late Pleistocene isolation event followed by extensive dispersal and admixture in recent times. This represents the clearest example of the dispersal opportunities afforded by

Pleistocene sea level changes on the Sunda Shelf. In addition, the relatively low genetic diversity detected (only 6 sequences in 42 fish) suggests a population bottleneck in the recent past, possibly related to population declines in the lower reaches of rivers flooded by increasing sea levels.

4.3. Future Directions. We have provided a global picture of the evolutionary taxonomy of the genus in Southeast Asia. Clearly, a more rigorous phylogenetic reconstruction of the genus *Hemibagrus*, requiring the construction of multiple gene trees based on the analysis of both mitochondrial and nuclear DNA markers from a wider taxonomic array of hemibagrids, will involve an even greater effort of sampling throughout the region. In particular, we require sampling at smaller spatial and temporal scales. For example, smaller scale sampling in the Padas River and the Segama River of North Borneo using both mitochondrial and nuclear genetic markers would contribute to clarifying the nature of the cryptic species complex believed to inhabit these rivers. The demonstration of reproductive isolation and/or microallopatric distribution among morphologically cryptic species may result in the description of greater species richness within the genus than presently described. The species status of *H. chrysops* merits reconsideration, given its highly divergent evolutionary history. Nuclear genetic markers in conjunction with mtDNA markers could reveal cases of historical introgression, particularly in the Kapuas River and the Mekong River. The hemibagrid catfishes continue to provide multiple opportunities for the study of the evolutionary taxonomy and biogeography of a major faunal component of the fluvial ecosystems of Southeast Asia.

Disclosure

Ethanol-preserved tissue samples of many of the specimens identified in Table S1 are freely available upon request from JJD.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Goh Yan Yih for assistance in collecting specimens and molecular analyses, Mrs. Françoise Colombani for molecular analyses, and Dr. Heok Hee Ng for species identification. This work was funded by the National University of Singapore grant to P. K. L. Ng and a Canadian Natural Sciences and Engineering Research Council grant to J. J. Dodson.

References

- [1] D. J. Lohman, M. de Bruyn, T. Page et al., "Biogeography of the Indo-Australian archipelago," *Annual Review of Ecology, Evolution, and Systematics*, vol. 42, pp. 205–226, 2011.

- [2] D. S. Woodruff, "Neogene marine transgressions, palaeogeography and biogeographic transitions on the Thai-Malay Peninsula," *Journal of Biogeography*, vol. 30, no. 4, pp. 551–567, 2003.
- [3] H. K. Voris, "Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations," *Journal of Biogeography*, vol. 27, no. 5, pp. 1153–1167, 2000.
- [4] S. Y. Yap, "On the distributional patterns of Southeast-East Asian freshwater fish and their history," *Journal of Biogeography*, vol. 29, no. 9, pp. 1187–1199, 2002.
- [5] S. K. J. McConnell, "Mapping aquatic faunal exchanges across the Sunda shelf, South-East Asia, using distributional and genetic data sets from the cyprinid fish *Barbodes gonionotus* (Bleeker, 1850)," *Journal of Natural History*, vol. 38, no. 5, pp. 651–670, 2004.
- [6] T. Mo, *Anatomy, Relationships and Systematics of the Bagridae (Teleostei: Siluroidei) with a Hypothesis of Siluroid Phylogeny*, vol. 17 of *Theses Zoologicae*, Koeltz Scientific Books, 1991.
- [7] H. H. Ng and M. Kottelat, "Revision of the Asian catfish genus *Hemibagrus* bleeker, 1862 (Teleostei: Siluriformes: Bagridae)," *Raffles Bulletin of Zoology*, vol. 61, no. 1, pp. 205–291, 2013.
- [8] P. K. L. Ng and H. H. Ng, "*Hemibagrus gracilis*, a new species of large riverine catfish (Teleostei: Bagridae) from Peninsular Malaysia," *Raffles Bulletin of Zoology*, vol. 43, pp. 133–142, 1995.
- [9] H. H. Tan and M. Kottelat, "Redescription of *Betta picta* (Teleostei: Osphronemidae) and description of *Betta falx* sp. n. from central Sumatra," *Revue Suisse de Zoologie*, vol. 105, no. 3, pp. 557–568, 1998.
- [10] J. J. Dodson, F. Colombani, and P. K. Ng, "Phylogeographic structure in mitochondrial DNA of a South-east Asian freshwater fish, *Hemibagrus nemurus* (Siluroidei: Bagridae) and Pleistocene sea-level changes on the Sunda shelf," *Molecular Ecology*, vol. 4, no. 3, pp. 331–346, 1995.
- [11] H. H. Ng and J. J. Dodson, "Morphological and genetic descriptions of a new species of catfish, *Hemibagrus chrysops*, from Sarawak, East Malaysia, with an assessment of phylogenetic relationships (Teleostei: Bagridae)," *Raffles Bulletin of Zoology*, vol. 47, no. 1, pp. 45–57, 1999.
- [12] P. D. N. Hebert, A. Cywinska, S. L. Ball, and J. R. DeWaard, "Biological identifications through DNA barcodes," *Proceedings of the Royal Society B: Biological Sciences*, vol. 270, no. 1512, pp. 313–321, 2003.
- [13] J. C. Carolan, T. E. Murray, Ú. Fitzpatrick et al., "Colour patterns do not diagnose species: quantitative evaluation of a DNA barcoded cryptic bumblebee complex," *PLoS ONE*, vol. 7, no. 1, Article ID e29251, 2012.
- [14] M. Kekkonen and P. D. N. Hebert, "DNA barcode-based delineation of putative species: efficient start for taxonomic workflows," *Molecular Ecology Resources*, vol. 14, no. 4, pp. 706–715, 2014.
- [15] D. M. Hillis, C. Moritz, and B. K. Marble, *Molecular Systematics*, Sinauer, Sunderland, Mass, USA, 2nd edition, 1996.
- [16] T. D. Kocher, W. K. Thomas, A. Meyer et al., "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*, vol. 86, no. 16, pp. 6196–6200, 1989.
- [17] T. A. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symposium Series*, vol. 41, pp. 95–98, 1999.
- [18] D. A. Benson, I. Karsch-Mizrachi, K. Clark, D. J. Lipman, J. Ostell, and E. W. Sayers, "GenBank," *Nucleic Acids Research*, vol. 40, no. D1, pp. D48–D53, 2012.
- [19] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar, "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods," *Molecular Biology and Evolution*, vol. 28, no. 10, pp. 2731–2739, 2011.
- [20] X. Ku, Z. Peng, R. Diogo, and S. He, "MtDNA phylogeny provides evidence of generic polyphyleticism for East Asian bagrid catfishes," *Hydrobiologia*, vol. 579, no. 1, pp. 147–159, 2007.
- [21] L. Yang and S. He, "Phylogeography of the freshwater catfish *Hemibagrus guttatus* (Siluriformes, Bagridae): implications for South China biogeography and influence of sea-level changes," *Molecular Phylogenetics and Evolution*, vol. 49, no. 1, pp. 393–398, 2008.
- [22] B. W. Bowen, A. L. Bass, L. A. Rocha, W. S. Grant, and D. R. Robertson, "Phylogeography of the trumpetfishes (*Aulostomus*): ring species complex on a global scale," *Evolution*, vol. 55, no. 5, pp. 1029–1039, 2001.
- [23] Y. P. Kartavtsev, "Divergence at Cyt-b and Co-1 mtDNA genes on different taxonomic levels and genetics of speciation in animals," *Mitochondrial DNA*, vol. 22, no. 3, pp. 55–65, 2011.
- [24] S. Wada, Y. Kameda, and S. Chiba, "Long-term stasis and short-term divergence in the phenotypes of microsnails on oceanic islands," *Molecular Ecology*, vol. 22, no. 18, pp. 4801–4810, 2013.
- [25] J. Colborn, R. E. Crabtree, J. B. Shaklee, E. Pfeiler, and B. W. Bowen, "The evolutionary enigma of bonefishes (*Albula* spp.): cryptic species and ancient separations in a globally distributed shorefish," *Evolution*, vol. 55, no. 4, pp. 807–820, 2001.
- [26] J.-D. Durand, K.-N. Shen, W.-J. Chen et al., "Systematics of the grey mullets (Teleostei: Mugiliformes: Mugilidae): molecular phylogenetic evidence challenges two centuries of morphology-based taxonomy," *Molecular Phylogenetics and Evolution*, vol. 64, no. 1, pp. 73–92, 2012.
- [27] K.-N. Shen, B. W. Jamandre, C.-C. Hsu, W.-N. Tzeng, and J.-D. Durand, "Plio-Pleistocene sea level and temperature fluctuations in the northwestern Pacific promoted speciation in the globally-distributed flathead mullet *Mugil cephalus*," *BMC Evolutionary Biology*, vol. 11, no. 1, article 83, 2011.
- [28] R. Hall, "The plate tectonics of Cenozoic SE Asia and the distribution of land and sea," in *Biogeography and Geological Evolution of SE Asia*, R. Hall and J. D. Holloway, Eds., pp. 99–131, Backhuys Publishers, Leiden, The Netherlands, 1998.
- [29] R. Hall, "Southeast Asia's changing palaeogeography," *Blumea*, vol. 54, no. 1–3, pp. 148–161, 2009.
- [30] R. F. Inger and P. K. Chen, "The freshwater fishes of North Borneo," *Fieldiana: Zoology*, vol. 45, pp. 1–268, 1962.
- [31] J. Bohlen, V. Šlechtová, H. H. Tan, and R. Britz, "Phylogeny of the Southeast Asian freshwater fish genus *Pangio* (Cypriniformes; Cobitidae)," *Molecular Phylogenetics and Evolution*, vol. 61, no. 3, pp. 854–865, 2011.
- [32] M. de Bruyn, E. Nugroho, M. M. Hossain, J. C. Wilson, and P. B. Mather, "Phylogeographic evidence for the existence of an ancient biogeographic barrier: the Isthmus of Kra Seaway," *Heredity*, vol. 94, no. 3, pp. 370–378, 2005.
- [33] J. Parnell, "The biogeography of the Isthmus of Kra region: a review," *Nordic Journal of Botany*, vol. 31, no. 1, pp. 001–015, 2013.
- [34] S. Usmani, S. G. Tan, S. S. Siraj, and K. Yusoff, "Population structure of the Southeast Asian river catfish *Mystus nemurus*," *Animal Genetics*, vol. 34, no. 6, pp. 462–464, 2003.
- [35] M. P. Tan, A. F. J. Jamsari, and M. N. Siti Azizah, "Phylogeographic pattern of the striped snakehead, *Channa striata* in Sundaland: ancient river connectivity, geographical and

anthropogenic signatures,” *PLoS ONE*, vol. 7, no. 12, Article ID e52089, 2012.

- [36] R. Abell, M. L. Thieme, C. Revenga et al., “Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation,” *BioScience*, vol. 58, no. 5, pp. 403–414, 2008.
- [37] P. U. Clark, D. Archer, D. Pollard et al., “The middle Pleistocene transition: characteristics, mechanisms, and implications for long-term changes in atmospheric pCO₂,” *Quaternary Science Reviews*, vol. 25, no. 23–24, pp. 3150–3184, 2006.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

