Journal of Zoological Systematics and Evolutionary Research

WILEY

RESEARCH ARTICLE OPEN ACCESS

Cryptic Diversity in the Slender Walking Catfish Inferred From Mitogenomes and DNA Barcodes Reveals a *Clarias nieuhofii* Species Complex Across Sundaland

Karnjanapond Sornchai[1,2] | Artem Lisachov[1,3] | Thitipong Panthum[1,2] | Suphada Kiriratnikom[4] | Narongrit Muangmai[1,5] | Prateep Duengkae[1,6] | Worapong Singchat[1,2,6,7] | Kornsorn Srikulnath[1,2,6,7]

[1]Animal Genomics and Bioresource Research Unit (AGB Research Unit), Faculty of Science, Kasetsart University, Bangkok 10900, Thailand | [2]Interdisciplinary Graduate Program in Bioscience, Faculty of Science, Kasetsart University, 50 Ngamwongwan, Chatuchak, Bangkok 10900, Thailand | [3]Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk 630090, Russia | [4]Modern Agriculture - Fishery, Faculty of Technology and Community Development, Thaksin University, Phatthalung 93210, Thailand | [5]Department of Fishery Biology, Faculty of Fisheries, Kasetsart University, Bangkok 10900, Thailand | [6]Special Research Unit for Wildlife Genomics (SRUWG), Department of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok 10900, Thailand | [7]Biodiversity Center Kasetsart University (BDCKU), Bangkok 10900, Thailand

Correspondence: Worapong Singchat (worapong.singc@ku.ac.th) | Kornsorn Srikulnath (kornsorn.s@ku.ac.th)

Received: 16 December 2025 | **Revised:** 12 March 2026 | **Accepted:** 2 April 2026

Academic Editor: Aron D. Katz

Keywords: *Clarias* | cryptic species | mitochondrial DNA | species delimitation

ABSTRACT

The complete mitochondrial genome of *Clarias nieuhofii* collected from Phatthalung, Thailand (16,502 bp) was assembled and annotated. The study aimed to understand the mitochondrial genome composition, examine genetic divergence, and explore the phylogenetic relationships within and between populations of *C. nieuhofii*. The goal was also to contribute to the understanding of Sundaland biogeography and inform future conservation strategies. The circular mitogenome, which contained 13 protein-coding genes (PCGs), two rRNAs, 22 tRNAs, and a control region (CR), had 28 genes encoded on the heavy strand, whereas NADH dehydrogenase subunit 6 (*ND6*) and eight tRNAs were encoded on the light strand. The genome was AT-biased (57.85%–57.87%), which is consistent with other *Clarias* species. DNA barcoding (cytochrome c oxidase subunit I [*COI*] and cytochrome b [*Cytb*]) and species delimitation analyses were conducted to assess intra- and interspecific divergence. A positive *COI* barcoding gap (1.91%) but a negative *Cytb* gap (−6.94%) were observed. Delimitation tests variably split the three *C. nieuhofii* groups, with *Cytb* separating all groups and *COI* distinguishing Thai sequences (Group 3: Sumatra, Thailand) from those from Kalimantan (Group 1: Sumatra, Kalimantan). Phylogenetic and molecular analyses recovered three well-supported, geographically structured clades (divergence ~18 million years ago [MYA]), in which Thai and Sumatran samples clustered together, and Kalimantan samples formed distinct clades. These results suggest cryptic diversity within *C. nieuhofii*, underscoring the biogeography of Sundaland and the need for broader sampling and region-specific conservation.

Karnjanapond Sornchai and Artem Lisachov contributed equally to this work.

1 | Introduction

The catfish family Clariidae, comprising approximately 120 recognized species in 16 genera, is distributed across Africa and Asia [1]. Members of this family are characterized by a suprabranchial organ formed by extensions of the second and fourth epibranchials [1]. Aerial oxygen uptake is facilitated by this organ, which allows the catfish to survive in hypoxic habitats such as swamps, floodplains, seasonally drying pools, and rainforests [2–5]. The genus *Clarias*, which is the most diverse within the family, comprises 57 species [6], of which 19 (*Clarias anfractus*, *C. batrachus*, *C. batu*, *C. fuscus*, *C. insolitus*, *C. intermedius*, *C. kapuasensis*, *C. leiacanthus*, *C. macrocephalus*, *C. meladerma*, *C. microspilus*, *C. microstomus*, *C. nieuhofii*, *C. nigricans*, *C. olivaceus*, *C. planiceps*, C. *pseudoleiacanthus*, *C. pseduonieuhofii*, and *C. sulcatus*) are currently recognized in the Southeast Asian group [7]. The genus *Clarias* is characterized by features that include an anguilliform body, long dorsal and anal fins, a flattened bony head, and a broad mouth bearing four pairs of barbels [8]. The slender walking catfish (*C. nieuhofii*) is distributed across southeast Asian countries, including Thailand and Indonesia. This species is characterized by a reddish-brown color with 13 or 14 vertical rows of white spots. Slender walking catfish are considered essential for maintaining a balanced ecosystem within swamp forests because they are one of the few species capable of thriving in this environment [9]. Owing to its attractive body pattern, taste, and high market value, it is often used for food and ornamental purposes. Over the past two decades, its population and density have declined significantly, mainly due to overfishing, habitat loss from peat swamp forest reclamation, and the unregulated use of insecticides in rice fields [10]. Slender walking catfish has been identified as a rare and vulnerable species on the Thailand Red Data Fish List since 2005; however, it remains assessed as least concern on the IUCN Red List, a discrepancy that highlights the urgent need for localized conservation efforts. Remarkably, three major groups of *C. nieuhofii* were observed to be allopatric, with overlapping evolutionary relationships with *C. meladerma*, *C. intermedius*, *C. Pseudonieuhofii*, and *C. kapuasensis*, raising questions about the structure of the species complex among these three groups [11, 12]. Before the resolution of the relationships among several species, *C. nieuhofii* was hypothesized to represent a species complex that contained cryptic species, which may be related to the evolutionary pathway of *Clarias* from the islands to mainland Southeast Asia. More comprehensive data, including specimens from both island and mainland locations across Southeast Asia, are required to clarify this issue.

To date, the characterization of *Clarias* mitogenomes has been limited in scope and less extensive. Eleven complete mitochondrial genomes (mitogenomes) from five species (*C. batrachus*, *C. dussumieri*, *C. fuscus*, *C. gariepinus*, and *C. macrocephalus*) have been reported [13]. No mitogenome data or annotations for *C. nieuhofii* were available. Only specific mitochondrial single nucleotide polymorphism (SNP) variations have been described in most previous studies [13], which lacked comprehensive comparisons and left the mitogenome structural variability in *Clarias* incompletely understood. A definitive evaluation of the hypothesis that the *C. nieuhofii* species complex contains multiple cryptic species is prevented by this knowledge gap, which restricts the availability of comparative data. This study aimed to assemble the complete mitogenome of *C. nieuhofii* from southern Thailand, marking the first complete mitogenome specimen from mainland Southeast Asia. Its structure and variations were compared to those of other *Clarias* mitogenomes. Additionally, a mitogenome-based phylogeny and time tree were constructed to explore the evolutionary patterns and diversification of *C. nieuhofii* and other *Clarias* species. These mitogenomic data are expected to clarify the population genetic structure of *Clarias* species across their distribution, aiding the sustainable conservation and aquaculture of native populations.

2 | Materials and Methods

2.1 | Sample Collection and DNA Extraction

Caudal fin tissue was obtained from four adult *C. nieuhofii* specimens (two males and two females) (Supporting Information 1: Figure S1), which were collected from the breeding stock of the Fisheries Science and Aquatic Resources Program, Faculty of Technology and Community Development, Thaksin University, Phatthalung, Thailand (7°37′0.696″ N, 100°4′22.3788″ E), maintained under the registered aquaculture laboratory (ESPReL Registration No. 2-0560-0003-2) in October 2024. These specimens had been originally captured in Narathiwat Province, Thailand (6°04′21.044″ N, 101°57′53.820″ E) were identified by morphological characters [14]. Tissue samples were excised, preserved in 95% ethanol, and stored at −80 °C until genomic DNA extraction. All animal care and experimental procedures were approved by the Animal Experiment Committee of Kasetsart University (Approval No. ACKU66-SCI-0014). All experiments were conducted in accordance with the Regulations on Animal Experiments at Kasetsart University and the ARRIVE guidelines (https://arriveguidelines.org). Genomic DNA was extracted from the preserved dorsal fin tissue using the salting-out protocol that was described by Supikamolseni et al. [15]. DNA concentration and purity were quantified using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA), and DNA integrity was assessed using 1% agarose gel electrophoresis.

2.2 | Whole Genome Sequencing, Mitogenome Assembly, and Annotation

DNA samples (100 ng/μL) were subjected to next-generation sequencing on an Illumina NovaSeq 6000 (2 × 150 bp paired-end) to a target coverage of 15× at Novogene Co., Ltd. (Singapore). Raw reads were preprocessed using fastp [16], which was used to remove low-quality sequences and generate clean data for assembly. High-quality reads were assembled using GetOrganelle Galaxy version 1.7.7.1, followed by manual inspection to compare the assembled sequences with reference mitogenomes [17]. Annotation was performed using MITOS2 de-novo annotation of metazoan mitochondrial genomes (version 2.1.9) [18]. Reference-based annotation was performed using the mitochondrial genome of *C. gariepinus* (GenBank Accession No. NC_027661) as the reference sequence. Secondary structures of all mitochondrial tRNA genes were predicted using two bioinformatics tools to confirm tRNA location and folding. RNAfold (Galaxy version 2.2.10.4), which calculates the minimum free-energy secondary structures and partition functions, was run on the Aragorn (Galaxy v1.2.36), which predicts tRNA genes, was executed with

default parameters [19]. The predicted secondary structures were visualized using Forna (http://rna.tbi.univie.ac.at/forna/) as described by Kerpedjiev et al. [20]. The genomic sequences were deposited in the NCBI database (https://www.ncbi.nlm.nih.gov, accessed on 01 April 2026) under the accession numbers PX395444, PX395445, PX395446, and PX395447. The mitogenome map was visualized using OGDRAW 1.3.1 [21]. Three conserved regions, the terminal-associated sequence (TAS), central conserved domain, and conserved sequence blocks (CSBs), were identified in both control regions (CRs) by recognizing motifs that were similar to those described in other vertebrates [22]. Tandem repeats in the CRs, including motifs, repeat length, and copy number, were detected using Tandem Repeats Finder v4.09. Additionally, raw sequence data for a *C. nieuhofii* sample from Kalimantan (SRR30682833, BioProject PRJNA1155604) were retrieved from GenBank (accessed July 22, 2025), which were used to perform a parallel mitogenome annotation.

2.3 | Comparative Sequence Analysis and Base Composition and Skewness Analysis

The four complete mitogenome sequences of *C. nieuhofii* generated in this study were included in the analyses, along with the mitogenomes of other *Clarias* species. Nucleotide composition (A%, T%, C%, and G%), AT content (A + T), and GC content (G + C) of each mitogenome were calculated using MEGA version 12 [23]. Strand asymmetry was computed using the formulas GC skew = (G − C)/(G + C) and AT skew = (A − T)/(A + T) [24]. The observed skewness, which may reflect differential mutational pressures arising from asymmetric mitogenome replication, has been previously recorded [25].

2.4 | Codon Usage and Selective Pressures Across Protein-Coding Gene (PCG)

The codon usage parameters of PCGs in *C. nieuhofii* and other *Clarias* species included the relative synonymous codon usage (RSCU), GC content at the first and second codon positions (GC12), GC content at the third codon position (GC3), and the effective number of codons (ENc). RSCU analysis was performed with the R package "cubar" [26], whereas GC12, GC3, and ENc were calculated with DAMBE7 [27]. RSCU is defined as the ratio between the observed frequency of a synonymous codon and the expected frequency when all codons for an amino acid are used evenly, indicating a deviation from even usage (RSCU >1 = preferred; RSCU <1 = less frequent; RSCU = 1 = no bias) [28]. Neutrality plots comparing GC12 and GC3 were used to assess the relative influence of mutation pressure and natural selection on codon usage [29]. ENc, which ranged from 20 (extreme bias) to 60 (neutral usage), was examined along with ENc − GC3s plots to infer mutational versus selective effects [30]. The P2 index, which estimates codon–anticodon interaction efficiency, was calculated as (WWC + SSU)/(WWY + SSY) (W = A or U; S = C or G; Y = C or U), and P2 >0.5 was interpreted as evidence of translational selection [31].

2.5 | Signature of Selection Analysis

The ratio of nonsynonymous (Ka) to synonymous (Ks) substitution rates (Ka/Ks) was used to assess the evolutionary pressure on the PCGs in *C. nieuhofii* and other *Clarias* species, with each gene analyzed separately. Ka/Ks ratios were interpreted as follows: values greater than 1 indicated positive selection, values less than 1 indicated purifying selection, and values less than 1 indicated neutral evolution [32]. The Ka and Ks, as well as their ratios, were calculated using DnaSP version 6.12.03 [33].

2.6 | Phylogenetic Tree and Molecular Dating

Twelve concatenated mitochondrial PCGs, excluding the NADH dehydrogenase subunit (*ND6*) sequence located in the light chain, along with the cytochrome c oxidase subunit I (*COI*) and cytochrome b (*Cytb*) gene datasets, were independently aligned using MAFFT v7.490 [34]. The best-fit substitution models were selected using ModelFinder, based on the Bayesian information criterion [35]. The models identified were GTR + F + I + G4 and TVM + F + G4 for the CR dataset, where GTR represents the general time-reversible model, TVM indicates the transversion model, F signifies the use of empirical frequencies, I indicates invariant sites, and G denotes the discrete gamma distribution. Phylogenetic trees were constructed using maximum likelihood (ML) with 10,000 ultrafast bootstrap replicates in IQ-TREE v1.6.12 [36, 37]. Bayesian inference (BI) was performed with Markov chain Monte Carlo (MCMC) runs of 2,000,000 generations, sampling every 1000 generations, with the first 25% discarded as burn-in, in MrBayes v3.2.6 [38] via the CIPRES Science Gateway. The analysis was conducted under the GTR + F + I + G4 model. For the *COI* dataset, *Heteropneustes fossilis* and *Bagrus bajad* (GenBank JX983313 and MK335908) were used as outgroups. For *Cytb*, *Heteropneustes fossilis* and *Bagre panamensis* (GenBank PV617374 and AJ580999) were used as outgroups. For the PCGs dataset, *Heteropneustes fossilis* and *Bagre panamensis* (GenBank KP345880 and NC_037470) served as outgroups. All the datasets were selected using the same analytical framework. The mean sequence divergence of the annotated PCGs of the studied mitogenomes was estimated using MEGA version 12 [23] with the Kimura 2-parameter (K2P) substitution model. Divergence times were estimated for three datasets (PCGs, *COI*, and *Cytb*) using BEAST2 in XSEDE version 2.7.3 [39], with calibration points derived from the fossil-calibrated nodes of Cobitoidei. The earliest cobitoid fossils from the family Catostomidae date back to the early Paleocene (61.1–65.5 million years ago [MYA]). Accordingly, the lower age constraint for Cobitoidei was set at 61.1 MYA, with a soft upper bound of 140.2 MYA, modeled using an exponential distribution with a mean of 21.45 [40]. Another calibration point was set at the first appearance of African Clariidae in the Lower Eocene (34–56 MYA, mean ≈ 45 MYA), based on fossil evidence [41]. The GTR model was used, along with a log-normal relaxed clock algorithm, for the Yule speciation process [42]. Markov chain MCMC simulations were run for 10 million iterations, and convergence was verified using LogCombiner version 1.8.4 [43]. A maximum clade credibility tree was generated using TreeAnnotator version 1.8.4.

2.7 | Species Identification With DNA Barcoding

Mitochondrial gene sequences, which focus on the *COI* and *Cytb* regions, commonly used for clariid catfish identification and phylogenetic analysis [44], were analyzed for *C. nieuhofii* specimens

obtained in this study. These data were combined with publicly available *C. nieuhofii* sequences retrieved from the NCBI, a Biotechnology Information database, for further analysis. Each dataset was aligned using the Geneious Prime software (version 2025.2.2). Sequences containing deletions or gaps were excluded from further analysis to avoid contamination with nuclear mitochondrial DNA. Because sequence lengths varied, the most frequently occurring lengths were selected for downstream analyses [45]. The selected *COI* and *Cytb* coding sequences were translated into amino acid sequences using MEGA version 12 to verify the presence of open reading frames without stop codons [23]. Alignment length, number of variable positions, and GC content were compared among the datasets. Sequence divergence among individual sequences was calculated using the K2P model, which is commonly used in barcoding studies, to evaluate the discriminatory power of the distance- and tree-based approaches [46]. Both distance- and tree-based approaches, which were used to examine genetic structure, were applied to the *C. nieuhofii* species Group. Pairwise divergences were computed using the dist.dna function in the R package ape [47–50]. The nearest-neighbor test was performed using the nearest neighbor function in the R package spider to determine whether the shortest divergences were among conspecific sequences, even in *C. nieuhofii* [49]. The percentage of correctly identified sequences was calculated by dividing the number of sequences with an intraspecific nearest neighbor by the total number of sequences. The barcoding gap was determined by calculating the difference between the minimum and maximum intraspecific divergence using the non-ConDist and maxInDist functions in spider [49]. The barcoding gaps among markers were compared using the Kruskal–Wallis test to assess significant differences between the *COI* and *Cytb* datasets [51], with Dunn's post hoc test and Bonferroni correction for multiple comparisons [52]. Markers with high discriminatory power were identified by a high percentage of correct identifications in the nearest-neighbor test and a positive barcoding gap.

Phylogenetic analyses were conducted using the BI method, which was applied to reconstruct phylogenies based on *COI* and *Cytb* datasets, allowing for the assessment of topological congruence across various programs and algorithms. The outgroup sequences for the *COI* dataset, *Heteropneustes fossilis* and *Bagrus bajad* (GenBank accessions JX983313 and MK335908), were used as outgroups, whereas *Heteropneustes fossilis* and *Bagre panamensis* (GenBank accessions PV617374 and AJ580999) served as outgroups for the *Cytb* dataset. Sequences were aligned using Clustal Omega in Geneious Prime 2025.2.2. Phylogenetic analysis was performed with BI using MCMC runs of 2,000,000 generations, sampling every 1000 generations, with the first 25% discarded as burn-in in MrBayes 3.2.6 [38] via the CIPRES Science Gateway, with the best-fit model GTR + F + I + G4. The resulting trees were visualized and edited using the online tools One Table (tvBOT) [53] and Interactive Tree of Life (iTOL) v6 [54]. The number of monophyletic Groups was calculated using the R package spider, which was used to evaluate monophyly with the "monophyly" function [49]. A tree-based barcoding efficiency test was performed to evaluate the ability of gene markers to recover monophyly among sequences of the same species, without considering the relationships among other taxa. This approach was also used to test whether all specimens of *C. nieuhofii* could be grouped regardless of their classification.

2.8 | Species Delimitation

The delimitation of *C. nieuhofii* was assessed using two methods: the Bayesian Poisson tree process (bPTP) [55] and the multirate Poisson tree process (mPTP) [56]. The *COI* and *Cytb* datasets were tested separately. For the bPTP (available in https://species.h-its.org/) and mPTP (https://mptp.h-its.org/#/tree) analyses, an ML tree built in the software RAxML version 8.2.10 [57] was used as input, using 100,000 Monte Carlo Markov chains (MCMC) generations, thinning = 100, and 10% of burn-in. The bPTP adds Bayesian support to the nodes from a tree considering similar diversification rates [55], while mPTP assumes distinct evolutionary rates and relies on the Akaike information criterion (AIC) to infer the number of species in the input trees [56].

3 | Results

3.1 | Mitogenome Structure of *C. nieuhofii*

The complete mitogenome of *C. nieuhofii*, which was 16,501–16,502 bp in length (Supporting Information 1: Figures S2–S5), included 37 genes and one CR. This genome consisted of 13 PCGs, two rRNA genes (12S and 16S rRNA), 22 tRNA genes, and a CR (Supporting Information 2: Table S1; Supporting Information 1: Figure S6). All PCGs within the mitogenome, except for the *COI* gene, which uses GTG as the start codon, begin with a common ATG start codon. Complete stop codons, such as TAA and TAG, were observed in *ND1*, *ND2*, *ND3*, *ND5*, *ND6*, *ND4L*, *COI*, *ATP8*, and *ATP6*, whereas incomplete stop codons (TA or T) were detected in *ND4*, *COII*, *COIII*, and *Cytb*. Among the PCGs, *ND5* was the longest at 1849–1850 bp and *ATP8* was the shortest at 170bp (Supporting Information 1: Figures S2–S5, Supporting Information 2: Table S1). The base composition of the *C. nieuhofii* PCGs was 30.41%–30.43% A, 27.43%–27.44% T, 27.41%–27.42% C, and 14.71%–14.74% G, which resulted in an A + T bias of 57.85%–57.87% (Supporting Information 2: Table S2; Supporting Information 1: Figure S7). A positive AT skew and a negative GC skew were exhibited by the PCGs, except *ND6* and *ND4L*, which showed a negative AT skew (Supporting Information 2: Table S3; Supporting Information 1: Figure S8). The large noncoding region located between the tRNA-Thr and tRNA-Pro genes in *C. nieuhofii* (936 bp) was assumed to be a putative CR (Supporting Information 1: Figures S2–S5; Supporting Information 2: Table S1). Three functional domains were identified in the putative CR. These included (1) the extended TAS (ETAS) domain, which contained ETAS regions, (2) the central domain, which included CB-1, CB-2, CB-3, and CB-4, and (3) the CSB domain, which comprised CSB-1, CSB-2, and CSB-3 (Supporting Information 1: Figure S9).

3.2 | Codon Usage and Selective Pressures Across PCGs

Codons encoding leucine and serine were most frequently used in *C. nieuhofii* (Supporting Information 1: Figure S10), and codon usage bias was observed across other *Clarias* species (Supporting Information 1: Figure S11). The highest RSCU values were observed for GUA (Val), AAA (Lys), AGC (Ser), GAA (Glu), and CUA (Leu), whereas the lowest values were observed for AAU (Asn) and CAC (His). Twenty-five codons were found to have RSCU >1 (preferential usage); of these, 14 were terminated

with A or T and 11 were terminated with G or C. Methionine and tryptophan, which lack synonymous codons, were unique, and two codons were observed with RSCU = 1 (unbiased usage). Neutrality-plot analysis, which compared GC12 versus GC3, revealed GC3 values that ranged from 23.1% to 46.5% and produced a slope of 0.16 (Supporting Information 1: Figure S12a). ENc–GC3s analysis that all PCGs in the *Clarias* mitogenome fell below the expected ENc curve (Supporting Information 1: Figure S12b), with factors other than mutation influencing codon usage. The average ENc values across the PCGs were greater than 40 (Supporting Information 1: Figure S12c; Supporting Information 2: Table S4), which was consistent with a moderate codon bias. A bias toward cytosine over uracil at the third codon position was indicated by the P2 index, which had a mean value of <0.5 for 13 PCGs of *C. nieuhofii* (Supporting Information 2: Table S5).

3.3 | Selection of Clarias Mitogenome and Genetic Distances

Ka/Ks analysis across *Clarias* species showed that *ND3* had the highest average Ks, whereas *COI* had the lowest average Ka (Supporting Information 2: Table S6). Across the 13 PCGs, the Ka/Ks ratios ranged from 0.003 to 0.251, consistent with strong purifying selection acting on these genes (Supporting Information 2: Table S7; Supporting Information 1: Figure S13).

3.4 | Phylogenetic Relationships and Molecular Dating

The evolutionary distance estimation among *Clarias* species indicated that *ND3* exhibited the largest K2P genetic distance, followed by *ND4* and *ND5*, while *ATP8* showed the smallest value (Supporting Information 2: Table S7). Bayesian divergence-time analyses indicated that Asian and African *Clarias* lineages diverged during the Eocene epoch, were estimated to have diverged at marker-specific times, which were 40–65 MYA based on *COI* and 35–51 MYA based on *Cytb*; the PCGs-based estimate was 37–42 MYA (Figures 1–3). Within the Asian *Clarias*, *C. nieuhofii* was observed as a species complex comprising three clades: *C. nieuhofii* Group 1, which is distributed in Kalimantan and Sumatra; *C. nieuhofii* Group 2, which was restricted to Kalimantan; and *C. nieuhofii* Group 3, which was found on Sumatra Island and in the southern part of Thailand. Groups 1 and 2 were placed in the same major clade, which clustered with *C. intermedius*, whereas Group 3 was located in a separate major clade closely related to *C. meladerma* and *C. maurus* (possibly *C. teijsmanni*).

3.5 | DNA Barcoding

COI and *Cytb* sequences compiled for barcoding varied in length, creating trade-offs. To account for intraspecific variability, most specimens were represented by multiple sequences, as previously reported [44]. The *COI* sequences had a length of 552 bp, with 134 variable sites, and a GC content of 43.39%. The *Cytb* sequence was 596 bp in length, containing 398 variable sites, and a GC content of 42.87%. For consideration with *C. nieuhofii*, intraspecific K2P divergences were 0%–8.49% (mean 1.27% ± 1%) and interspecific K2P divergences between *C. nieuhofii* and other *Clarias* species were 10.40%–16.24% (mean 13.53% ± 1%) based on the *COI* dataset. For *Cytb*, intraspecific K2P divergences were also 0%–9.17% (mean 1.96% ± 2%), and interspecific K2P divergences were 2.23%–13.46% (mean 8.57% ± 2), producing a positive barcoding gap in *COI* (1.91%) and a negative barcoding gap in *Cytb* (−6.94%) (Supporting Information 1: Figure S14; Supporting Information 2: Table S8). The two datasets differed significantly in barcoding gap size (Kruskal–Wallis test, $\chi^2 = 3.94$, $p < 0.05$), and pairwise comparison indicated significant differences between *COI* and *Cytb* barcoding gaps. Nearest-neighbor analysis of *C. nieuhofii* showed that 100% of *COI* and *Cytb* sequences had an intraspecific nearest neighbor (Supporting Information 1: Figure S14). Within- and between-groups of *C. nieuhofii*, *Cytb* sequence divergences were observed, which ranged from 0.70%–1.09% (within groups) and 3.75%–7.92% (between Groups). The pairwise sequence divergences were 3.75% (Group 1 vs. Group 2), 6.93% (Group 1 vs. Group 3), and 7.92% (Group 2 vs. Group 3). Pairwise comparisons produced $Z = 0.71$ ($p = 1$) for Group 1 versus Group 2, $Z = 1.41$ ($p = 0.472$) for Group 1 versus Group 3, and $Z = 0.71$ ($p = 1$) for Group 2 versus Group 3, indicating that the observed *Cytb* divergences among the three Groups were not statistically significant. The maximum *COI* divergence, which was observed between Group 3 (Thailand) and Group 1 (Kalimantan), resulted in a pairwise comparison $Z = 2.05$ ($p = 0.385$), indicating that the divergence between these Groups was not statistically significant. *COI* sequence data for Group 2 were unavailable from the public data repository. Phylogenetic relationships reconstructed from the *COI* and *Cytb* datasets were strongly supported for each of the analyzed *Clarias* species. Well-supported monophyletic clades were formed by the *C. nieuhofii* specimens from each of the three Groups. *C. intermedius* and *C. meladerma* were positioned between these groups, with *C. intermedius* being closely related to Group 3 (Supporting Information 1: Figures S15, S16).

3.6 | Species Delimitation

Species delimitation was performed using the bPTP and mPTP methods applied to the *COI* and *Cytb* datasets. The bPTP method supported 17 and 44 species in the *COI* and *Cytb* sequence datasets, respectively. The mPTP method supported 1 and 15 species for the *COI* and *Cytb* sequence datasets, respectively (Figures 1 and 2). Considering the *C. nieuhofii* species complex, bPTP delimited the *COI* sequences of *C. nieuhofii* from Thailand (Group 3) as distinct from those from Kalimantan (Group 1), whereas mPTP grouped Groups 1 and 3 into the same species. For the *Cytb* dataset, bPTP separated all three *C. nieuhofii* Groups as distinct species, in contrast to mPTP, in which Groups 1 and 2 were inferred to be the same species, whereas Group 3 was recognized as distinct.

4 | Discussion

The complete mitogenome of *C. nieuhofii* from Phatthalung, Thailand (16,501–16,502 bp) was determined, and the mitogenome from Borneo, Indonesia (16,514 bp) was analyzed, which showed a slight difference in length between the populations. Genome sizes similar to those of *C. gariepinus* (16,508 bp), *C. camerunensis* (16,511 bp), *C. batrachus* (16,511 bp), *C. dussumieri* (16,514 bp), *C. fuscus* (16,525 bp), and *C. macrocephalus* (16,511 bp) were observed [13]. The mitogenome structure and

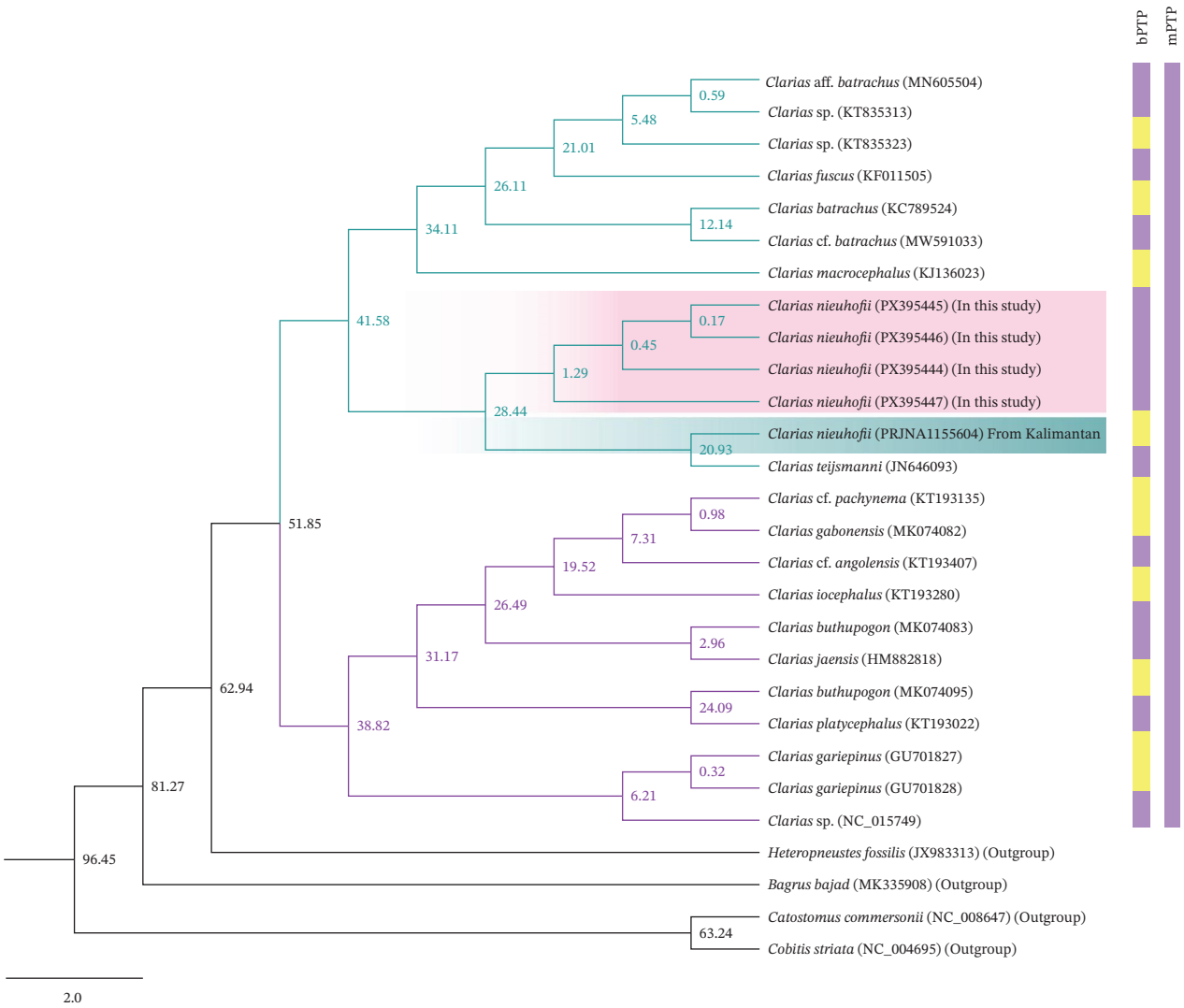

FIGURE 1 | Bayesian phylogenetic tree with divergence time estimates. The fossil-calibrated chronogram was inferred from mitochondrial cytochrome c oxidase subunit I (*COI*) sequences under a relaxed molecular clock. Branch lengths represent divergence times, and node bars indicate the 95% credible intervals for node ages. Posterior probabilities are shown above the branches. Fossil-calibrated nodes for Cobitoidei. *Heteropneustes fossilis* and *Chrysichthys auratus* were used as outgroups. Color bars on the right of each clade correspond to the results of species delimitation analyses based on bPTP and mPTP model.

organization of *C. nieuhofii* were consistent with the highly conserved mitogenome organization of 37 genes and one CR in *Clarias* species [13]. The use of start and stop codons in the mitogenomes of *C. nieuhofii* was analyzed to gain insights into their translational regulation and evolution. Codon usage in mitochondrial genes was evaluated using RSCU, which revealed diverse patterns; leucine and serine were the most frequently used amino acids in *Clarias*. This pattern, which reflects a preference for leucine, a branched-chain amino acid (BCAA) involved in mitochondrial signaling pathways [58]—may be related to natural selection. Further highlighting the conserved role of leucine usage across diverse taxa, several studies have shown that leucine and serine are the most abundant amino acids in the mitochondrial PCGs of fishes. In clupeoid fish, leucine accounts for the highest proportion of amino acids, reflecting a codon usage bias toward this residue [59]. Similarly, in the Reba carp (*Cirrhinus reba*), comparative mitogenomic analyses demonstrated that leucine and serine were the most frequent amino acids across 13 PCGs [60, 61]. Analyses of three Carangidae species have also confirmed that serine and leucine, each encoded by six codons, are the most abundant residues in mitochondrial PCGs [62]. Comparable findings have been reported in novel gene rearrangement studies of teleost mitogenomes, in which leucine constituted the highest proportion of amino acids [63]. In avian mitogenomes, accipitrids exhibit the highest RSCU values for CUA (Leu), with a parallel codon bias at the third codon position [64]. While a strong codon usage bias toward leucine is observed in vertebrates, arthropod lineages show preferential use of other amino acids, such as serine and arginine, reflecting lineage-specific codon biases [65]. BCAAs, such as leucine, isoleucine, and valine, are required by a wide range of eukaryotes [66]. All Ka/Ks ratios of the 13 PCGs were below 1, indicating strong purifying selection, which is consistent with previous findings for all PCGs in *Clarias* species [67, 68]. This agreed with the neutrality plot, which indicated a GC3 slope of 0.16, suggesting that 16% mutation pressure and 84% selection pressure influenced the codon usage bias in the mitogenome of *Clarias* species. ENC versus GC3 analysis, which was performed

FIGURE 2 | Bayesian phylogenetic tree with divergence time estimates. The fossil-calibrated chronogram was inferred from mitochondrial cytochrome b (*Cytb*) sequences under a relaxed molecular clock. Branch lengths represent divergence times, and node bars indicate the 95% credible intervals for node ages. Posterior probabilities are shown above the branches. Fossil-calibrated nodes for Cobitoidei. *Heteropneustes fossilis* and *Chrysichthys auratus* were used as outgroups. Color bars on the right of each clade correspond to the results of species delimitation analyses based on bPTP and mPTP model.

FIGURE 3 | Bayesian phylogenetic tree with divergence time estimates. The fossil-calibrated chronogram was inferred from mitochondrial protein-coding genes (PCGs) sequences under a relaxed molecular clock. Branch lengths represent divergence times, and node bars indicate the 95% credible intervals for node ages. Posterior probabilities are shown above the branches. Fossil-calibrated nodes for Cobitoidei. *Heteropneustes fossilis* and *Chrysichthys auratus* were used as outgroups.

to assess codon bias, showed that most points for the 13 PCGs of *Clarias* species fell below the expected curve, indicating the involvement of selection, possibly driven by adaptive evolution, significantly influences codon usage patterns in mitochondrial genes, rather than mutations alone [69]. The ENc, which measures codon bias and ranges from 20 (strong bias) to 60 (neutral), was used to assess codon usage. Values above 40 for the 13 PCGs of *Clarias* species indicated relatively neutral codon usage with low bias. A high ENc value, which reflects the use of synonymous codons, suggests no strong preference for specific codons during translation [30].

The potential presence of cryptic species in *C. nieuhofii* was assessed using a barcoding approach. It has been indicated that in over 1000 teleost species, the intraspecific sequence divergence in the *COI* gene is 3% or less [70]. Previously, a positive barcoding gap between interspecific and intraspecific sequence divergences of *Clarias* was observed in the *Cytb* dataset, which was not detected in the *COI* dataset. Intraspecific sequence divergence is typically <4.4% based on *Cytb* sequences, whereas interspecific sequence divergence is generally >66.9% among *Clarias* species [44].

Considering only a small dataset of *C. nieuhofii* in this study, a positive *COI* barcoding gap was observed because samples from only *C. nieuhofii* Groups 1 and 3 (a total of five samples) were included. In contrast, the *Cytb* dataset showed a negative barcoding gap (−6.94%) because it included representatives from *C. nieuhofii* Groups 1 (Sumatra–Kalimantan), 2 (Kalimantan), and 3 (Sumatra–Thailand), resulting in greater variation in intraspecific sequence divergence. These results are consistent with the species delimitation analyses. In the *Cytb* dataset, bPTP separated all three *C. nieuhofii* Groups as distinct species, whereas mPTP inferred Groups 1 (Sumatra–Kalimantan) and 2 (Kalimantan) as the same species, but recognized Group 3 (Sumatra–Thailand) as distinct (Supporting Information 1: Figure S4),

supporting the interpretation of distinct lineages. The observed negative barcoding gap in the *Cytb* dataset suggests that standard distance-based thresholds may not be fully applicable in this context. Furthermore, as the pairwise *Z*-tests did not reach statistical significance, these results are interpreted as reflecting deep phylogenetic structure within the *C. nieuhofii* complex rather than providing definitive evidence for species delimitation.

The differing barcoding gap patterns between *COI* and *Cytb* in this study highlight how dataset composition and marker choice affect species delimitation. The positive *COI* gap was due to limited sampling (Groups 1 and 3), reducing intraspecific variation and exaggerating divergence differences. In contrast, *Cytb* included more Groups (1, 2, and 3), showing a negative gap from higher intraspecific divergence. Therefore, broader sampling and the integration of nuclear genomic data will be required to more accurately resolve species boundaries within the *C. nieuhofii* complex. This shows that barcoding gaps depend heavily on sampling and population structure. Relying on a single marker or limited data can be misleading; *Cytb* revealed hidden diversity within *C. nieuhofii* not detected by *COI*. Thus, using multiple markers and broad sampling is key for accurate species delimitation and evolutionary insights [71]. Such low sequence divergence within species complexes or cryptic species has been predicted in clariid catfish, such as African catfish or walking catfish [44, 72]. High levels of intraspecific sequence divergence observed in clariid catfish, including *C. nieuhofii*, suggest the presence of cryptic species. Although pairwise comparisons between *C. nieuhofii* Groups showed no significant differences ($p > 0.05$), the divergence between groups exceeded the threshold (4.4%), indicating that the groups, particularly Group 3 versus Groups 1 and 2, may represent distinct species. This finding was supported by tree-based analysis that clustered each *C. nieuhofii* Group into a single monophyletic clade. Therefore, it is likely that the three *C. nieuhofii* Groups represent deeply divergent mitochondrial lineages: Group 3 versus Groups 1–2, and a closely related species complex between Groups 1 and 2, which are closely related. Future integration of nuclear genomic data will be essential to further validate the taxonomic status of these divergent lineages.

Remarkably, all *C. nieuhofii* groups occur in Sundaland, consistent with the region's strong phylogeographic structuring among islands [11]. Although *C. nieuhofii* Group 1 comprised populations from Kalimantan and Sumatra, which were estimated to have diverged at 3–7 MYA based on *Cytb* sequences, they were assessed as conspecifics based on sequence divergence and species delimitation [11]. Comparable island-driven divergences have been reported, with forest rodents dated at 3–4 MYA [73], orangutans at 2–5 MYA [74], and bony lip barb, *Osteochilus vittatus*, at 2 MYA [75]. A divergence of less than 5 MYA may not represent sufficient time for complete lineage separation under allopatric conditions, which is required for species-level divergence [76]. Sundaland's present biodiversity was established in the late Pliocene and was only minimally affected by Pleistocene sea-level fluctuations, as supported by paleoenvironmental evidence [77]. During low sea-level periods, a savanna corridor, as proposed by Heaney [78], would have been critical for *Clarias*, some species (*C. intermedius*, *C. leiacanthus*, and *C. nieuhofii*) that depend on blackwater habitats maintained by closed rainforests. A 50–150 km wide savanna corridor across Sundaland was inferred by Bird et al. [79] based on multiple lines of evidence, which implies that blackwater streams were unlikely to have connected Sumatra, Java, and Kalimantan during glacial periods. Because blackwater *Clarias* lineages generally exhibit limited dispersal ability across unsuitable habitats [80], this phenomenon may explain why populations of *C. nieuhofii* from the two islands remained conspecific despite their geographic separation. Similar to the historical isolation observed in other Sundaic blackwater taxa, such as the leaf catfish (*Silurichthys insulanus*) and walking catfish (*Clarias rennyae*) [80–83], the restricted dispersal capacity of these catfish is likely to have prevented interisland colonization and maintained genetic cohesion within the species. However, the spatial and temporal extent of open vegetation across Sundaland remains debatable and requires further study [84]. By contrast, compelling evidence was obtained that the *C. nieuhofii* populations from Kalimantan (Groups 1 and 2), which were estimated to have diverged at 2–5 MYA based on *Cytb* sequences, and from Sumatra, Thailand (Group 3), which were estimated to have diverged at approximately 15–19 MYA, represent genetically distinct lineages, indicating true evolutionary separation rather than merely geographic variation. These results were supported by the species delimitation analysis. Sequence divergence within *C. nieuhofii* Group 3, which is distributed in Sumatra and Thailand, was 1% based on *Cytb*, with an estimated divergence time of approximately 3 MYA. This short divergence time is comparable to that observed in other species or animal groups, which are located on different sites between Sumatra Island and mainland Southeast Asia, but are still considered the same species. Similarly, *Osteochilus vittatus* (bony lip barb), was divided into Indochina (Vietnam and Thailand) and Sunda-shelf (Malaysia and Indonesia) clades that were estimated to have split at approximately 2 MYA. Despite the high dispersal capacity of this species, its phylogeographic structure across Sundaland has been shaped by historical events, long-term geographic isolation associated with Pleistocene sea-level fluctuations, and restricted gene flow resulting from limited habitat connectivity [75]. The cryptic diversity within *C. nieuhofii* indicates a complex evolutionary history and highlights the need for region-specific conservation strategies.

5 | Conclusions

Four complete mitogenomes of *Clarias nieuhofii* obtained from Phatthalung (mainland Thailand) and Borneo were assembled and annotated, thereby expanding the mitogenomic resources for *Clarias*. The mitogenome structure and gene content were highly conserved among *Clarias* species, displaying an A + T bias and codon usage patterns dominated by leucine and serine. Codon usage and Ka/Ks analyses indicated a predominantly purifying selection of mitochondrial PCGs. Phylogenetic and molecular dating analyses identified Asian and African *Clarias* as distinct lineages and revealed three well-supported, geographically structured *C. nieuhofii* clades. Species delimitation and barcoding results indicated that Group 3 (Sumatra–Thailand) is different from Groups 1 to 2 (Kalimantan–Sumatra), which may form a species complex. Broader sampling and nuclear data are recommended to resolve the species boundaries and guide conservation.

Nomenclature

COI:	Cytochrome c oxidase subunit I
Cytb:	Cytochrome b
PCGs:	Protein-coding genes
RSCU:	Relative synonymous codon usage
Enc:	Effective number of codons
ETAS:	extended terminal-associated sequences.

Author Contributions

Conceptualization, methodology, writing – original draft preparation, visualization: Karnjanapond Sornchai, Worapong Singchat, and Kornsorn Srikulnath. Validation: Karnjanapond Sornchai, Suphada Kiriratnikom, Worapong Singchat, and Kornsorn Srikulnath. Formal analysis: Karnjanapond Sornchai, Artem Lisachov, and Thitipong Panthum. Investigation: Karnjanapond Sornchai, Narongrit Muangmai, Prateep Duengkae, Worapong Singchat, and Kornsorn Srikulnath. Resources: Suphada Kiriratnikom, Thitipong Panthum, and Worapong Singchat. Writing – review and editing: Karnjanapond Sornchai, Artem Lisachov, Thitipong Panthum, Suphada Kiriratnikom, Narongrit Muangmai, Prateep Duengkae, Worapong Singchat, and Kornsorn Srikulnath. Supervision, project administration, funding acquisition: Worapong Singchat and Kornsorn Srikulnath.

Acknowledgments

Data processing and mitochondrial genome assembly were performed using the Europe Galaxy server (https://usegalaxy.eu), which is supported by the German Federal Ministry of Education and Research (Grant 031L0101C) and de.NBI-epi. We acknowledge the Galaxy Community for providing an open platform that enables accessible, reproducible, and collaborative data analysis (The Galaxy Community, 2024. *Nucleic Acids Research*, gkae410. https://doi.org/10.1093/nar/gkae410).

Funding

This study was granted by the Faculty of Science, Kasetsart University through the Graduate Fellowship Program (Grant 6717400111), awarded to Karnjanapond Sornchai and Worapong Singchat, the Kasetsart University Institute for Advanced Studies (KUAIS) through the Reinventing University Program (Grant RUP69_Act1_No.03) awarded to Karnjanapond Sornchai and Worapong Singchat, the Postdoctoral Fellowship from Kasetsart University awarded to Artem Lisachov and Worapong Singchat, the National Research Council of Thailand: High-Potential Research Team Grant Program (Grant N42A660605) awarded to Worapong Singchat, Thitipong Panthum, Narongrit Muangmai, Prateep Duengkae, and Kornsorn Srikulnath, the Program Management Unit for Human Resources and Institutional Development and Innovation (PMU-B) (Contract Number B13F670053), awarded to Worapong Singchat, Artem Lisachov, Thitipong Panthum, Narongrit Muangmai, Prateep Duengkae, and Kornsorn Srikulnath, the Program Management Unit for Competitiveness (PMU-C) under the Global Partnership Program (Contract Number C23F670224), awarded to Worapong Singchat, Thitipong Panthum, Narongrit Muangmai, Prateep Duengkae, and Kornsorn Srikulnath, a grant from the Kasetsart University Research and Development Institute (Contract Number FF(KU)61.69) awarded to Worapong Singchat and Kornsorn Srikulnath, the International SciKU Branding (ISB), Faculty of Science, Kasetsart University awarded to Worapong Singchat and Kornsorn Srikulnath, the Ministry of Science and Higher Education of the Russian Federation (Grant FWNR-2026-0024) awarded to Artem Lisachov.

Disclosure

All authors have read and agreed to the published version of the manuscript. No funding source was involved in the study design, collection, analysis, and interpretation of the data, writing of the report, or decision to submit the article for publication.

Ethics Statement

All animal care, specimen handling, and experimental procedures were approved by the Animal Experiment Committee of Kasetsart University (Approval Number ACKU66-SCI-0014). Fin tissues were obtained from the breeding stock of the Fisheries Science and Aquatic Resources Program, Faculty of Technology and Community Development, Thaksin University, Phatthalung, Thailand, which is maintained under the registered aquaculture laboratory (ESPReL Registration Number 2-0560-0003-2). All experiments were conducted in accordance with the Regulations on Animal Experiments at Kasetsart University and complied with the ARRIVE guidelines (https://arriveguidelines.org).

Consent

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The datasets analyzed during the current study are available from the corresponding author upon reasonable request.

References

1. G. G. Teugels and D. Adriaens, "Taxonomy and Phylogeny of Clariidae–An Overview," in *Catfishes*, eds. G. Arratia, B. G. Kapoor, M. Chardon, and R. Diogo, (Science Publishers, 2003): 465–487.

2. P. H. Greenwood, "A Revision of the Genus, *Dinotopterus*, Blgr. (Pisces; Clariidae) With Notes on the Comparative Anatomy of Suprabranchial Organs in the Clariidae," *Bulletin of the British Museum (Natural History) Geology* 7 (1961): 215–241.

3. J. B. Graham, *Air-Breathing Fishes: Evolution, Diversity and Adaptation* (Academic Press, 1997).

4. H. Matthes, "Les poissons du la Tumba et de la region d'Ikela: etude systematique, ecologique et zoogeographique: resultats scientifiques d'une mission IRSAC dans la cuvette centrale congolaise," *Annales du Musée Royal de l'Afrique Centrale, Série Zoologie* 8 (1964): 1–199.

5. W. E. Burgess, *An Atlas of Freshwater and Marine Catfishes: A Preliminary Survey of the Siluriformes* (TFH Publications, 1989).

6. H. H. Ng and M. Kottelat, "*Clarias serniosus*, a New Walking Catfish (Teleostei: Clariidae) From Laos," *Zootaxa* 3884, no. 5 (2014): 437–444.

7. H. H. Ng, K. H. Dang, and V. T. Nguyen, "*Clarias gracilentus*, a New Walking Catfish (Teleostei: Clariidae) From Vietnam and Cambodia," *Zootaxa* 2823, no. 1 (2011): 61–68.

8. N. H. Hee, "Two New Species of Catfishes of the Genus *Clarias* From Borneo (Teleostei: Clariidae)," *The Raffles Bulletin of Zoology* 47, no. 1 (1999): 17–32.

9. M. Lertworapreecha, C. Tuyarum, A. Tanomtong, C. Suwannapoom, and S. Kiriratnikom, "Establishment of Cell Lines From Slender Walking Catfish *Clarias nieuhofii* (Valenciennes, 1840)," *Walailak Journal of Science and Technology (WJST)* 18, no. 12 (2021): 10857.

10. J. Miettinen, C. Shi, and S. C. Liew, "Two Decades of Destruction in Southeast Asia's Peat Swamp Forests," *Frontiers in Ecology and the Environment* 10, no. 3 (2012): 124–128.

11. L. Pouyaud, Sudarto, and E. Paradis, "The Phylogenetic Structure of Habitat Shift and Morphological Convergence in Asian *Clarias* (Teleostei, Siluriformes: Clariidae)," *Journal of Zoological Systematics and Evolutionary Research* 47, no. 4 (2009): 344–356.

12. S. Lavoué, J. A. F. Jamaluddin, A. H. Muhammad-Rasul, M. L. Ilham-Norhakim, and K. M. Zain, "Mitochondrial Evidence on the Phylogenetic Position of the Southeast Asian Catfish Genus, *Encheloclarias*, Myers, 1937 (Actinopterygii: Siluriformes: Clariidae): Evolutionary and Conservation Implications," *Acta Ichthyologica et Piscatoria* 54 (2024): 235–241.

13. P. S. De Alwis, S. Kundu, F. Z. Gietbong, et al., "Mitochondriomics of *Clarias* Fishes (Siluriformes: Clariidae) With a New Assembly of *Clarias camerunensis*: Insights Into the Genetic Characterization and Diversification," *Life* 13, no. 2 (2023): 482.

14. G. G. Sudarto and L. Pouyaud, "Description of a New Clarii Catfish, *Clarias pseudonieuhofii* From West Borneo (Siluriformes: Clariidae)," *Zoological Studies* 43, no. 1 (2004): 8–19.

15. A. Supikamolseni, N. Ngaoburanawit, M. Sumontha, et al., "Molecular Barcoding of Venomous Snakes and Species-Specific Multiplex PCR Assay to Identify Snake Groups for Which Antivenom is Available in Thailand," *Genetics and Molecular Research* 14, no. 4 (2015): 13981–13997.

16. S. Chen, Y. Zhou, Y. Chen, and J. Gu, "Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor," *Bioinformatics* 34, no. 17 (2018): i884–i890.

17. J. J. Jin, W. B. Yu, J. B. Yang, et al., "GetOrganelle: A Fast and Versatile Toolkit for Accurate de Novo Assembly of Organelle Genomes," *Genome Biology* 21, no. 1 (2020): 241.

18. M. Bernt, A. Donath, F. Jühling, et al., "MITOS: Improved de Novo Metazoan Mitochondrial Genome Annotation," *Molecular Phylogenetics and Evolution* 69, no. 2 (2013): 313–319.

19. D. Laslett and B. Canback, "ARAGORN, a Program to Detect tRNA Genes and tmRNA Genes in Nucleotide Sequences," *Nucleic Acids Research* 32, no. 1 (2004): 11–16.

20. P. Kerpedjiev, S. Hammer, and I. L. Hofacker, "Forna (Force-Directed RNA): Simple and Effective Online RNA Secondary Structure Diagrams," *Bioinformatics* 31, no. 20 (2015): 3377–3379.

21. S. Greiner, P. Lehwark, and R. Bock, "OrganellarGenomeDRAW (OGDRAW) Version 1.3.1: Expanded Toolkit for the Graphical Visualization of Organellar Genomes," *Nucleic Acids Research* 47, no. W1 (2019): W59–W64.

22. Y. Kumazawa and H. Endo, "Mitochondrial Genome of the Komodo Dragon: Efficient Sequencing Method With Reptile-Oriented Primers and Novel Gene Rearrangements," *DNA Research* 11, no. 2 (2004): 115–125.

23. S. Kumar, G. Stecher, M. Suleski, M. Sanderford, S. Sharma, and K. Tamura, "MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing," *Molecular Biology and Evolution* 41, no. 12 (2024): 263.

24. N. T. Perna and T. D. Kocher, "Patterns of Nucleotide Composition at Fourfold Degenerate Sites of Animal Mitochondrial Genomes," *Journal of Molecular Evolution* 41, no. 3 (1995): 353–358.

25. J. J. Faith and D. D. Pollock, "Likelihood Analysis of Asymmetrical Mutation Bias Gradients in Vertebrate Mitochondrial Genomes," *Genetics* 165, no. 2 (2003): 735–745.

26. H. Zhang and B. Zi, "Cubar: Codon Usage Bias Analysis; R Package Version 0.4.0," 2023, https://CRAN.R-project.org/package=cubar.

27. X. Xia and S. Kumar, "DAMBE7: New and Improved Tools for Data Analysis in Molecular Biology and Evolution," *Molecular Biology and Evolution* 35, no. 6 (2018): 1550–1552.

28. N. Goldman and Z. A. Yang, "Codon-Based Model of Nucleotide Substitution for Protein-Coding DNA Sequences," *Molecular Biology and Evolution* 11 (1994): 725–736.

29. N. Sueoka, "Directional Mutation Pressure and Neutral Molecular Evolution," *Proceedings of the National Academy of Sciences* 85, no. 8 (1988): 2653–2657.

30. F. Wright, "The Effective Number of Codons Used in a Gene," *Gene* 87, no. 1 (1990): 23–29.

31. M. Gouy and C. Gautier, "Codon Usage in Bacteria: Correlation With Gene Expressivity," *Nucleic Acids Research* 10, no. 22 (1982): 7055–7074.

32. M. Nei and S. Kumar, *Molecular Evolution and Phylogenetics* (Oxford University Press, 2000).

33. J. Rozas, A. Ferrer-Mata, J. C. Sánchez-DelBarrio, et al., "DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets," *Molecular Biology and Evolution* 34, no. 12 (2017): 3299–3302.

34. K. Katoh and D. M. Standley, "MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability," *Molecular Biology and Evolution* 30, no. 4 (2013): 772–780.

35. S. Kalyaanamoorthy, B. Q. Minh, T. K. Wong, A. Von Haeseler, and L. S. Jermiin, "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates," *Nature Methods* 14, no. 6 (2017): 587–589.

36. B. Q. Minh, M. A. T. Nguyen, and A. Von Haeseler, "Ultrafast Approximation for Phylogenetic Bootstrap," *Molecular Biology and Evolution* 30, no. 5 (2013): 1188–1195.

37. L. T. Nguyen, H. A. Schmidt, A. Von Haeseler, and B. Q. Minh, "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies," *Molecular Biology and Evolution* 32, no. 1 (2015): 268–274.

38. J. P. Huelsenbeck and F. Ronquist, "MRBAYES: Bayesian Inference of Phylogenetic Trees," *Bioinformatics* 17, no. 8 (2001): 754-755.

39. R. Bouckaert, J. Heled, D. Kühnert, et al., "BEAST 2: A Software Platform for Bayesian Evolutionary Analysis," *PLoS Computational Biology* 10, no. 4 (2014): e1003537.

40. I. Kappas, S. Vittas, C. N. Pantzartzi, E. Drosopoulou, and Z. G. Scouras, "A Time-Calibrated Mitogenome Phylogeny of Catfish (Teleostei: Siluriformes)," *PLoS ONE* 11, no. 12 (2016): e0166988.

41. M. Gayet, M. Jégu, J. Bocquentin, and F. R. Negri, "New Characoids From the Upper Cretaceous and Paleocene of Bolivia and the Mio-Pliocene of Brazil: Phylogenetic Position and Paleobiogeographic Implications," *Journal of Vertebrate Paleontology* 23, no. 1 (2003): 28–46.

42. M. Hasegawa, H. Kishino, and T. A. Yano, "Dating of the Human–Ape Splitting by a Molecular Clock of Mitochondrial DNA," *Journal of Molecular Evolution* 22, no. 2 (1985): 160–174.

43. A. J. Drummond, M. A. Suchard, D. Xie, and A. Rambaut, "Bayesian Phylogenetics With BEAUti and the BEAST. 1.7," *Molecular Biology and Evolution* 29, no. 8 (2012): 1969–1973.

44. P. Chalermwong, T. Panthum, P. Wattanadilokcahtkun, et al., "Overcoming Taxonomic Challenges in DNA Barcoding for Improvement of Identification and Preservation of Clariid Catfish Species," *Genomics & Informatics* 21, no. 3 (2023): e39.

45. D. X. Zhang and G. M. Hewitt, "Nuclear Integrations: Challenges for Mitochondrial DNA Markers," *Trends in Ecology & Evolution* 11, no. 6 (1996): 247–251.

46. A. Srivathsan and R. Meier, "On the Inappropriate Use of Kimura-2-Parameter (K2P) Divergences in the DNA-Barcoding Literature," *Cladistics* 28, no. 2 (2012): 190–194.

47. P. D. Hebert, A. Cywinska, S. L. Ball, and J. R. DeWaard, "Biological Identifications Through DNA Barcodes," *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, no. 1512 (2003): 313–321.

48. E. Paradis, J. Claude, and K. Strimmer, "APE: Analyses of Phylogenetics and Evolution in R Language," *Bioinformatics* 20, no. 2 (2004): 289-290.

49. S. D. Brown, R. A. Collins, S. Boyer, et al., "Spider: An R Package for the Analysis of Species Identity and Evolution, With Particular Reference to DNA Barcoding," *Molecular Ecology Resources* 12, no. 3 (2012): 562–565.

50. R Core Team, "R: A Language and Environment for Statistical Computing," (R Foundation for Statistical Computing2025, https://www.R-project.org/.

51. E. Ostertagova, O. Ostertag, and J. Kováč, "Methodology and Application of the Kruskal–Wallis Test," *Applied Mechanics and Materials* 611 (2014): 115–120.

52. A. Dinno, "Nonparametric Pairwise Multiple Comparisons in Independent Groups Using Dunn's Test," *The Stata Journal: Promoting Communications on Statistics and Stata* 15, no. 1 (2015): 292–300.

53. J. Xie, Y. Chen, G. Cai, R. Cai, Z. Hu, and H. Wang, "Tree Visualization by One Table (tvBOT): A Web Application for Visualizing, Modifying and Annotating Phylogenetic Trees," *Nucleic Acids Research* 51, no. W1 (2023): W587–W592.

54. I. Letunic and P. Bork, "Interactive Tree of Life (iTOL) v6: Recent Updates to the Phylogenetic Tree Display and Annotation Tool," *Nucleic Acids Research* 52, no. W1 (2024): W78–W82.

55. J. Zhang, P. Kapli, P. Pavlidis, and A. Stamatakis, "A General Species Delimitation Method With Applications to Phylogenetic Placements," *Bioinformatics* 29, no. 22 (2013): 2869–2876.

56. P. Kapli, S. Lutteropp, J. Zhang, et al., "Multi-Rate Poisson Tree Processes for Single-Locus Species Delimitation Under Maximum Likelihood and Markov Chain Monte Carlo," *Bioinformatics* 33, no. 11 (2017): 1630–1638.

57. A. Stamatakis, "RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies," *Bioinformatics* 30, no. 9 (2014): 1312-1313.

58. A. Morio, R. Tsutsumi, S. Satomi, et al., "Leucine Imparts Cardioprotective Effects by Enhancing mTOR Activity and Mitochondrial Fusion in a Myocardial Ischemia/Reperfusion Injury Murine Model," *Diabetology & Metabolic Syndrome* 13, no. 1 (2021): 139.

59. Y. Zhang, Z. Shen, X. Meng, et al., "Codon Usage Patterns Across Seven Rosales Species," *BMC Plant Biology* 22, no. 1 (2022): 65.

60. M. N. Islam and S. Sultana, "Codon Usage Bias and Purifying Selection Identified in *Cirrhinus reba* Mitogenome," *Journal of Advanced Biotechnology and Experimental Therapeutics* 5, no. 3 (2022): 605–614.

61. M. A. Siddika, K. A. Ahmed, M. S. Alam, J. Bushra, and R. A. Begum, "Complete Mitogenome and Intra-Family Comparative Mitogenomics Showed Distinct Position of Pama Croaker *Otolithoides pama*," *Scientific Reports* 14, no. 1 (2024): 13820.

62. J. Li, G. Fan, and Y. He, "Predicting the Current and Future Distribution of Three *Coptis* Herbs in China Under Climate Change Conditions Using the MaxEnt Model and Chemical Analysis," *Science of the Total Environment* 698 (2020): 134141.

63. L. Zhang, Guo LiF, Q. Duan, Y. Wang, W. Yang, and Y. etal, "Different Proportions of Branched-Chain Amino Acids Modulate Lipid Metabolism in a Finishing Pig Model," *Journal of Agricultural and Food Chemistry* 69, no. 25 (2021): 7037–7048.

64. J. Sonongbua, T. Thong, T. Panthum, et al., "Insights Into Mitochondrial Rearrangements and Selection in Accipitrid Mitogenomes, With New Data on *Haliastur indus* and *Accipiter badius Poliopsis*," *Genes* 15, no. 11 (2024): 1439.

65. O. Rota-Stabelli, N. Lartillot, H. Philippe, and D. Pisani, "Serine Codon-Usage Bias in Deep Phylogenomics: Pancrustacean Relationships as a Case Study," *Systematic Biology* 62, no. 1 (2013): 121–133.

66. H. S. Souza, S. Marsiccobetre, R. O. O. Souza, et al., "Branched-Chain Amino Acids Catabolism as a Source of New Drug Targets in Pathogenic Protists," *Experimental Parasitology* 249 (2023): 108499.

67. Z. Yang and J. P. Bielawski, "Statistical Methods for Detecting Molecular Adaptation," *Trends in Ecology & Evolution* 15, no. 12 (2000): 496–503.

68. M. Yang, Z. Yang, C. Liu, X. Lee, and K. Zhu, "Characterization of the Complete Mitochondrial Genome of the Spotted Catfish *Arius maculatus* (Thunberg, 1792) and Its Phylogenetic Implications," *Genes* 13, no. 11 (2022): 2128.

69. J. Gu, X. Zhou, C. Song, et al., "Codon Usage Bias in Mitochondrial Genomes Across Three Species of *Siphonaria* (Mollusca: Gastropoda)," *Genes* 16, no. 7 (2025): 747.

70. R. D. Ward, R. Hanner, and P. D. Hebert, "The Campaign to DNA Barcode All Fishes (FISH-BOL)," *Journal of Fish Biology* 74, no. 2 (2009): 329–356.

71. A. Tsoupas, S. Papavasileiou, S. Minoudi, et al., "DNA Barcoding Identification of Greek Freshwater Fishes," *PLoS One* 17, no. 1 (2022): e0263118.

72. I. Parvez, R. A. Rumi, P. R. Ray, et al., "Invasion of African *Clarias gariepinus* Drives Genetic Erosion of the Indigenous *C. batrachus* in Bangladesh," *Biology* 11, no. 2 (2022): 252.

73. A. J. Gorog, M. H. Sinaga, and M. D. Engstrom, "Vicariance or Dispersal? Historical Biogeography of Three Sunda Shelf Murine Rodents (*Maxomys surifer*, *Leopoldamys sabanus* and *Maxomys whiteheadi*)," *Biological Journal of the Linnean Society* 81, no. 1 (2004): 91–109.

74. M. E. Steiper, "Population History, Biogeography, and Taxonomy of Orangutans (*Pongo*) Based on a Population Genetic Meta-Analysis of Multiple Loci," *Journal of Human Evolution* 50, no. 5 (2006): 509–522.

75. I. Imron, F. Anggraeni, W. Pamungkas, et al., "Intraspecific Diversity and Phylogeography of Bony Lip Barb, *Osteochilus vittatus*, in Sundaland, as Revealed by Mitochondrial Cytochrome Oxidase I (mtCOI)," *Fisheries and Aquatic Sciences* 27, no. 3 (2023): 145–158.

76. S. Singhal and C. Moritz, "Reproductive Isolation Between Phylogeographic Lineages Scales With Divergence," *Proceedings of the Royal Society B: Biological Sciences* 280, no. 1772 (2013): 20132246.

77. J. A. Leonard, R. J. den Tex, M. T. Hawkins, V. Muñoz-Fuentes, R. Thorington, and J. E. Maldonado, "Phylogeography of Vertebrates on the Sunda Shelf: A Multi-Species Comparison," *Journal of Biogeography* 42, no. 5 (2015): 871–879.

78. L. R. Heaney, "A Synopsis of Climatic and Vegetational Change in Southeast Asia," *Climatic Change* 19, no. 1-2 (1991): 53–61.

79. M. I. Bird, D. Taylor, and C. Hunt, "Palaeoenvironments of Insular Southeast Asia During the Last Glacial Period: A Savanna Corridor in Sundaland?" *Quaternary Science Reviews* 24, no. 20-21 (2005): 2228–2242.

80. A. Sholihah, E. Delrieu-Trottin, F. L. Condamine, et al., "Impact of Pleistocene Eustatic Fluctuations on Evolutionary Dynamics in Southeast Asian Biodiversity Hotspots," *Systematic Biology* 70, no. 5 (2021): 940–960.

81. A. H. Bornbusch, "Phylogenetic Relationships Within the Eurasian Catfish Family Siluridae (Pisces: Siluriformes), With Comments on Generic Validities and Biogeography," *Zoological Journal of the Linnean Society* 115, no. 1 (1995): 1–46.

82. H. K. Voris, "Maps of Pleistocene Sea Levels in Southeast Asia: Shorelines, River Systems and Time Durations," *Journal of Biogeography* 27, no. 5 (2000): 1153–1167.

83. B. W. Low, H. H. Ng, and H. H. Tan, "Two New Species of Blackwater Catfishes (Siluriformes: Siluridae and Clariidae) From the Natuna Archipelago, Indonesia," *The Raffles Bulletin of Zoology* 70 (2022): 1–15.

84. Z. Cheng, J. Wu, C. Luo, et al., "Southeast Asian Rainforest Lost Biodiversity During the Range Expansion to Ice-Age Sunda Shelf," *Global and Planetary Change* 242 (2024): 104597.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Supporting Information 1**. Figure S1. Photograph of the *Clarias nieuhofii* specimen from Phatthalung Province, Thailand. (a) Lateral view, (b) dorsal view, and (c) head and barbels. Scale bar represents 5 cm. Figure S2. The circular mitochondrial genome map of *Clarias nieuhofii* (Accession Number PX395444). Genes encoded on the H-strand are shown outside the circle, whereas those on the L-strand are shown inside. Different gene types are represented as colored boxes, and the gray inner circles indicate the GC content of the mitogenome. Figure S3. The circular mitochondrial genome map of *Clarias nieuhofii* (Accession Number PX395445). Genes encoded on the H-strand are shown outside the circle, whereas those on the L-strand are shown inside. Different gene types are represented as colored boxes, and the gray inner circles indicate the GC content of the mitogenome. Figure S4. The circular mitochondrial genome map of *Clarias nieuhofii* (Accession Number PX395446). Genes encoded on the H-strand are shown outside the circle, whereas those on the L-strand are shown inside. Different gene types are represented as colored boxes, and the gray inner circles indicate the GC content of the mitogenome. Figure S5. The circular mitochondrial genome map of *Clarias nieuhofii* (Accession Number PX395447). Genes encoded on the H-strand are shown outside the circle, whereas those on the L-strand are shown inside. Different gene types are represented as colored boxes, and the gray inner circles indicate the GC content of the mitogenome. Figure S6. Putative secondary structure of the 22 tRNA genes identified in the mitogenome of *Clarias nieuhofii*. Figure S7. AT-skew versus AT % and GC-skew versus GC% in *Clarias* species on protein-coding genes. The *x*-axis indicates the skew values and the *y*-axis provides the AT or GC percentages. Figure S9. Comparison of AT-skew and GC-skew across 13 protein-coding genes (PCGs) of *Clarias* species. The *x*-axis represents individual PCGs and the *y*-axis shows the skew values (positive or negative) for each gene. Figure S9. Structural organization of the mitochondrial CR (D-loop) in *Clarias nieuhofii* from Thailand. The alignment highlights conserved sequence motifs and functional domains, including the extended termination-associated sequence (ETAS), the central conserved domain (CD), and CSBs (CSB-1, CSB-2, and CSB-3). The sequence logo at the top represents nucleotide conservation, with letter height proportional to information content (bits). Figure S10. Relative synonymous codon usage (RSCU) of the 13 mitochondrial protein-coding genes in *Clarias nieuhofii*. (a) RSCU values across codon families, with codon families shown on the *x*-axis and RSCU values on the *y*-axis. (b) Heatmap of RSCU patterns across the mitochondrial protein-coding genes. Figure S11. Relative synonymous codon usage (RSCU) of the 13 mitochondrial protein-coding genes in *Clarias* species. (a) RSCU values across codon families, with codon families shown on the *x*-axis and RSCU values on the *y*-axis. (b) Heatmap of RSCU patterns across mitochondrial protein-coding genes. Figure S12. Codon usage analysis of the 13 mitochondrial protein-coding genes in *Clarias* species. (a) Neutrality plot, (b) effective number of codons (ENc) plot, and (c) boxplot of ENc values for each protein-coding gene across *Clarias* species. Figure S13. Nonsynonymous to synonymous substitution ratios (Ka/Ks) across mitochondrial protein-coding genes in *Clarias* species. Bars indicate Ka/Ks values for each gene. Asterisks (***) denote significant differences ($p < 0.05$), while ns indicates nonsignificance. Figure S14. Barcoding gap analysis for COI and Cytb sequences of *Clarias* species. Panels (A, B) show the distribution of intraspecific divergence (A) and nearest neighbor distances (B) for each gene. Panels (C–F) illustrate scatterplots comparing intraspecific and interspecific distances to nearest neighbor distances: minimum intraspecific vs. nearest neighbor (C), maximum intraspecific vs. nearest neighbor (D), minimum interspecific vs. nearest neighbor (E), and maximum interspecific vs. nearest neighbor (F). Panel (G) presents the distribution of the barcoding gap (nearest neighbor minus maximum intraspecific distance) for each gene. Red dashed lines indicate threshold values for species delimitation (2% for divergence and 0% for barcoding gap). Figure S15. Molecular phylogeny of *Clarias* species inferred from cytochrome c oxidase subunit I (*COI*) sequences using Bayesian inference, showing monophyletic status of each lineage. Figure S16. Molecular phylogeny of *Clarias* species inferred from cytochrome b (*Cytb*) sequences using Bayesian inference, showing monophyletic status of each lineage.

Supporting Information 2. Table S1. Gene organization of the mitochondrial genome of slender walking catfish (*Clarias nieuhofii*). Table S2. Nucleotide composition and skewness of protein-coding genes (PCGs) in *Clarias* species. Table S3. Nucleotide composition and skewness of 13 protein-coding genes (PCGs) in *Clarias* species. Table S4. Codon usage bias indices (GC12, GC3, and effective number of codons) of the protein-coding genes (PCGs) in *Clarias* species. Table S5. P2 index values of mitochondrial protein-coding genes (PCGs) in *Clarias nieuhofii*. Table S6. Classification and accession numbers of species used in the sequence analysis of genetic diversity and nonsynonymous/synonymous (Ka/Ks) values for 13 protein-coding genes across the *Clarias* mitogenome. Table S7. Genetic diversity and nonsynonymous/synonymous (Ka/Ks) values for 13 protein-coding genes (PCGs) across the *Clarias* mitogenome. Table S8. Summary of intraspecific and interspecific genetic distances and barcoding gaps for each marker in *Clarias* species.