Mitophylogeny of Pangasiid Catfishes and its Taxonomic Implications for Pangasiidae and the Suborder Siluroidei

Thuy Yen Duong[1], Linh Thi Khanh Pham[2,3], Xuyen Thi Kim Le[2,3], Ngoc Tran Thi Nguyen[1], Siti Azizah Mohd Nor[4], and Thanh Hoa Le[2,3,*]

[1]*College of Aquaculture and Fisheries, Can Tho University, 3/2 street, Can Tho City, Vietnam. E-mail: thuyyen@ctu.edu.vn (Duong)*

[2]*Immunology Department, Institute of Biotechnology (IBT), Vietnam Academy of Science and Technology (VAST). 18. Hoang Quoc Viet Rd., Cau Giay, Hanoi, Vietnam. *Correspondence: E-mail: imibtvn@gmail.com (Le)*
E-mail: ntntran@ctu.edu.vn (Pham); lethikimxuyen71@gmail.com (Kim Le)

[3]*Graduate University of Science and Technology (GUST), Vietnam Academy of Science and Technology (VAST), 18. Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam*

[4]*Institute of Marine Biotechnology, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia. E-mail: s.azizah@umt.edu.my (Nor)*

Pangasiidae (catfish order: Siluriformes) comprises 30 valid catfish species in four genera: *Pangasius*, *Pangasianodon*, *Helicophagus*, and *Pseudolais*. Their systematics is frequently revised due to the addition of newly described species. Although Pangasiidae is known to be a monophyletic family, the generic and phylogenetic relationships among the taxa are poorly resolved. This study characterized three newly obtained complete mitogenomes of Mekong River catfishes from Vietnam (*Pangasius mekongensis*, *Pangasius krempfi*, and *Pangasianodon hypophthalmus)*, as well as the inter- and intrafamilial relationships of the Pangasiidae and catfish families in Siluroidei. The genomic features of their mitogenomes were similar to those of previously reported pangasiids, including all regulatory elements, extended terminal associated sequences (ETAS), and conserved sequence blocks (CSBs) (CSB-1, CSB-2, CSB-3, and CSBs, A to F) in the control region. A comprehensive phylogeny constructed from datasets of multiple 13 PCG sequences from 117 complete mitogenomes of 32 recognized siluriform families established Pangasiidae as monophyletic and a sister group of Austroglanididae. The [Pangasiidae + Austroglanididae] + (Ictaluridae + Cranoglanididae) + Ariidae] clade is a sister to the "Big Africa" major clade of Siluriformes. Furthermore, both phylogenies constructed from the single barcodes (83 partial *cox*1 and 80 partial *cyt*B, respectively) clearly indicate

genus relationships within Pangasiidae. *Pangasianodon* was monophyletic and a sister to the (*Pangasius* + *Helicophagus* + *Pseudolais*) group. Within the genus *Pangasius*, *P. mekongensis* was placed as a sister taxon to *P. pangasius*. *Pangasius sanitwongsei* was found to be related to and grouped with *Pangasianodon*, but in single-gene phylogenies, it was assigned to the *Pangasius* + *Helicophagus* + *Pseudolais* group. The datasets in this study are useful for studying pangasiid systematics, taxonomy and evolution.

Key words: *Pangasius*, *Pangasianodon*, Pangasiidae, Mitogenomes, Mitophylogenetic analysis

BACKGROUND

Siluroidei (Teleostei: Siluriformes), one of the three suborders of Siluriformes (Siluroidei, Loricarioidei, and Diplomystoidei), comprises catfishes throughout the world that have not been exhaustively phylogenetically classified (Diogo et al. 2004; Betancur et al. 2017; Fricke et al. 2023; Schedel et al. 2022). Improvements in phylogeny testing using molecular markers and the increasing availability of mitogenome data have indicated the Pangaean origin of the siluriform catfishes (Hardman, 2005; Sullivan et al. 2006; Nakatani et al. 2011; Kappas et al. 2016; Moreira et al. 2017; Schedel et al. 2022). Interfamilial relationships within Siluroidei are divided into two major clades, "Big Africa" and "Big Asia", first reported by Sullivan et al. (2006) based on the analysis of the *rag1* and *rag2* nuclear gene sequences. These big groups have been clarified by mitophylogeny (Kappas et al., 2016) and, more recently, by the additional mitogenomic data of mochokids (Mochokidae) and austroglanidids (Austroglanididae) (Schedel et al. 2022). Furthermore, when phylomitogenomic data from a number of loricarioid species were analyzed, the sub-order position of Siluroidei, Diplomystoidei, and Loricarioidei in the Siluriformes was recognized (Moreira et al. 2017). However, in Siluroidei, the intra and interrelationships of the pangasiid species and the monophyly of the family Pangasiidae remain uncertain. Schedel et al. (2022) recovered the (Astroglanididae + Pangasiidae) interfamilial group as a sister to a clade encompassing Ictaluridae and Cranoglanididae. The tentative "Big Africa" membership status of Pangasiidae requires additional support. Morphology data have led to a division of the family Pangasiidae Bleeker, 1858, comprising 30 valid species, into four recognized genera: *Pangasius* Valenciennes, 1840; *Pangasianodon* Chevey, 1931; *Helicophagus*

Bleeker, 1858; and *Pseudolais* Vaillant, 1902 (https://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp) (Fricke et al. 2023). The genus *Pangasius* possesses 23 species, the highest number of species in Pangasiidae, while the genus *Pangasianodon* includes only two species, the genus *Helicophagus* three species and *Pseudolais* two species, respectively (Froese and Pauly 2021; Fricke et al. 2023). Catfishes of the family Pangasiidae are widely distributed in river systems throughout South Asia (Pakistan and India), Southeast Asia (Myanmar, Malaysia, Thailand, Cambodia, Laos, Vietnam, and Indonesia), the southern part of China, and recently in a river in South Africa (Mäkinen et al. 2013; Wei et al. 2020; Fricke et al. 2023). In general, pangasiids are primarily restricted to freshwater, but a few species, including *Pangasius mekongensis* Gustiano, Teugels & Pouyaud, 2003 and *Pangasius krempfi* Fang & Chaux, 1949, spend part of their life cycle in brackish water in the Lower Mekong Basin (Rainboth 1996; Poulsen et al. 2004; Hogan et al. 2007). Some large-sized pangasiid catfish species, particularly migratory species inhabiting the Mekong River system, have been considered vulnerable or critically endangered in the wild. Among these threatened species are *Pangasianodon hypophthalmus* Sauvage, 1878; *Pangasius mekongensis*; and *Pangasius krempfi* (Vidthayanon 2012; Vidthayanon and Hogan 2011; Baird 2011).

In addition to their ecological importance, *Pangasius* and *Pangasianodon* catfish have become economically important aquaculture species in Vietnam, Thailand, Indonesia, and India (Phan et al. 2009). *Pangasianodon hypophthalmus* (*Pn. hypophthalmus*) is the most widely cultured pangasiid species, accounting for the eighth-highest freshwater finfish production in the world (FAO 2022). In fish taxonomy, this species was formerly listed in the genus *Pangasius* (Roberts and Vidthayanon 1991; Pouyaud and Teugels 2000; Zhao et al. 2014). However, it was later revised to *Pangasianodon* based on the unique character of 8 to 9 pelvic fin rays, compared to 6 pelvic fin rays in other genera of the same family (Rainboth 1996; Gustiano 2009; Kottelat 2013). *Pangasius krempfi* and *Pangasius mekongensis* are important for capture fisheries in the Lower Mekong River basin (Poulsen et al. 2004; Hogan et al. 2007). Both species exhibit anadromous behavior (Hogan et al. 2007; Vu et al. 2022) and similar morphological characteristics, particularly at early stages (Karinthanyakit and Jondeung 2012; Tran and Duong 2019; NAGAO 2021). These characteristics make these two species suitable subjects for studying the congeneric relationships of the genus *Pangasius* and their phylogenetic relationships with *Pangasianodon* species in the Siluroidei suborder. There are still some difficulties surrounding the taxonomic classification and phylogenetic relationships of several species in the genera *Pangasius* and *Pangasianodon*. Furthermore, the systematics of Pangasiidae is frequently revised due to the addition of newly described species (Gustiano et al. 2003; 2021; Mohindra et al. 2015; Dwivedi et al. 2017; Fricke et al. 2023).

Mitochondrial markers have been tested to infer the mitophylogenetic interfamilial relationship in the suborder Siluroidei, which involves the placement of Pangasiidae. Many previous studies have investigated the phylogenetic relationships of the familial and multifamilial clades of Siluriformes and reported the intra- and inter-continental diversification within catfishes. These analyses comprised partial or complete mitochondrial DNA (mtDNA) sequences, which were used to construct the phylogeny of Siluriformes (Sullivan et al., 2006; Hardman 2005; Karinthanyakit and Jondeung, 2012; Miya and Nishida 2015; Kappas et al. 2016; Moreira et al. 2017; Kim et al. 2018; Quyen et al. 2018; Zhang et al. 2021; Schedel et al. 2022). The partial or complete single-gene sequence and phylogenetic analyses included cytochrome *b* (*cyt*B, approximately 1,150 bp) (Hardman 2005; Karinthanyakit and Jondeung 2012), cytochrome oxidase subunit 1 (*cox*1, 551 bp or 1,551 bp) (Quyen et al., 2018; Zhang et al. 2021), and a growing number of works using the complete mitogenomes (Jondeung et al. 2007; Miya and Nishida 2015; Kappas et al. 2016; Moreira et al. 2017; Kim et al. 2018; Schedel et al. 2022). Although partial mtDNA sequences effectively resolve relationships among catfish taxa (Karinthanyakit and Jondeung 2012; Hardman 2005; Kartavtsev et al. 2007; Nakatani et al. 2011), full-length mtDNA sequences provide a higher level of resolution. In some studies, the complete mitogenomes were used to clarify that clarifies the intergeneric, interfamilial, and intersuborder phylogenetic relationships related to Siluroidei, Diplomystoidei, and Loricarioidei in Siluriformes (Miya and Nishida 2015; Kappas et al. 2016; Villela et al. 2017; Kim et al. 2018; Schedel et al. 2022).

Representatives of pangasiids, however, were still placed in a non-stable phylogenetic relationship within the family Pangasiidae and between the associated siluroid families in Siluroidei (Jondeung et al. 2007; Nakatani et al., 2011; Kappas et al. 2016; Kim et al. 2018; Wei et al. 2020). Concerning genus-level relationships in Pangasiidae, based on a complete *cytochrome b* analysis of 13 pangasiids and six schilbids in Thailand, Karinthanyakit and Jondeung (2012) determined that the four genera have *Pangasianodon* as a basal group and *Pseudolais* and *Helicophagus* as sister groups of *Pangasius*. This conclusion was aligned with findings by Quyen et al. (2018), based on partial *cox*1 and 16S sequences of 30 species belonging to nine families reported distributed in the Lower Mekong Basin, including 14 species of Pangasiidae. In both studies, *Pangasius pangasius* was absent from the analysis. However, in a phylogenetic analysis using the partial *cyt*B by Tran et al. (2017), this species showed a sister taxon relationship with *P. mekongensis*. Another Indian *Pangasius silasi* was found to be a sister species to *P. pangasius* (Dwivedi et al. 2017), but no molecular data were available from this newly described species in GenBank for further investigation.

Insufficient numbers of complete mtDNA sequences of Pangasiidae have been used for phylogenetic analyses of pangasiids and siluriforms, therefore awaiting further work (Jondeung et al. 2007; Miya and Nishida 2015; Kappas et al. 2016). Early molecular phylogenetic studies by Jondeung et al. (2007) and Nakatani et al. (2011) used two complete mtDNA sequences of *Pangasianodon gigas*

(AY762971) and *Pangasius larnaudii* (AP012018), respectively. In later taxonomic assessments, a few more complete mitogenomes of pangasiids were used. For example, three sequences of *Pn. gigas*, (AY762971), *P. larnaudii* (AP012018), and *Pn. hypophthalmus* (NC_021752) were reported in Kappas et al. (2016), and two *Pangasius* and three *Pangasianodon* mtDNA sequences were used in Kim et al. (2018). In recent studies, Wei et al. (2020) employed *Pn. hypophthalmus* and three *Pangasius* species, including *P. sanitwongsei* (MN809630), while Villela et al. (2017) and Schedel et al. (2022) included *Pangasius pangasius* in their phylogenetic analyses. More coverage of complete mitogenomic sequences from each species, as well as more comprehensive sequences and phylogenetic analyses within and between *Pangasius* and *Pangasianodon*, is required.

This study aimed to reconstruct a mitophylogeny of pangasiids and all representative siluriform species in order to resolve Pangasiidae's placement within Siluriformes, and clarify inter- and intrageneric relationships within the monophyletic Pangasiidae among the key families within the suborder Siluroidei. To complete these analyses, we used the available siluriform mitogenomic data and expanded the topology presentation of the siluriform mitophylogeny, combined with our recently sequenced mitosequences from the three Mekong River catfishes (i.e., *Pangasius mekongensis*, *Pangasius krempfi*, and *Pangasianodon hypophthalmus*). Furthermore, we used the available partial *cox*1 and *cyt*B datasets, plus newly obtained sequence data, to reconstruct phylogenies showing the most comprehensive interspecific and intergeneric relationships of the pangasiid species.

MATERIALS AND METHODS

Sample collection and species identification

Samples were obtained from wild-caught adults from the Mekong River in Vietnam, including *Pangasius mekongensis* (*P. mekongensis*) from Vam Nao River in An Giang province (9°53'30.2"N, 105°57'39.3"E); *Pangasianodon hypophthalmus* (*Pn. hypophthalmus*) from Hau Giang River in Can Tho City (10.0452°N, 105.7469°E); and *Pangasius krempfi* (*P. krempfi*) from Phong Nam in Soc Trang province (10°32'16.79"N, 105°19'22.20"E). After the morphological examination, a piece (50 mg) of muscle was excised from the fish and individually stored at −20°C until use. Total genomic DNA was extracted from each fish sample according to the manufacturer's instructions using the GeneJET™ Genomic DNA Purification Kit (Thermo Scientific Inc., MA, USA). DNA extract was eluted in 100 μL and stored at −20 °C until use.

Subsequently, species' identities were confirmed by molecular phylogenetic analysis using mitochondrial DNA sequences and compared with the sequences available in GenBank, *e.g.*, for *P.*

mekongensis (*cyt*B: KY451465; KY451466; KY451467; *cox*1: KT289880), for *Pn. hypophthalmus* (*cyt*B: GQ856796/KC846907; *cox*1: MK216612/MK216603), and for *P. krempfi* (*12S-16S*: HM355773/MG076881; *cyt*B: MN087451/ MN087461/ MN087471/ HM236386/ HM236390) (Karinthanyakit and Jondeung 2012; Tran and Duong 2019; Zhao et al. 2014; Kim et al. 2018; Quyen et al. 2018).

The annotation and sequence analysis of the pangasiid mitogenomes

Long-range PCRs (L-PCRs) were applied using commercial kits (Thermo Fisher Scientific Inc., Waltham, MA, USA) and sequenced using primers listed in **table S1**. The complete mitogenome was obtained after assembling all the sequences from the sequencing. Protein-coding genes (PCGs) were identified by alignment with available mitogenomes of other *Pangasius/Pangasianodon* catfish species, with ATG/GTG as start and TAA/TAG as stop codons. For some genes, T-- or TA-incomplete stops were considered. PCGs were translated using the *vertebrate mitochondrial genetic code* (translation Table 2 in GenBank). The tRNAscan-SE 1.21 program (www.genetics.wustl.edu/eddy/tRNAscan-SE/) (Lowe and Chan 2016) and ARWEN (http://mbio-serv2.mbioekol.lu.se/ARWEN/) were used to identify transfer RNA genes (Laslett and Canback 2008). The mitoribosomal genes (MRG), *i.e.*, 16S (*rrn*L) and 12S (*rrn*S) RNA genes, were recognized by the data inference from the previous publication by Jondeung et al. (2007). The circular map was generated by the MitoAnnotator at: http://mitofish.aori.u-tokyo.ac.jp/annotation/input.html (Iwasaki et al. 2013).

The stem-loop secondary structure of the O_L origin site in each mitogenome was identified based on the prediction from RNAfold with minimum free energy (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). The extended terminal associated sequences (ETAS), central conserved sequence blocks (CSB-F, CSB-E, CSB-D, CSB-C, CSB-B, CSB-A), conserved sequence blocks (CSB-1, CSB-2, CSB-3), and putative promoters of the mtDNA control region were identified by alignment of the conserved domains in teleost fishes (Guo et al. 2003; Li et al. 2012; Fischer et al. 2013; Villela et al. 2017).

The MEGA X program (Kumar et al. 2018) was used to determine the mitogenomic characteristics of pangasiid species. These included the nucleotide composition, AT and GC content of the complete mtDNAs and 13 PCGs, and the pairwise nucleotide comparison (%) of each PCG and each MRG among *P. mekongensis*, *Pn. hypophthalmus*, *P. krempfi*, *Pn. gigas* and *P. larnaudii*. Genetic distances were inferred by the analysis of 15,566–15,576 coding nucleotide sequences. The AT and GC skewness values (ranging from −1 to +1) were calculated using the following formula: AT skew = $(A - T)/(A + T)$ and GC skew = $(G - C)/(G + C)$ (Perna and Kocher 1995).

Mitophylogenetic and synteny analyses

We used 117 complete or near-complete mitogenomes from 109 species of 32 families of Siluriformes for ingroup data for phylogenetic analysis, and two species from Gonorynchiformes as an outgroup (information and author reference for each is given in Table S2), which provided all 13 PCGs for nucleotide sequence alignment. The siluriform families were Pangasiidae ($n = 13$), including one each of the newly obtained mitogenomes from *P. mekongensis*, *Pn. hypophthalmus*, and *P. krempfi*. Other siluriform families were Ailiidae ($n = 1$), Amblycipitidae ($n = 4$), Amphiliidae ($n = 1$), Ariidae ($n = 6$), Aspredinidae ($n = 1$), Astroblepidae ($n = 1$), Auchenipteridae ($n = 3$), Auchenoglanididae (1), Austroglanididae ($n = 3$), Bagridae ($n = 10$), Callichthyidae ($n = 2$), Cetopsidae ($n = 2$), Chacidae ($n = 1$), Clariidae ($n = 5$), Claroteidae ($n = 2$), Cranoglanididae ($n = 1$), Diplomystidae ($n = 1$), Doradidae ($n = 2$), Heteropneustidae ($n = 1$), Horabagridae ($n = 2$), Ictaluridae ($n = 9$), Loricariidae ($n = 9$), Malapteruridae ($n = 1$), Mochokidae ($n = 4$), Pimelodidae ($n = 5$), Plotosidae ($n = 2$), Ritidae ($n = 1$), Schilbeidae ($n = 2$), Siluridae ($n = 11$), Sisoridae ($n = 9$), and Trichomycteridae ($n = 1$). Outgroup species were taken from Gonorynchiformes (*Chanos chanos* and *Gonorynchus greyi*) (Saitoh et al. 2003), as used in a previous report (Kim et al. 2018). The taxon collection was targeted at the taxa/families surrounding the Pangasiidae and the representatives of the catfish clade "Big Africa" and "Big Asia" (Sullivan et al. 2006), including the newly described families (Austroglanididae and Mochokidae) by Schedel et al. (2022). The growing catfish mitodataset enabled us to include representative members of all three major lineages, e.g., Diplomystoidei ($n = 1$), Siluroidei (n = 102), and Loricarioidei (n = 14), for analysis. The nucleotide sequences of the PCGs were concatenated in the order of *ND1*, *ND2*, *COX1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *ND3*, *ND4L*, *ND4*, *ND5*, *ND6*, and *cytB* (Schedel et al. 2022). The complementary ND6 sequence was automatically converted into a sense sequence when downloaded from GenBank. The final block alignment of about 11,410 bp was used for phylogenetic analysis according to the procedure described in Schedel et al. (2022). We used a software package for phylogenetic services, available at https://ngphylogeny.fr/workflows/ (Lemoine et al. 2019) for constructing a maximum likelihood tree. Briefly, the input 119 sequences in fasta format were uploaded for multiple alignment by MAFFT v7.407 (Katoh and Standley 2013), then curated by BMGE v1.12 (Criscuolo and Gribaldo 2010), and inferenced by PhyML v3.3 with Maximum Likelihood phylogeny with 1000 bootstrap replicates (Guindon et al. 2010). The output final tree was rendered and displayed in the Newick v1.6 format (Junier and Zdobnov 2010), and this Newick tree (nwk format) was visualized and parameterized with the FigTree v1.4.4 program (Rambaut 2018).
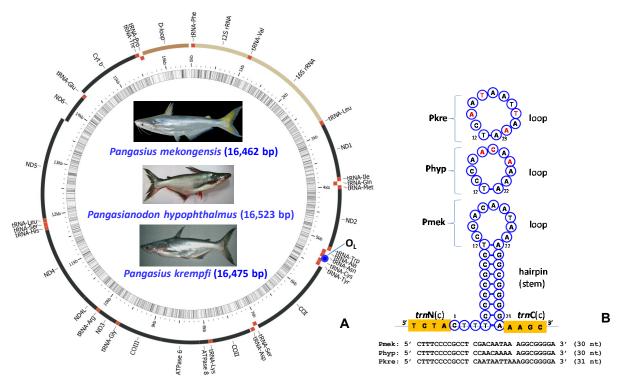
To investigate the taxonomic and generic relationships of the pangasiid taxa, we used the partial *cox*1 and *cyt*B sequences. These sequences were downloaded from GenBank and extracted from the mitogenomes of the newly sequenced pangasiids and the associated families (Cranoglanididae, Schilbeidae, Austroglanididae, and Ictaluridae), and two outgroup sequences were used for each phylogeny (information on mtDNA sequences and country of origin was given in Table S3). Due to the constraints in the composition of marker sequences, we took the outgroup sequences from the distant orders, such as Gymnotiformes for the *cox*1 and Clupeiformes for the *cyt*B analsysis. These supplementary datasets contained 83 partial *cox*1 (551 bp) and 80 partial *cyt*B (634 bp), respectively, for phylogenetic analyses. The phylogenetic tree was reconstructed based on the MAFFT 7.471 alignment by the MEGA X program using the maximum likelihood method with the GTR + I + G model and 1000 bootstrap resamplings (Kumar et al. 2018).

RESULTS

Mitogenome features of the pangasiid species

The complete mitogenome was 16,462 bp in length for *P. mekongensis* (GenBank: MZ272451), 16,523 bp for *Pn. hypophthalmus* (MZ272452), and 16,475 bp for *P. krempfi* (MZ272453). The circular mtDNA comprised 13 PCGs (*atp*8, *atp*6, *cox*1−3, *cyt*B, *nad*1−6, *nad*4L), two MRGs (12S or *rrn*S and 16S or *rrn*L), and 22 tRNAs, similar in gene order and gene length to those of other fish mitogenomes (Fig. 1A) (Jondeung et al. 2007; Zhao et al. 2014; Satoh et al. 2016). Two MRGs, including *12S* (957 bp for *P. mekongensis* and 958 bp for both *Pn. hypophthalmus* and *P. krempfi*), and *16S* (1,677 bp for *P. mekongensis*; 1,674 bp for *Pn. hypophthalmus*; and 1,681 bp for *P. krempfi*) were identified. The twenty-two tRNAs in the mtDNAs ranged from 67 to 75 nucleotides in length, and one tRNA specifying serine (tRNA$^{Ser(GCT)}$, that lacked the DHU-arm (dihydrouridine-arm). All the remaining twenty-one tRNAs had the common '*cloverleaf*' secondary structures with the complete four arms.

Twelve PCGs in *P. mekongensis, Pn. hypophthalmus,* and *P. krempfi* used ATG as a start codon, only *cox*1 used GTG. Nine PCGs used complete TAG or TAA and four (*cox*2, *cox*3, *nad*4, and *cyt*B) used incomplete T-- or TA- codons for gene termination. Short intergenic tracts of 1−5 bp were common, and larger spacers of 14 bp were observed between PCGs.

Fig. 1. A schematic circular map of the mitochondrial genome of three Mekong River-pangasiid catfishes in Vietnam, *Pangasius mekongensis, Pangasianodon hypophthalmus*, and *Pangasius krempfi*, and the O_L origin site of the light (L) strand's replication. A, The circular map and gene abbreviations were generated by the MitoAnnotator software in the MitoFish database (http://mitofish.aori.u-tokyo.ac.jp/annotation/input.html). Protein-coding genes (PCGs) are denoted by two capital letters or full names, and transfer RNA genes (tRNAs) are marked with three-letter amino-acid abbreviations. The heavy (H) strand is indicated by the upper line of the circle and the light (L) strand by the lower line. The D-loop (control region) is located between tRNA^{Pro} and tRNA^{Phe}. The panagsiid photos were taken by the authors from the naturally caught fish on site. B, A schematic presentation of the stem-loop secondary structure of the O_L origin site in mitogenomes of three pangasiid species based on the RNAfold predicted structure with the lowest free energy (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). On the L-strand, between the two flanking tRNAs (*trn*N (c) and *trn*C (c)), there is a conserved stem (hairpin) formed by 9-nucleotide (nt) base-pairing and ending with a loop of 9 nt (in Pmek and Phyp) and 10 nt (in Pkre).

The base composition and skewness values for the mtDNA of seven pangasiid species are listed in table S4. Similar to other Siluroidei and common fishes, for example, in the mtDNA of *P. mekongensis*, the use of A+T was 55.67% and G+C was 44.03%; the AT-skew was 0.096/mtDNA and 0.030/PCGs, and the GC-skew was −0.293/mtDNA and −0.309/PCGs. A similar pattern for base composition and skewness with a minimal difference was observed in other species in the genera *Pangasius* and *Pangasianodon*. Pairwise nucleotide differences (%) for each PCG and MRG among five species, *P. mekongensis*, *P. krempfi*, *P. larnaudii*, *Pn. hypophthalmus*, and *Pn. gigas* are shown in table S5. The lowest sequence divergence (2.44%) was found in *atp*8 between *P. mekongensis* and *P. krempfi*, whereas the highest (17.47%) was in *nad*3 between *Pn. gigas* and *P. krempfi*.

Genetic distances (p-distances) were estimated from the alignment of nucleotide sequences from 11 mitogenomes from eight pangasiid species (using sequences of 15,566 bp–15,571 bp, excluding

tRNA^{Pro} and the CR) (Table S6). The lowest level of divergence was between *P. mekongensis* (Vietnam) and *P. pangasius* (India) at 4.24%–4.31%, and the highest was between *Pangasianodon* species and both *P. krempfi* (at 9.19%–9.34%) and *P. larnaudii* (at 9.25%–9.31%). Within each pangasiid species, pairwise genetic distances were small, only 0.07%–0.34% within *Pn. hypophthalmus* (three mitogenomes available); and 0.29% within *P. pangasius* (two mitogenomes available). Interestingly, *P. sanitwongsei* exhibited a lower distance to *Pn. hypophthalmus* (6.32%–6.45%) than other members of *Pangasius* (7.32%–7.68%), and the distance between *Pn. gigas* and *Pn. hypophthalmus* (7.57%–7.66%) (Table S6).
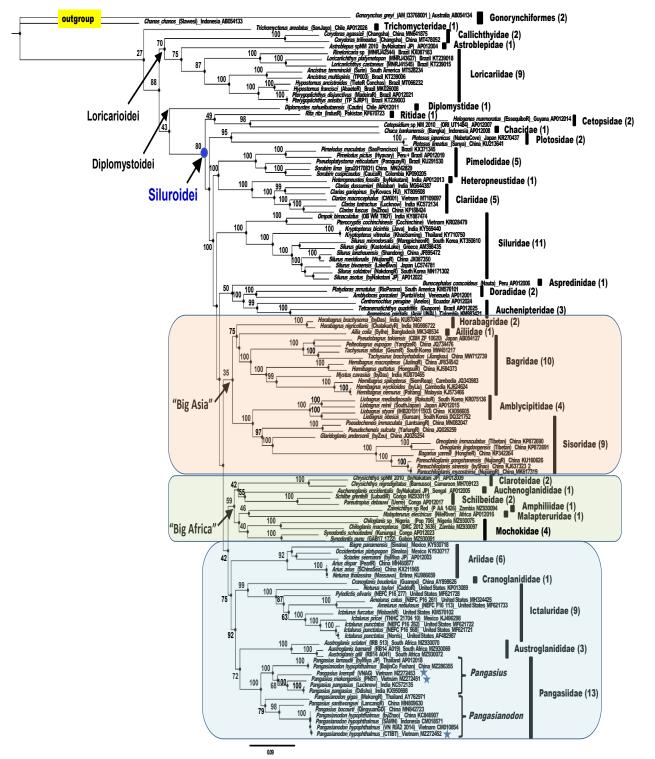
The origin (O$_L$ site) of L-strand replication was a short, non-coding sequence of 30 or 31 nucleotides in the typical tRNA$^{W-A-N-C-Y}$ region (Kartavtsev et al. 2007; Boore 1999). As is the case in all vertebrates, including fish, it has the potential to form a stable stem-loop structure. The conserved stem (hairpin) was rich in G and C and was formed by base-pairing of 9 nucleotides (nt) and ending with an A-rich loop of 9 nt (in *P. mekongensis* and *Pn. hypophthalmus*) and 10 nt (in *P. krempfi*) (Fig. 1B).

The major non-coding region, termed the "control region" (CR), is 823 bp in length for *P. mekongensis*, 887 bp for *Pn. hypophthalmus*, 830 bp for *P. krempfi*, and 827 bp for *P. bocourti*. It is 829 bp for *P. pangasius*, 822 bp for *P. larnaudii*, 887 bp for *P. sanitwongsei*, and 897 bp for *Pn. gigas* (Table S3). The CRs of *Pangasius* were 44–48 nucleotides shorter than those of *Pangasianodon* species. By comparative alignment and sequence analysis of the conserved domains in teleost fishes reported in the previous studies (Guo et al. 2003; Li et al. 2012; Fischer et al. 2013; Villela et al. 2017; Cui et al. 2020), we successfully identified all regulatory elements in the mtDNAs of the pangasiid taxa. These included ETAS, six central conserved sequence blocks (CSB-F, CSB-E, CSB-D, CSB-C, CSB-B, and CSB-A) and three conserved sequence blocks (CSB-1, CSB-2, and CSB-3). The origin site of H-strand replication (O$_H$), a putative promoter, and the TATA boxes in the aligned mtDNA CRs of eight species (*P. mekongensis*, *P. krempfi*, *P. pangasius*, *P. larnaudii*, *P. bocourti*, *P. sanitwongsei*, *Pn*. *hypophthalmus* (3 mitogenomes), and *Pn. gigas*) (Fig. S1). The ETAS is a palindromic sequence of 47 nucleotides, containing TAS (TACAT) and cTAS (reverse complementary ATGTA) motifs. The CSB-2, CSB-3, CSB-A, and CSB-B are the most conserved. Other CSBs share 90%–95% sequence identity among the pangasiids. Interestingly, *P. sanitwongsei* possesses a more similar CR in terms of length (887 bp), sequence composition, and CSB patterns to *Pn. hypophthalmus* than to *Pangasius* species (Fig. S1).

Phylogenetic relationships of Pangasiidae within Siluroidei and Siluriformes

The ML tree (Fig. 2) clearly demonstrated the monophyly of Siluroidei in Silurifomes with 80% boostrap support, distinct from Loricarioidei and Diplomystoidei with 70% and 43% support, respectively. The family Trichomicteridae was not recovered within the Siluroidei. Some siluroid families, such as Ritidae and Ceptopsidae, were shown to be polyphyletically split at the beginning of the descending topology of the tree with a low bootstrap (49%). *Rita rita* catfish, uniquely reported from Pakistan (GenBank: KF670723), was classified into the family Bagridae (Fricke et al. 2023) but was not recovered in this family in the present study. This species represented the separate Ritidae family, recently used by Schedel et al. (2022), which showed its phylogenetic placement far from Bagridae (Fig. 2).

The remaining siluroids were classified into eight multiple family clusters, as follows: i) the Plotosidae and Chacidae with 95% bootstrap; ii) the polyphyletic Pimelodidae, Heteropneustidae, and Clariidae (100% bootstrap); iii) the monophyletic Siluridae (100% bootstrap); iv) the polyphyletic Aspredinidae, Doradidae, and Auchenipteridae with low bootstrap (50%); v) the Bagridae, Horabagridae, and Aillidae (75% bootstrap); vi) the Amblycipitidae and Sisoridae (100% bootstrap); vii) the Claroteidae, Auchenoglanididae, Schilbeidae, Amphillidae, Malapteruridae, and Mochokidae, or also termed as the "Big Africa" major group (59% bootstrap) (Sullivan et al. 2006); and viii) the Pangasiidae, Austroglanididae, Ictaluridae, Cranoglanididae, and Ariidae (72% bootstrap). The "Big Asia" major clade contained the clusters (Amblycipitidae + Sisoridae) and (Bagridae + Horabagridae + Aillidae) (35% bootstrap) (Sullivan et al. 2006), which is positioned as a sister group to the "Big Africa" major clade (Fig. 2). The Pangasiidae and its related Austroglanididae group are monophyletically sistered with the (Ictaluridae and Cranoglanididae) (92% bootstrap), and this major group is placed as a sister to the Ariidae with a moderate boostrap value (75%), according to the topology depicted in figure 2.

Fig. 2. PhyML-phylogeny of the order Siluriformes, including 32 catfish families (117 sequences) of three suborders, Siluroidei, Loricarioidei, and Diplomystoidei based on the complete concatenated nucleotide sequences of all 13 mitochondrial protein coding genes (about 11,408 bp in length) (**Table S2**). Two sequences of Gonorynchiformes were used as outgroup. The alignment was performed by MAFFT (Katoh and Standley 2013), curated by BMGE v1.12 (Criscuolo and Gribaldo 2010), the tree was reconstructed in PhyML 3.3 (Guindon et al. 2010) using a maximum likelihood method and 1000 bootstrap resamplings, and the output Newick tree was extracted and visualized using FigTree v1.4.4 (Rambaut 2018). The nodal bootstrap support values (shown at each node) were interpreted from the concurrently constructed tree using the above MAFFT-BMGE alignment by MEGA X (Kumar et al. 2018). The basal nodes of the three suborders (Diplomystoidei, Loricarioidei, and Siluroidei) as well as the two major "Big Asia" and "Big Africa" groups (background highlighted) are shown by arrows.

The *Pangasius mekongensis, Pangasianodon hypophthalmus*, and *Pangasius krempfi* sequences in this study were indicated by stars and with the associated families background highlighted. The taxa were presented with their full names. The abbreviations of the isolates are given in brackets, including the geographical origin or voucher records of each sequenced specimen (where available), which were retrieved from the previous studies (Saitoh et al. 2003; Nakatani et al. 2011; Kappas et al. 2016; Zhang et al. 2021; Schedel et al. 2022). The country of origin or geographical regions where the sample was reported is given in full name, if available. Accession numbers are given at the end of each sequence label. The scale bar represents the number of substitutions per site.

Phylogenetic relationships within Pangasiidae

The complete mtDNA sequence datasets (concatenated 13 PCGs) recovered *Pangasius* and *Pangasianodon* as sister subclades and the Pangasiidae as a monophyletic clade with a 100% bootstrap value (Fig. 2). The ML tree clarified the monophyly of *Pangasianodon* and *Pangasius* with 88% bootstrap support. The ML tree clarified the monophyly of *Pangasianodon*, with 88% bootstrap support. *Pangasianodon gigas* is identified as a sister taxon to a group of *P. sanitwongsei* and *Pn. hypophthalmus* (5 mitogenomes), and *Pangasius bocourti* (as named in GenBank under no. MN842723) is positioned in the *Pn. hypophthalmus* group with 100% bootstrap support. Within the *Pangasius*, *P. mekongensis* was recovered as a sister taxon to the Indian *Pangasius pangasius* Hamilton-Buchanan, 1822 (Hossain et al. 2009) with a very high boostrap (100%), while *P. krempfi* was placed in between *P. larnaudii* and the two above mentioned *Pangasius* species (*P. mekongensis* and *P. pangasius*) with a moderate bootstrap value (68%). The most concerning feature is that *P. sanitwongsei* was positioned as a sister taxon to a subgroup of all the *Pn. hypophthalmus* sequences (bootstrap 98%), and was in between this subgroup and *Pn. gigas* (Fig. 2). Several samples may have been misidentified and, therefore, phylogenetically misplaced. In fact, the sequence named "*Pangasianodon_hypophthalmus_*(BaijinCo-Foshan)-China-MZ286355" is *Pangasius larnaudii*, and the "*Pangasius bocourti* (QingyuanGD)-China-MN842723" sample is *Pangasianodon hypophthalmus*. These pangasiid sequences were, in fact, grouped into their corresponding generic clades (Fig. 2).

Furthermore, using single-gene datasets, we investigated the close phylogenetic relationships between *Pangasius* and *Pangasianodon*. Multiple *cox*1 and *cyt*B barcode sequences are available in GenBank and in previous publications, therefore, we downloaded all *cox*1 (551 bp) and *cyt*B (634 bp) and extracted *cox*1 and *cyt*B, respectively, from the complete mitogenomes (listed in Table S3). The *cox*1- and *cyt*B-topologies also revealed that the Pangasiidae were a sister group to the Austroglanididae. All three families (Pangasiidae, Cranoglanididae, and Austroglanididae) together with the Ictaluridae formed a large group that was always a sister group to the Ariidae, as discovered in the mitophylogeny constructed based on the complete mitogenome data in our current study and as in the previously reported analysis (Schedel et al. 2022).

The *cox*1-phylogenetic tree (Fig. 3) indicated that *P. mekongensis* (4 sequences) is a sister taxon to *P. pangasius* (8 sequences) with a relatively high bootstrap support (88%). In the *cyt*B tree (Fig. 4), this species is sistered with *P. pangasius* with 100% nodal support. *Pangasius krempfi*, in the *cox*1-tree, was shown to be close to the *Helicophagus* species (*Helicophagus leptorhynchus* and *Helicophagus waandersii*) with a low bootstrap of 44%, and in the *cyt*B-tree, was placed as a sister taxon, in a non-stable phylogenetic relationship with a 63% bootstrap, to *P. macronema* and *P. polyuranodon*. The partial *cox*1 datasets (four sequences, consisting of two from South Africa, one from China, and one from Cambodia) and *cyt*B datasets (two sequences from Thailand and one from China) of the correctly identified *P. sanitwongsei* species placed this taxon into the *Pangasius* clade, although with low support (bootstrap 36% and 29%, respectively) (Figs. 3 and 4). It should be noted that several *P. sanitwongsei* sequences, including Psan-(PB2)-(China)-JN020073 and Psan-(PB1)-CN-JN020086, as well as a *P. bocourti* sequence, Pboc-(QingyuanGD)-China-MN842723, were placed in the *Pn. hypophthalmus* cluster, which could be attributed to missampling or misidentification (Figs. 3 and 4).

In both the *cox*1- and *cyt*B-phylogenies, the *Pangasianodon* was clearly resolved as a sister to the *Pangasius* with medium nodal support (bootstrap: 86% for *cox*1 and 52% for *cyt*B). While the monophyletic *Pangasianodon* was resolved with its two members (*Pn. gigas* was the sister taxon to *Pn. hypophthalmus* in both the *cox*1- and the *cyt*B-phylogenetic topologies), the *Pangasius* cluster included not only *Pangasius* but *Helicophagus* and *Pseudolais* species as well. Although with low boostrap (28% for *cox*1- and 29% for *cyt*B-phylogeny), *Helicophagus* and *Pseudolais* species were grouped as siter taxa to some of the *Pangasius* species.
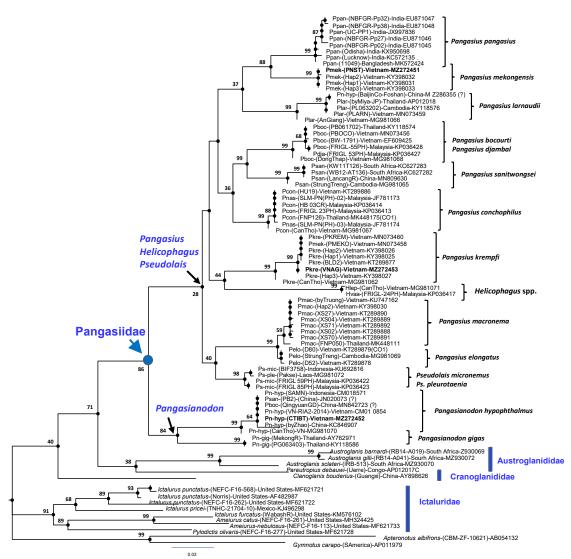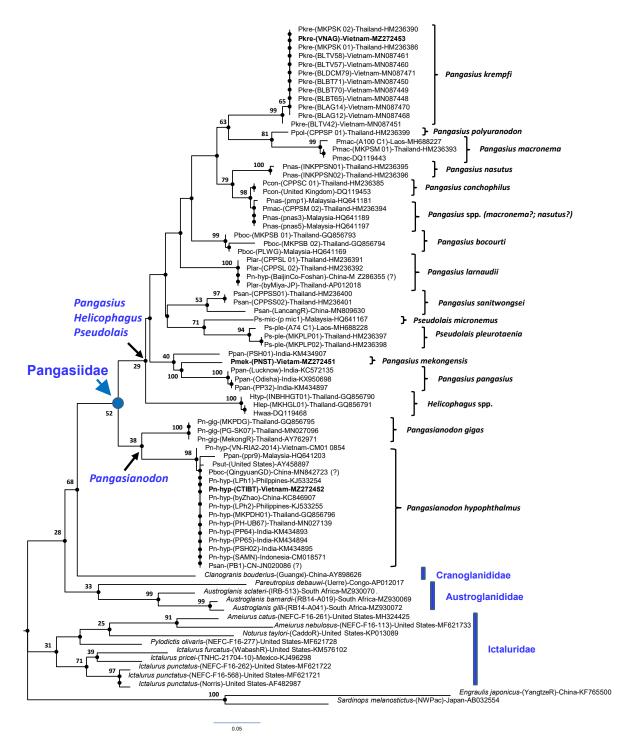
Fig. 3. Detailed PhyML-phylogeny based on the analysis of the partial *cox*1 sequences (551 bp) showing the detailed relationships of the family Pangasiidae and related families (Austroglanididae, Ictaluridae, and Cranoglanididae). In total, 83 sequences, including 81 from *Pangasius* and *Pangasianodon* and 2 outgroup sequences from the order Gymnotiformes, were included (Table S3). The alignment was performed by MAFFT (Katoh and Standley 2013), curated by BMGE v1.12 (Criscuolo and Gribaldo 2010), the tree was reconstructed in PhyML 3.3 (Guindon et al. 2010) using a maximum likelihood method and 1000 bootstrap resamplings, and the output Newick tree was extracted and visualized using FigTree v1.4.4 (Rambaut 2018). The basal nodes of the Pangasiidae and two sister groups (*Pangasianodon* and (*Pangasius* + *Helicophagus* + *Pseudolais*)) are shown by arrows. The *Pangasius mekongensis, Pangasianodon hypophthalmus*, and *Pangasius krempfi* sequences in this study were bolded. The taxon-misidentified sequences were added with a question mark at the end. The taxa from Pangasiidae were shortened and those from other related families were presented with their full names. The abbreviations of the isolates are given in brackets, including the geographical origin or voucher records of each sequenced specimen (where available), which were retrieved from the previous studies (Karinthanyakit and Jondeung 2012; Tran and Duong 2019; Schedel et al. 2022). The country of origin or where the sample was reported is given in full or in brackets, if available. Accession numbers are given at the end of each sequence label. The scale bar represents the number of substitutions per site.

Fig. 4. Detailed PhyML-phylogeny based on the analysis of the partial *cyt*B sequences (634 bp) showing the detailed relationships of the family Pangasiidae and related families (Austroglanididae, Ictaluridae, and Cranoglanididae). In total, 80 sequences, including 78 from *Pangasius* and *Pangasianodon* and 2 outgroup sequences from the order Clupeiformes, were included (**Table S3**). The alignment was performed by MAFFT (Katoh and Standley 2013), curated by BMGE v1.12 (Criscuolo and Gribaldo 2010), the tree was reconstructed in PhyML 3.3 (Guindon et al. 2010) using a maximum likelihood method and 1000 bootstrap resamplings, and the output Newick tree was extracted and visualized using FigTree v1.4.4 (Rambaut 2018). The basal nodes of the Pangasiidae and two sister groups (*Pangasianodon* and (*Pangasius* + *Helicophagus* + *Pseudolais*)) are shown by arrows. The *Pangasius mekongensis, Pangasianodon hypophthalmus*, and *Pangasius krempfi* sequences in this study were bolded. The taxon-misidentified sequences were added with a question mark at the end. The taxa from Pangasiidae were shortened and those from other related families were presented with their full names. The abbreviations of the isolates are given in brackets, including the

geographical origin or voucher records of each sequenced specimen (where available), which were retrieved from the previous studies. The country of origin or where the sample was reported is given in full or in brackets, if available. Accession numbers are given at the end of each sequence label. The scale bar represents the number of substitutions per site.

DISCUSSION

The characteristics of the 13 pangasiid mitogenomes were discussed, including the individual genes, non-coding regions, and comprehensive gene and genome features between members of the *Pangasius* and *Pangasianodon* species in the Pangasiidae. Sequence analysis has clarified the clearly associated generic interrelationships of the *Pangasius* and *Pangasianodon* species. This study used the concatenated 13 PCG sequences of 32 catfish families for the first time to resolve the mitophylogeny of Pangasiidae and the interfamilial and inter-major clades within Siluriformes. Additionally, single-gene phylogenies (based on partial *cox*1 and *cyt*B, respectively) were reconstructed in order to check the taxonomic and phylogenetic relationships within and between genera of Pangasiidae.

Pangasiid mitogenome structure and characteristics

The mtDNA length of *P. mekongensis* was the shortest of all Pangasiidae spp. sequenced to date, and four *Pangasius* spp. (*i.e.*, *P. mekongensis*, *P. krempfi*, *P. larnaudii*, and *P. pangasius*) appeared to have the shortest mitogenome among those complete siluroid mitogenomes obtained to date (Table S2) (Jondeung et al. 2007; Kim et al. 2018; Zhao et al. 2014; Mohindra et al. 2015). The base composition of 13 PCGs and mtDNAs and their skewness values were consistent among all the pangasiids, siluriform species, and other fishes (Satoh et al. 2016; Sharma et al. 2020). The structure is characterized by a similar pattern of equal use of A and C, a slight A over T, and much higher A, T, and C than G (C $\approx$ A > T > G). The nucleotide usage is clearly biased toward GC; thus, their negative skewness wholly contrasts with the T-rich nucleotides and the AT-biased codon usage in other metazoans. This unique mitogenomic feature is due to the imbalance of cytosine nucleotides, with more on the H-strand than on the L-strand (Satoh et al. 2016). All species in the genera *Pangasius* and *Pangasianodon* exhibit a similar pattern for base composition and skewness, with minimal differences. Between *Pangasius* and *Pangasianodon* species, relatively high divergence was found in *cyt*B, *nad*3, and *nad*4, and the most conserved PCG was found to be *atp*8. Among the PCGs, therefore, *cyt*B has been the most frequently used genetic marker (or barcode) for the phylogenetic analysis of catfish. This barcode marker, along with *cox*1, has significantly contributed to the investigation of the

appropriate evolutionary rates for this purpose (Karinthanyakit and Jondeung 2012; Miya and Nishida 2015; Hardman 2005; Kartavtsev et al. 2007; Tran et al. 2017).

The A-rich loop in the O_L sequence and the conserved –GGCGG– motif, which are reportedly required for L-strand replication in all vertebrates (Satoh et al. 2016), were found in all pangasiids studied so far. The CR in mtDNA of eight species (*P. mekongensis*, *P. krempfi*, *P. pangasius*, *P. larnaudii*, *P. bocourti*, *P. sanitwongsei*, *Pn. hypophthalmus*, and *Pn. gigas*) exhibits repeat-rich ETAS and CSBs, with the CSB-3 being almost identical among pangasiids. The CSB-2 and CSB-3 of these eight pangasiids retained the core sequence, which is identically conserved among siluriforms, for example, *Pseudoplatystoma reticulatum* and species in Pimelodidae (Villela et al. 2017) and *Silurus lanzhouensis* in Siluridae (Lian et al. 2015). The H-strand replication site sequence was found to be associated with the pyrimidine tract, which is highly conserved. A putative promoter for transcription of the L- and H-strands was found to be identical among all the studied pangasiids, indicating the essential function of these tracts in pangasiids and all bony fishes (Wang et al. 2007; Zhuang et al. 2013; Villella et al. 2017; Moreira et al. 2017). A distinguishing feature is that *Pangasius sanitwongsei* Smith, 1931 evidently appears to have a more similar CR to *Pn. hypophthalmus* in terms of length, sequence composition, and CSB patterns than other *Pangasius* species, although this taxon is a nominal species of the *Pangasius* genus (Wei et al. 2020; Fricke et al. 2023).

Mitophylogenetic analyses and taxonomic implications
Phylogenetic resolution of the Pangasiidae in Siluroidei

A maximum likelihood (ML) phylogenetic tree of the siluriform catfishes was successfully reconstructed. This was based on the analysis of the concatenated 13 PCGs of the complete mitogenomes, which were available from 117 sequences of 109 species in 32 families, with Gonorynchiformes as an outgroup (Table S2). In this study, the increasing coverage of complete mitogenomic sequences from Pangasiidae, as well as those from African *Austroglanis* catfishes, has resulted in a comprehensive reconstruction of the phylogenetic relationships of the Asian Pangasiidae, Cranoglanididae, and African Austroglanididae, with their relationship to the Ictaluridae (**Fig. 2**).

The monophyly of Pangasiidae was predicted, although few complete mitogenome datasets for phylogeny construction exist (Jondeung et al. 2007; Nakatani et al. 2011; Kappas et al. 2016; Kim et al. 2018). In previous studies, Jondeung et al. (2007) used the protein sequence dataset recovered from the entire mtDNA of *Pangasianodon gigas* (AY762971) to analyze the phylogenetic relationships among 15 families of Siluriformes. They found that the Mekong giant catfish of Pangasiidae was closer in relationship to Siluridae than to Bagridae. In other studies, the best-scoring ML tree of the 66 otophysan and 44 outgroup species, based on unambiguously aligned whole mitogenome sequences,

positioned *Pangasius* (*P. larnaudii*, AP012018) between the *Sciades* and paraphyletic to *Ictalurus* (Nakatani et al. 2011). These analyses were based only on the single mtDNA of one pangasiid species, *P. larnaudii*, as a Pangasiidae entry input. The outgoing results seemed to render Pangasiidae polyphyletic, but only a few mtDNA sequences of pangasiid taxa were available, and they lacked representatives in both the *Pangasius* and *Pangasianodon* genera. Kim et al. (2018) presented clearer generic and interfamilial resolutions by analyzing the complete mtDNAs of five pangasiid catfishes: two from *Pangasius* (*P. larnaudii*, AP012018, and *P. pangasius*, KC572135), and three from *Pangasianodon* (one *Pn. gigas*, AY762971, and two *Pn. hypophthalmus* (KC846907, and CM010854). However, the interfamilial and sistership relationships of the Pangasiidae were not yet well resolved.

In these previous studies, when only a limited number of pangasiid mitogenomes and no complete mitogenome from Austroglanididae were used, Pangasiidae seemed to position itself as a sister group to Cranoglanididae and Ictaluridae (Kappas et al. 2016; Kim et al. 2018). However, when complete mitogenome data from the *Austroglanis* catfish became available, Pangasiidae and Austroglanididae appeared to be sister families, at least according to our current study and a recent construction initiative reported by Schedel et al. (2022). Furthermore, the clade of Pangasiidae and Austroglanididae was monophyletically placed with the clade of Cranoglanididae and Ictaluridae (as sister groups) and with Ariidae to form a major clade with a reliable bootstrap support (75%) in Siluriformes (**Fig. 2**). Our complete mitogenomic datasets, therefore, support the consistent relationship of the (Pangasiidae + Austroglanididae + Cranoglanididae + Ictaluridae + Ariidae) group, which was recovered as a sister major group to the "Big Africa" major clade, as previously demonstrated based on nuclear gene datasets (Sullivan et al. 2006). This was consistent across all molecular analyses, including short and entire gene sequences, and all comprehensive mitogenomic datasets reported to date (Kappas et al. 2016; Zhang et al. 2021; Schedel et al. 2022). The current basal, intermediate position of the above major group (Pangasiidae + Austroglanididae + Cranoglanididae + Ictaluridae + Ariidae) between "Big Asia" and "Big Africa" requires more evidence once sufficient mitogenomic datasets (for example, from Heptapteridae or related families) are available to be used for reconstruction. Nevertheless, Austroglanididae and Pangasiidae contributed to the systematic resolution of the monophyly of Cranoglanididae and Ictaluridae in Siluroidei. The phylogenetic positions of *Diplomystes nahuelbutaensis* (Diplomystidae, of the suborder Diplomystoidei) and *Rita rita* (in the former grouping: family Bagridae), respectively, showed relatively low stability in the different analyses using different sequence markers. In our study, the Diplomystoidei were placed as a sister group to all the remaining Siluroidei species, excluding Trichomycteridae and families of the suborder Loricarioidei. The Ritidae family, as recently used by Schedel et al. (2022), is represented by only one species, *Rita rita* from Pakistan, which was previously systematically classified into the Bagridae family (Punhal et al. 2015; Fricke et al. 2023). This species

has a constraining and labile phylogenetic position, far from Bagridae and yet now rooted with Cetopsidae, forming the subclade of which is placed as a sister to all the remaining Siluroidei taxa. *Rita rita* (of the family Bagridae or Ritidae) is still a labile species with an uncertain systematic position in Siluroidei, and its phylogenetic relationship needs to be further investigated (Hardman 2005; Sullivan et al. 2006; Kappas et al. 2016; Kim et al. 2018; Zhang et al. 2022; Schedel et al. 2022).

To classify catfishes' familial relationships, Kappas et al. (2016) have conclusively recovered the two "Big Asia" and "Big Africa" multifamilial clades, in which Pangasiidae was revealed to be a sister-group of the "Big Africa" clade, distinct from other representatives of the "Big Asia" clade. Our study clearly resolves the mitophylogenetic classification of seven *Pangasius* and *Pangasianodon* species in Siluroidei. By incorporating more pangasiid mitogenomes, Pangasiidae's monophyletic status and taxonomic relationships have been better clarified. However, we have only set out to assess the intra- and intergeneric and phylogenetic relationships of the *Pangasius* and *Pangasianodon* species, not the origination of the Asian and Mekong River pangasiids.

Inter- and intra-generic phylogenetic relationships of the Pangasiidae

In this study, the complete mitogenome sequences from 13 samples of seven Pangasiidae species were available for mitophylogeny construction (Table S2), which revealed the clade from the South and Southeast Asian *Pangasius* and *Pangasianodon* species. With 100% bootstrap support, the complete mitogenome datasets recovered *Pangasianodon* Chevey, 1931, as sister to *Pangasius* Valenciennes, 1840 (Fig. 2). The monophyletic genera *Pangasianodon* and *Pangasius* were also recovered, as shown in the previous phylogeny reconstructions (Kappas et al. 2016; Kim et al. 2018; Schedel et al. 2022). Unfortunately, all these phylogenies, including our reconstruction, were performed solely with mitogenomes from two genera, *Pangasianodon* and *Pangasius*. No complete mitogenome datasets of the other two valid genera of Pangasiidae, *Helicophagus* and *Pseudolais*, were yet available.

Within Pangasiidae, four *Pangasius* species (*P. mekongensis*, *P. krempfi*, *P. pangasius*, and *P. larnaudii*) were found to comprise one well-supported group (100% bootstrap support), where *P. mekongensis* was a sister taxon to *Pangasius pangasius*. It should be noted that the species "*Pangasianodon_hypophthalmus_*(BaijinCo-Foshan)-China-MZ286355" was identified as *Pangasius larnaudii* and assigned a support value of 100% to *P. larnaudii*. The topology showed the taxonomic placement in this *Pangasius* subclade as (*P. larnaudii* + (*P. krempfi* + (*P. mekongensis* + *P. pangasius*)). All the *Pangasius* species except *Pangasius sanitwongsei* (from the Lancang River, China, GenBank: MN809630) formed a subclade that was monophyletically situated as sister to the *Pangasianodon* subgroup. The phylogeny based on the studied complete mitogenome sequences has

not recovered *Pangasius sanitwongsei* Smith, 1931 (Chinese sample, GenBank: MN809630) in the assigned *Pangasius* genus, as morphologists have demonstrated (Roberts and Vidthayanon, 1991; Hogan et al. 2009). The grouping of *P. sanitwongsei* into the *Pangasianodon* cluster in the mitogenome-dataset tree was well supported by a high bootstrap (98%) (**Fig. 2**). In a previous study (Wei et al. 2020), *P. sanitwongsei* was positioned as a sister taxon to *P. larnaudii*, and *P. pangasius* was far removed from Pangasiidae, being placed between the families Bagridae and Sisoridae. This constraining phylogeny might be reconsidered because it was constructed using the neighbor-joining method with 1000 bootstrap replicates and because many other *Pangasius* and *Pangasianodon* mitogenomes were absent. The limited data coverage might affect the phylogenetic status, but it did not reflect the precise taxonomic relationship of this taxon. In our study, using the ML method, *P. sanitwongsei* was absent in the *Pangasius* subclade, and the subcluster encompassing *P. mekongensis, P. pangasius*, and *P. krempfi* fully separated *P. sanitwongsei* from *P. larnaudii* with 100% bootstrap support.

The partial single PCG (*cox*1 and *cyt*B) phylogenetic analyses (Figs. 3 and 4) revealed phyletic arrangements within the Pangasiidae clade from *Pangasius* to *Pangasianodon*. When partial *cox*1 and *cyt*B sequences available from *Helicophagus* and *Pseudolais* species were included, the phylogenetic trees did not recover *Pangasius* as monophyletic or as a sister subclade to *Pangasianodon.* In both single-gene phylogenies, the *Pangasius* species were clustered with *Helicophagus* and *Pseudolais* species in the same subclade and appeared to be paraphyletic. This paraphyletic appearance of the *Pangasius* species in the presence of species from four genera in the Pangasiidae has also been reported by Quyen et al. (2018). A study of pangasiids in Thailand also concluded that *Pangasius* is the sister group to *Helicophagus* and *Pseudolais* (Karinthanyakit and Jondeung 2012).

Another notable phyletic position was related to *P. sanitwongsei*. The representative samples of *P. sanitwongsei* (Table S3) came from China (Lancang River) (Wei et al. 2018), South Africa (GenBank: KC627282, Breede River, near Bonnievale; and KC627283, KwaZulu-Natal, Durban) (Mäkinen et al. 2013), Cambodia (Stung Treng) (Quyen et al. 2018), and Thailand (Chao Phraya River, Nakhon Sawan) (Karinthanyakit and Jondeung 2012). The phylogenies recovered all of these sequences as a sister group to *Pangasius bocourti* (*cox*1-based) or to *Pseudolais micronemus* and *Pseudolais pleurotaenia* species (*cyt*B-based phylogeny), respectively, and all were placed within the (*Pangasius + Helicophagus + Pseudolais*) subclade. Given that the samplings and morphological identifications for *P. sanitwongsei* were correct, the contradictory phylogenies may suggest caution when approaching barcode data and selecting genetic markers, particularly single-gene barcodes, for taxonomic phylogeny reconstruction. Integrating comprehensive DNA barcodes, robust NGS-based data, and morphological characterization is the best method for the complex systematic and taxonomy clarification of fishes, including siluriforms (Hubert et al. 2015).

However, in the sequence characterization of the complete mitogenome, *P. sanitwongsei* exhibited *Pangasianodon*-like characteristics, including identical CSBs in the CR (Fig. S1), a highly similar pattern of genetic distances in the mtDNA coding region (Table S6), and close generic and species relationships in the phylogenetic tree (Fig. 2). These characteristics lead to the contentious species relationship of *Pangasius sanitwongsei* within Pangasiidae. Further investigation with the complete mitogenomes of different geographical samples of *P. sanitwongsei* (Karinthanyakit and Jondeung 2012; Mäkinen et al. 2013; Wei et al. 2018; Quyen et al. 2018) may help clarify this taxon's taxonomic and phylogenetic relationships.

The pangasiid generic and familial datasets significantly contributed to their comprehensive, growing data-based sequence and phylogenetic analyses in Siluriformes. With a high nodal bootstrap (72%), Pangasiidae, which is monophyletic, was placed as a sister group to the Austroganididae clade comprising African *Austroglanis* catfishes that was originally analyzed and first reported by Schedel et al. (2022). Our analyses confirmed these phylogenetic relationships, which were based on the complete genome or on individual gene data using *cox*1 and *cyt*B (Figs. 1, 2, and 3). It should be noted, however, that all complete mitodataset-based phylogenies reconstructed to date have only used mitogenomes from the *Pangasius* Valenciennes, 1840, and *Pangasianodon* Chevey, 1931 species, with no representatives from *Helicophagus* Bleeker, 1858, or *Pseudolais* Vaillant, 1902 (the other two recognized genera in Pangasiidae). With the future inclusion of the complete mitogenome data from the *Helicophagus* and *Pseudolais* genera for the phylogenetic analyses, we anticipate that the *Pangasianodon* intrageneric relationships (between *Pn. gigas* and *Pn. hypophthalmus*) will not change, but the *Pangasius* intrageneric and the *Pangasianodon* intergeneric relationships will most likely be rearranged. Unfortunately, there is no complete mtDNA sequence for *Helicophagus* and *Pseudolais* species such as *Helicophagus leptorhynchus, Helicophagus waandersii, Helicophagus typus,* and/or *Pseudolais micronemus*, *Pseudolais pleurotaenia*, or for other related *Pangasius* species that could help resolve the phylogenetic position of *Pangasius* species, especially for unsolved or contradictory species such as *Pangasius sanitwongsei*.

CONCLUSIONS

This study presents the fully annotated mitogenomes of *Pangasius mekongensis*, *Pangasius krempfi*, and *Pangasianodon hypophthalmus* catfishes (from the Mekong River Basin in Vietnam) and describes their genomic features. Comparative sequence analyses have characterized the base composition, genetic distance, codon usage, regulatory elements, and conserved blocks of these species in comparison with those of other pangasiid congeners. The CR tends to be divided into two

types of sequences, *Pangasius* and *Pangasianodon*, although the elements and blocks of mtDNA are highly conserved. The monophyletic Pangasiidae and its related Austroglanididae group, which has been monophyletically recovered as a sister group to the (Ictaluridae + Cranoglanididae) group, were found in the complete mitodataset phylogeny. The [Pangasiidae + Austroglanididae] + (Ictaluridae + Cranoglanididae) + Ariidae] clade is positioned as a sister group to the "Big Africa" major clade in Siluriformes. Single-gene (*cox*1 and *cyt*B) phylogenetic analyses corroborated the previously reported monoplyletic *Pangasianodon* and paraphyletic groups of *Pangasius*, *Helicophagus*, and *Pseudolais*. Although phylogenetic analyses support the monophyly of the family Pangasiidae in Siluroidei, the phylogenetic and taxonomic relationships of *Pangasius* species, including *Pangasius sanitwongsei*, still remain to be clarified. Additional complete pangasiid mitogenomes are required to resolve these points. The datasets and the mitophylogenetic analyses of pangasiids reported in this study establish a foundation for reappraising pangasiid taxonomy, with implications for identification, DNA barcoding, systematics, phylogenetics, population genetics, and timeline and mode of diversification studies of siluriform catfishes.

Authors' contributions: TYD and NTTN: sample collection, examination and species identification, draft preparation and reviewing. LTKP and XTKL: laboratory work, raw sequence analysis and data curation. SAMN: draft reviewing and validation. LTH: fine data analyses, figures/tables preparation, draft writing-reviewing-editing and approval.

Competing interests: The authors declare no conflict of interest.

Availability of data and materials: Complete sequences have been deposited in GenBank with accession numbers provided in the text. Additional supporting information can be found online in the Supplemental Information section.

Consent for publication: Not applicable.

Ethics approval consent to participate: Not applicable.

REFERENCES

Baird I. 2011. *Pangasius krempfi*, The IUCN Red List of Threatened Species, e.T181328A7668262. Downloaded on 24 August 2021. doi:10.2305/IUCN.UK.2011-1.RLTS.T181328A7668262.en.

Betancur RR, Wiley EO, Arratia G, Acero A, Bailly N, Miya M, Lecointre G, Ortí G. 2017. Phylogenetic classification of bony fishes. BMC Evol Biol **17:**162. doi:10.1186/s12862-017-0958-3.

Boore JL. 1999. Animal mitochondrial genomes. Nucleic Acids Res **27:**1767–1780. doi:10.1093/nar/27.8.1767.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol **10:**210. doi:10.1186/1471-2148-10-210.

Cui L, Dong Y, Cao R, Zhou X, Lu S. 2020. Characterization of the Complete Mitochondrial Genome of *Arius dispar* (Siluriformes: Ariidae) and Phylogenetic Analysis Among Sea Catfishes. J Ocean Univ China **19(5):**1198–1206. doi:10.1007/s11802-020-4409-1.

Diogo R. 2004. Phylogeny, origin and biogeography of catfishes: support for a Pangean origin of 'modern teleosts' and reexamination of some Mesozoic Pangean connections between the Gondwanan and Laurasian supercontinents. Anim Biol **54(4):**331–351. doi:10.1163/1570756042729546.

Dwivedi AK, BK Gupta, RK Singh, V Mohindra, S Chandra, S Easawarn, Jena J, Lai KK. 2017. Cryptic diversity in the indian clade of the catfish family pangasiidae resolved by the description of a new species. Hydrobiologia **797:**351–370. doi:10.1007/s10750-017-3198-z.

FAO. 2022. The State of World Fisheries and Aquaculture (SOFIA) 2022 - Towards blue transformation (https://www.fao.org/publications/sofia/2022/en/).

Fischer C, Koblmüller S, Gülly C, Schlötterer C, Sturmbauer C, Thallinger GG. 2013. Complete mitochondrial DNA sequences of the threadfin cichlid (*Petrochromis trewavasae*) and the blunthead cichlid (*Tropheus moorii*) and patterns of mitochondrial genome evolution in cichlid fishes. PLoS ONE **8(6):**e67048. doi:10.1371/journal.pone.0067048.

Fricke R, Eschmeyer WN, van der Laan (eds). 2023. Eschmeyer's catalog of fishes: Genera, Species, References. (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version. Accessed 12 June 2023.

Froese R, Pauly D (Eds). 2021. FishBase, Pangasiidae Bleeker, 1858. Accessed through: World Register of Marine Species at: http://www.marinespecies.org/aphia.php?p=taxdetails&id=154648 on 2021-07-01

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol **59(3):**307–321. doi:10.1093/sysbio/syq010.2010.

Guo X, Liu S, Liu Y. 2003. Comparative analysis of the mitochondrial DNA control region in cyprinids with different ploidy level. Aquaculture **224:**25–38. doi:10.1016/S0044-8486(03)00168-6.

Gustiano R, Prakoso VA, Ath-thar MHF, Kusmini II, Radona D. 2021. Similarity and genetic relationship analysis of 28 species of Pangasiidae (Siluriformes, Ostariophysi). IOP Conference Series: Earth and Environmental Science **674:**012014. doi:10.1088/1755-1315/674/1/012014.

Gustiano R, Teugels GG, Pouyaud L. 2003. Revision of the *Pangasius kunyit* catfish complex, with description of two new species from South-East Asia (Siluriformes; Pangasiidae). J Nat Hist **37:**357–376. doi:10.1080/713834687.

Gustiano R. 2009. Pangasiid Catfishes of Indonesia. Bul Plasma Nutfah **15:**91–100.

Hardman M. 2005. The phylogenetic relationships among non-diplomystid catfishes as inferred from mitochondrial cytochrome b sequences; the search for the ictalurid sister taxon (Otophysi: Siluriformes). Mol Phylogenet Evol **37(3):**700–720. doi:10.1016/j.ympev.2005.04.029.

Hogan Z, Baird IG, Radtke R, Vander Zanden MJ. 2007. Long distance migration and marine habitation in the tropical Asian catfish, *Pangasius krempfi.* J Fish Biol **71:**818–832. doi:10.1111/j.1095-8649.2007.01549.x.

Hogan Z, Na-Nakorn U, Kong H. 2009. Threatened fishes of the world: *Pangasius sanitwongsei* Smith, 1931 (Siluriformes: Pangasiidae). Environ Biol Fishes **84(3):**303–304. doi:10.1007/s10641-008-9419-6.

Hossain MY, Rhaman MM, Mollah MFA. 2009. Threatened fishes of the world: *Pangasius pangasius* Hamilton-Buchanan, 1822 (Pangasiidae). Environ Biol Fishes **84(3):**315–316. doi:10.1007/s10641-008-9422-y.

Hubert N, Wibowo AK, Busson F, Caruso D, Sulandari S, Nafiqoh N, Pouyaud L, Rüber L, Avarre JC, Herder F, Hanner R, Keith P, Hadiaty RK. 2015. DNA Barcoding Indonesian freshwater fishes: challenges and prospects. DNA Barcodes **3:**144–169.

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, Sado T, Mabuchi K, Takeshima H, Miya M, Nishida M. 2013. MitoFish and MitoAnnotator: a mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Mol Biol Evol **30(11):**2531–2540. doi:10.1093/molbev/mst141.

Jondeung A, Sangthong P, Zardoya R. 2007. The complete mitochondrial DNA sequence of the Mekong giant catfish (*Pangasianodon gigas*), and the phylogenetic relationships among Siluriformes. Gene **387(1-2):**49–57. doi:10.1016/j.gene.2006.08.001.

Junier T, Zdobnov EM. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the Unix shell. Bioinformatics **26(13):**1669–1670. doi:10.1093/bioinformatics/btq243.

Kappas I, Vittas S, Pantzartzi CN, Drosopoulou E, Scouras ZG. 2016. A time-calibrated mitogenome phylogeny of catfish (Teleostei: Siluriformes). PLoS ONE **11(12):**e0166988. doi:10.1371/journal.pone.0166988.

Karinthanyakit W, Jondeung A. 2012. Molecular Phylogenetic Relationships of Pangasiid and Schilbid Catfishes in Thailand. J Fish Biol **80:**2549–2570. doi:10.1111/j.1095-8649.2012.03303.x.

Kartavtsev YP, Jung SO, Lee YM, Byeon HK, Lee JS. 2007. Complete mitochondrial genome of the bullhead torrent catfish, *Liobagrus obesus* (Siluriformes, Amblycipididae): Genome description and phylogenetic considerations inferred from the Cyt b and 16S rRNA genes. Gene **396(1):**13–27. doi:10.1016/j.gene.2007.01.027.

Katoh K, Standley DM. 2013. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30:**772–780. doi:10.1093/molbev/mst010.

Kim OTP, Nguyen PT, Shoguchi E, Hisata K, Vo TTB, Inoue J, Shinzato C, Le BTN, Nishitsuji K, Kanda M, Nguyen VH, Nong HV, Satoh N. 2018. A draft genome of the striped catfish, *Pangasianodon hypophthalmus*, for comparative analysis of genes relevant to development and a resource for aquaculture improvement. BMC Genomics **19(1):**733. doi:10.1186/s12864-018-5079-x.

Kottelat M. 2013. The fishes of the inland waters of Southeast Asia: A catalogue and core bibliography of the fishes known to occur in freshwaters, mangroves and estuaries. Raffles Bull Zool Supplement **27:**1–663.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol **35:**1547–1549. doi:10.1093/molbev/msy096.

Laslett D, Canback B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics **24:**172–175. doi:10.1093/bioinformatics/btm573.

Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, Gascuel O. 2019. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. Nucleic Acids Res **47(W1):** W260–W265. doi:10.1093/nar/gkz303.

Li HY, Xu TJ, Cheng YZ, Sun DQ, Wang RX. 2012. Genetic diversity of Setipinna taty

(Engraulidae) populations from the China Sea based on mitochondrial DNA control region sequences. Genet Mol Res **11(2):**1230–1237. doi:10.4238/2012.

Lian ZQ, Wu XD, Xiao W, Sai QY, Gun SB. 2015. Complete sequence and characterization of the *Silurus lanzhouensis* (Siluriformes: Siluridae) mitochondrial genome. A DNA Mapp Seq Anal **27:**2483–2484.

Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: Search and Contextual Analysis of Transfer RNA Genes. Nucleic Acids Res **44:**W54–57. doi:10.1093/nar/gkw413.

Mäkinen T, Weyl OLF, van der Walt KA, Swartz ER. 2013. First Record of an Introduction of the giant *Pangasius*, *Pangasius sanitwongsei* Smith 1931, into an African river. Afr Zool **48(2):**388–391. doi:10.1080/15627020.2013.11407606.

Miya M, Nishida M. 2015. The mitogenomic contributions to molecular phylogenetics and evolution of fishes: a 15-year retrospect. Ichthyol Res **62:**29–71. doi:10.1007/s10228-014-0440-9.

Mohindra V, Singh RK, Kumar R, Sah RS, Lal KK. 2015. Complete mitochondrial genome sequences of two endangered Indian catfish species, *Clarias batrachus* and *Pangasius pangasius*. Mitochondrial DNA **26(5):**678–679. doi:10.3109/19401736.2013.840604.

Moreira DA, Buckup PA, Furtado C, Val AL, Schama R, Parente TE. 2017. Reducing the information gap on Loricarioidei (Siluriformes) mitochondrial genomics. BMC Genomics **18(1):**345. doi:10.1186/s12864-017-3709-3.

NAGAO 2021. Fishes of the Indochinese Mekong. Nagao Natural Environmental Foundation, Tokyo, xii+546 pages.

Nakatani M, Miya M, Mabuchi K, Saitoh K, Nishida M. 2011. Evolutionary history of Otophysi (Teleostei), a major clade of the modern freshwater fishes: Pangaean origin and Mesozoic radiation. BMC Evol Biol **11(1):**177. doi:10.1186/1471-2148-11-177.

Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol **41:**353–358. doi:10.1007/BF00186547.

Phan LT, Bui TM, Nguyen TT, Gooley GJ, Ingram BA, Nguyen HV, Nguyen PT, De Silva SS. 2009. Current status of farming practices of striped catfish, *Pangasianodon hypophthalmus* in the Mekong Delta, Vietnam. Aquaculture **296(3-4):**227–236. doi:10.1016/j.aquaculture.2009.08.017.

Poulsen AF, Hortle KG, Valbo-Jorgensen J, Chan S, Chuon CK, Viravong S, Bouakhamvongsa K, Suntornratana U, Yoorong N, Nguyen TT, Tran BQ. 2004. Distribution and ecology of some important riverine fish species of the Mekong River Basin. MRC Technical Papers **10:**1–116. MRC Secretariat. doi:10.52107/mrc.ak8j7v.

Pouyaud L, Teugels GG. 2000. Description of a new pangasiid catfish from east Kalimantan,

Indonesia (Siluriformes: Pangasiidae). Ichthyol Explor freshwaters **11:**193–200.

Punhal L, Laghari MY, Xu P, Zhao Z, Jiang L, Narejo NT, Deng Y, Sun X, Zhang Y. 2015. Complete mitochondrial genome of the Freshwater Catfish *Rita rita* (Siluriformes, Bagridae). Mitochondrial DNA **26(6):**817–818. doi:10.3109/19401736.2013.855908.

Quyen V, T Oanh, TLP Thai, LT Tran, TD Binh. 2018. Molecular phylogeny of catfishes (Teleostei: Siluriformes) inferred from mitochondrial markers –implications for lower Mekong River basin. EJARBLS **6:**1–12.

Rainboth WJ. 1996. Fishes of the Cambodian Mekong, FAO species identification field guide for fishery purposes. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy. doi:10.1017/CBO9781107415324.004.

Rambaut A. 2018. FigTree, version 1.4.4. http://tree.bio.ed.ac.uk/software/figtree/.

Roberts TR, Vidthayanon C. 1991. Systematic revision of the Asian catfish family Pangasiidae, with biological observations and descriptions of 3 new species. Proc Acad Nat Sci Philadelphia **143:**97–144. doi:10.2307/4064995.

Saitoh TP, Miya M, Mabuchi K, Nishida M. 2016. Structure and variation of the mitochondrial genome of fishes. BMC Genomics **17(1):**719. doi:10.1186/s12864-016-3054-y.

Schedel FDB, Chakona A, Sidlauskas BL, Popoola MO, Usimesa Wingi N, Neumann D, Vreven EJWMN, Schliewen UK. 2022. Newphylogenetic insights into the African catfish families Mochokidae and Austroglanididae.Journal of Fish Biol **100(5):**1171–1186. doi:10.1111/jfb.

Sharma A, Siva C, Ali S, Sahoo PK, Nath R, Laskar MA, Sarma D. 2020. The complete mitochondrial genome of the medicinal fish, *Cyprinion semiplotum*: Insight into its structural features and phylogenetic implications. Int J Biol Macromol **164:**939–948. doi:10.1016/j.ijbiomac.2020.07.142

Sullivan JP, Lundberg JG, Hardman M. 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using *rag1* and *rag2* nuclear gene sequences. Mol Phylogenet Evol **41(3):**636–62. doi:10.1016/j.ympev.2006.05.044.

Tran HTT, Tran TN, Tran HNA, Nguyen HT. 2017. DNA Barcoding and phylogenetic relationships of nine catfish species from Mekong Basin, Vietnam. J Mol Biomark Diagn **8:**363. doi:10.4172/2155-9929.1000363.

Tran TML, Duong TY. 2019. Differentiation of two *Pangasius* species, *Pangasius krempfi* and *Pangasius mekongensis* using inter-sequence simple repeat markers. Int J Fish Aquat Stud **7(4):**116–120.

Vidthayanon C, Hogan Z. 2011. *Pangasianodon hypophthalmus*, The IUCN Red List of Threatened Species, e.T180689A7649971. https://dx.doi.org/10.2305/IUCN.UK.2011-1.RLTS.T180689A7649971.en. Downloaded on 24 May 2021.

Vidthayanon C. 2012. *Pangasius mekongensis*, The IUCN Red List of Threatened Species, e.T181260A171473. Downloaded on 24 May 2021. doi:10.2305/IUCN.UK.2012-1.RLTS.T181260A1714731.en.

Villela LC, Alves AL, Varela ES, Yamagishi ME, Giachetto PF, da Silva NM, Ponzetto JM, Paiva SR, Caetano AR. 2017. Complete mitochondrial genome from South American catfish *Pseudoplatystoma reticulatum* (Eigenmann & Eigenmann) and its impact in Siluriformes phylogenetic tree. Genetica **145(1):**51–66. doi:10.1007/s10709-016-9945-7.

Vu AV, Baumgartner LJ, Limburg KE, et al. 2022. Life history strategies of Mekong pangasiid catfishes revealed by otolith microchemistry. Fish Res **249:**106239. doi:10.1016/j.fishres.2022.106239.

Wang TY, Tzeng CS, Teng HY, Chang T. 2007. Phylogeography and identification of a 187-bp-long duplication within the mitochondrial control region of *Formosania lacustre* (teleostei: Balitoridae). Zool Stud **46:**569.

Wei L, Ye X, Lv Y, Teng Z, Gan B, Zou H, Mo F, Zhang S. 2020. Complete mitochondrial genome and phylogenetic position of *Pangasius sanitwongsei* (Siluriformes: Pangasiidae). Mitochondrial DNA B: Res **5(1):**945–946. doi:10.1080/23802359.2020.1719915.

Zhang K, Liu Y, Chen J, Zhang H, Gong L, Jiang L, Liu L, Lü Z, Liu B. 2021. Characterization of the complete mitochondrial genome of *Macrotocinclus affinis* (Siluriformes; Loricariidae) and phylogenetic studies of Siluriformes. Mol Biol Rep **48:**677–689. doi:10.1007/s11033-020-06120-z.

Zhang R, Deng L, Lv X, Tang Q. 2022. Complete mitochondrial genomes of two catfishes (Siluriformes, Bagridae) and their phylogenetic implications. Zookeys **1115:**103–116. doi:10.3897/zookeys.1115.85249.

Zhao H, Kong X, Zhou C. 2014. The mitogenome of *Pangasius sutchi* (Teleostei, Siluriformes: Pangasiidae). Mitochondrial DNA **25(5):**342–344. doi:10.3109/19401736.2013.800492.

Zhuang X, Qu M, Zhang X, Ding S. 2013. A Comprehensive Description and Evolutionary Analysis of 22 Grouper (Perciformes, Epinephelidae) Mitochondrial Genomes with Emphasis on Two Novel Genome Organizations. PLoS ONE **8(8):**e73561. doi:10.1371/journal.pone.0073561.

Supplementary materials

Fig. S1. Alignment of ten control region sequences from eight species, *Pangasius mekongensis* (Pmek), *Pangasius krempfi* (Pkre), *Pangasius pangasius* (Ppan), *Pangasius larnaudii* (Plar), *Pangasius bocourti* (Pboc), *Pangasius sanitwongsei* (Psan), *Pangasianodon hypophthalmus* (Phyp, three sequences), and *Pangasianodon gigas* (Pgig). Types of conserved sequence blocks (CSBs) within a species (CSB-C), within a genus (CSB-B), and within both genera (*Pangasius* and *Pangasianodon*) (CSB-A) are presented. Aligned nucleotide sequences and typical CSBs are shown as boxes. For species, see Table S2; for descriptions of the blocks, see text. (download)

Table S1. Primers for amplification and sequencing of fragments of the mitochondrial genome of *Pangasius mekongensis, Pangasianodon hypophthalmus*, and *Pangasius krempfi* of the Mekong River, Vietnam. (download)

Table S2. List and information of 117 siluriform sequences (109 species in 32 families) and two outgroup species with complete mitogenomes providing 13 protein-coding sequences used in this study for phylogenetic and sequence analysis of catfishes of the Pangasiidae and families in the order Siluriformes. (download)

Table S3. Accession numbers and country report for the reference *cox*1 and *cyt*B markers from the GenBank database and those from this study used for the pangasiid comparative phylogenetic studies. (download)

Table S4. Nucleotide composition and skewness value for the complete mitochondrial genome (mtDNA) and the protein-coding genes (PCGs) of seven species members of the family Pangasiidae. (download)

Table S5. Pairwise nucleotide differences (%) among five catfish species of the family Pangasiidae present in the Mekong Basin for individual mitochondrial protein-coding genes and mitochondrial ribosomal genes (12S and 16S). (download)

Table S6. Pairwise genetic distances (%) estimated between *Pangasius mekongensis*, *Pangasianodon hypophthalmus*, *Pangasius krempfi*, and the published or GenBank-deposited representative Pangasiidae species based on mitogenome coding nucleotide sequences. (download)