Genetic Differentiation and Isolation by Distance in Mekong River Fishes With Typical Migration Patterns

OANH THI TRUONG[1,2], SANG QUANG TRAN[2], QUYEN VU DANG HA[2], VAN NGO THAI BICH[1], BINH THUY DANG[2,*]

[1]Faculty of Chemical Engineering, The University of Danang, University of Science and Technology, Da Nang, Vietnam

[2]Department of Biotechnology, Institute for Biotechnology and Environment, Nha Trang University, Khanh Hoa, Vietnam

*E-mail: binhdt@ntu.edu.vn

| Received: 26/06/2024; Accepted: 21/05/2025

E-ISSN: 2073-3720
https://doi.org/10.33997/j.afs.2025.38.3.001

Abstract

Riverine fishes exhibit diverse life history traits that are the result of evolutionary processes and local adaptation. Geographical sub-populations have been influenced by complex biological and physical factors. This study evaluates the level of genetic differentiation and isolation by distance (IBD) in three Mekong fishes with typical migratory patterns. *Labeo chrysophekadion* (Bleeker, 1849) and *Pangasius larnaudii* Bocourt, 1866, are both potamodromous "white" fish, the former is considered facultative in migration, while the latter undertakes long-distance migration. *Macrognathus siamensis* (Günther, 1861) is a "black", sedentary and non-migratory fish. Fish species from 7–9 natural populations were genotyped for over 800 single nucleotide polymorphisms (SNPs). Genetic differentiation indexes (F_{ST} values) revealed spatial patterns among geographical populations. Mantel tests supported strong IBD signals in both *M. siamensis* and *L. chrysophekadion*, while no IBD tendency was observed in *P. larnaudii* at increasing spatial scales. All three species showed a positive correlation between genetic clusters and distance in dbMEM analyses. This study highlights spatial genetic patterns and varying IBD signals corresponding to different fish migratory patterns, supporting species-specific and multi-species management strategies.

Keywords: genetic differentiation, isolation by distance, Mekong fishes, migration patterns, single nucleotide polymorphisms

Introduction

The Mekong River, the largest tropical river in Asia, originates in the Tibetan Plateau of China and flows through six countries before finally draining into the East Sea in Vietnam (Adamson et al., 2009). The section passing through Lao People's Democratic Republic (PDR), Thailand, Cambodia, and Vietnam is known as the Lower Mekong Basin (LMB), which is divided into five ecological regions: The Upper Mekong, Middle Mekong, Floodplain, Tonlé Sap, and Delta regions (Grill et al., 2014).

The Mekong River is characterised by its vast range of geographic and climatic zones (Winemiller et al., 2016), and supports multiple types of habitats, including mainstream, tributaries, lakes, deep pools, seasonal floodplains, reservoirs, wetlands, caves, and estuaries (Grill et al., 2014; Kang and Huang, 2021). Its hydrology is highly dynamic, influenced by seasonal flood pulses, making it one of the rivers with the most diverse fish populations and inland fisheries (Valbo-Jørgensen et al., 2009). There are 1,148 recognised species, of which 87 % being migratory fish (Poulsen et al., 2002). Based on their migratory patterns, fish species are further classified into three guilds – black, grey, and white fishes (Kang and Huang, 2021). Information on the distribution and migration routes of fish species is the basis for identifying six distinct but interconnected systems in the LMB: long-distance migration (The Upper, Middle, and Lower systems), and lateral migration (Great Lake, Se San-Se Kong-Sre Pok (3S), and Korat systems) (Kang and Huang, 2021).

In the context of complex ecosystems and migration systems in the LMB, *Macrognathus siamensis* (Günther,

1861) (Synbranchiformes: Mastacembelidae), *Labeo chrysophekadion* (Bleeker, 1849) (Cypriniformes: Cyprinidae), and *Pangasius larnaudii* Bocourt, 1866 (Siluriformes: Pangasiidae) may be considered suitable model fish species representing typical migration patterns. Among these, *M. siamensis* is a "black", sedentary and non-migratory species (Rainboth, 1996). In contrast, the other two species, *Labeo chrysophekadion* (Bleeker, 1849) and *Pangasius larnaudii* Bocourt, 1866, are potamodromous "white" fish. *Labeo chrysophekadion* is considered a facultative migratory species, an opportunistic spawner in a variety of habitats, including man-made impoundments (Poulsen et al., 2004; Vu et al., 2022). *Pangasius larnaudii,* on the other hand, is known as an extensive longitudinal migration fish (Poulsen et al., 2004; Hogan et al., 2007).

For Mekong fish species, the dispersal strategy of eggs and larvae is to drift along floodwaters to downstream floodplains (Hughes, 2024), so they often spawn at the beginning of the rainy season, and the three fish species studied are no exception. The Khone Falls – a 40 km wide series of waterfalls, rapids, and channels located near the Lao PDR-Cambodian border – is a well-known physical barrier between the middle and lower migration systems, which caused a different seasonal migration pattern of two "white" fish species (Baran et al., 2005). At the beginning of the rainy season, fish above the Khone Falls migrate upstream, while those below migrate downstream. At the onset of the dry season, widespread upstream migration was recorded below the Khone Falls, and only a few scattered routes were observed above the Khone Falls (see figures in pages 59 and 90, Poulsen et al., 2004). Several fish species have shown the ability to bypass the Khone Falls (Baird et al., 2004), and spawned individuals of *P. larnaudii* have been discovered, and their spawning ground was said to be right above the Khone Falls (Poulsen et al., 2004). Other biological characteristics of these species are presented in Table 1.

To develop effective strategies for resource conservation and fishery management, understanding geographic variation and movement patterns is crucial as they can reflect biological population structure and admixture. Population structures are influenced by various factors, including specific biological characteristics (such as egg and larval dispersal, spawning behaviour, habitats, and migration pathways), environmental parameters (pollution and ecological reservoirs), and physical features (like barriers and geographical distances) (Nguyen and Sunnucks, 2012). Spatially limited gene flow (Wright, 1943) can result in population genetic variation patterns due to isolation by distance (IBD). Current biological populations are experiencing local selection in response to changing environmental conditions, potentially leading to the emergence of genetically distinct sub-populations and separate stocks. Genomic technologies, including high-throughput sequencing and polymorphic molecular markers (e.g., SNPs), are widely utilised to investigate the level of genetic differentiation, thereby exploring population connectivity and viability in increasingly fragmented environmental habitats (Hendricks et al., 2018).

Given the vast expanse of the Mekong River Basin and the presence of ecological and physical barriers to fish migration within the basin, different patterns of subpopulations corresponding to geographic distance could likely occur. The correlation between population structure and IBD pattern has been studied not only in Mekong fish species with different life histories (such as the long-migratory *Henicorhynchus lobatus* (Hurwood et al., 2007); *Cirrhinus molitorella* (Nguyen and Sunnucks, 2012)), short-migratory *Henicorhynchus siamensis* (Adamson et al., 2009), and non-migratory *Mastacembelus favus* (Jamaluddin et al., 2019); *Channa striata* (Duong et al., 2019)), but also using various molecular markers (e.g., mitochondrial DNA (Hurwood et al., 2007; Adamson et al., 2009; Jamaluddin et al., 2019; Duong et al., 2019), microsatellites (Nguyen and Sunnucks, 2012)). In long-migratory fish, high population connectivity consistent with a non-significant IBD was confirmed (Hurwood et al., 2007; Nguyen and Sunnucks, 2012); in contrast, multiple genetic populations with a strong IBD pattern were reported in short and non-migratory fish (Adamson et al., 2009; Jamaluddin et al., 2019; Duong et al., 2019). The resulting research question was what is the correlation between the IBD and fish migratory patterns to the population connectivity?

This study aims to assess the level of spatial genetic differentiation of target fish species representing three typical migration patterns in the LMB and test the correlation of IBD to the level population connectivity.

Materials and Methods

Ethical approval

All the fish samples utilised in this study were collected from local fishermen or market vendors. As food fishes, they were dead at the time of sampling; therefore, no ethical approval was required for sampling.

Sample collection

Three fish species (*M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*) were field-identified and collected from Luang Prabang (Northern Lao PDR) to the Mekong Delta (Vietnam), spanning about 2,100 km. Our sampling strategy included mainstem sites (Paksan, Pakse - Lao PDR, Kratie - Cambodia, and Dong Thap, An Giang - Vietnamese Mekong Delta), tributaries (Khan River - Luang Prabang, Lao PDR, Mun River - Ubon Ratchathani, Thailand, and Chi River - Roi Et, Thailand); confluence of Mekong and 3S rivers (Stung Treng - Cambodia), and Tonlé Sap Lake (Siem Reap - Cambodia) (Fig. 1).

Table 1. Summary of biological characteristics for three fish species (Fish images were taken in this study).

Species	Fish guild and lifestyle	Egg and larval duration (Days)	Life span (Years)	Age at first maturity (year)
 *Macrognathus siamensis*	A "black" fish found at bottom depths in slow-moving and still-water bodies such as swamps, canals, streams, ponds, reservoirs, paddy fields, and floodplains	22[a]	8–18*	2[a]
 *Labeo chrysophekadion*	A "white" fish occurs in rivers, streams, canals, and floodplains	30–50[b]	28.7[c]	6.2[c]
 *Pangasius larnaudii*	A "white" fish found in medium and large-sized rivers, floodplains, and deep pools	33–35[b]	2.7[c]	0[c]

[a]Saowakoon and Saowakoon (2007).
[b]Hortle et al. (2015).
[c]Baran (2006).
*Non-official information (https://www.fishlaboratory.com/fish/peacock-eel/)

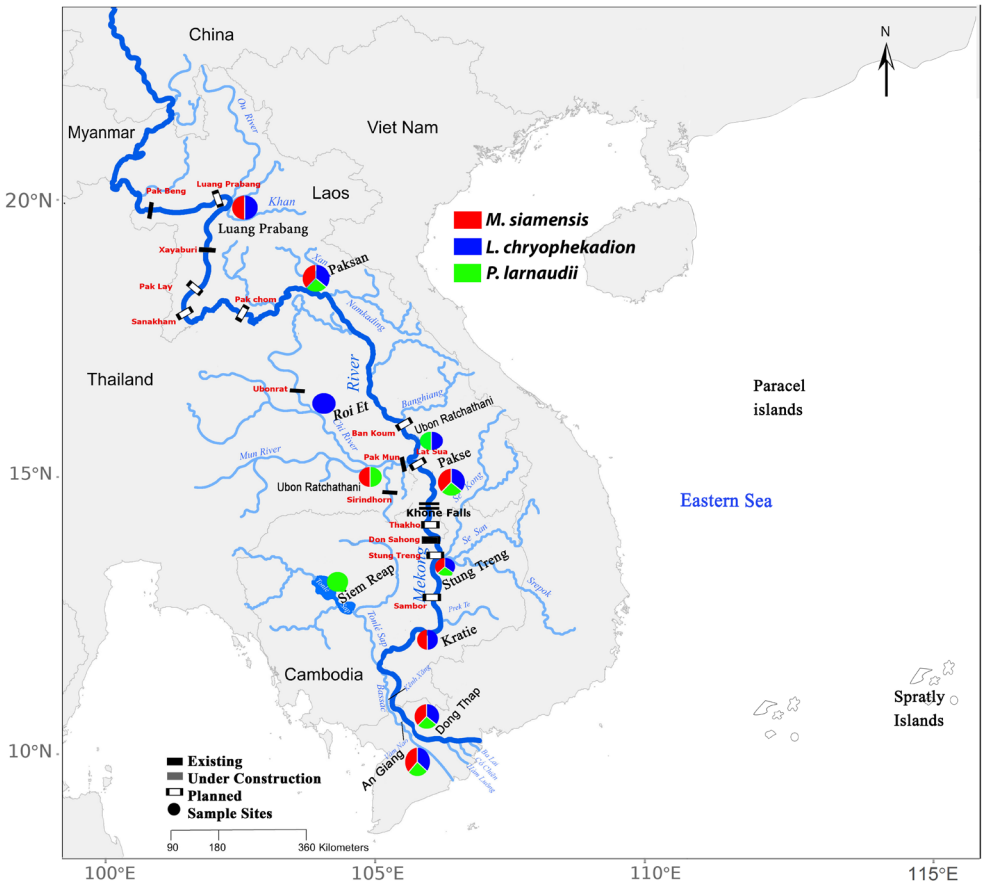


Fig. 1. Map of sampling locations for *Macrognathus siamensis* (red colour), *Labeo chrysophekadion* (blue), and *Pangasius larnaudii* (green) across the Lower Mekong River Basin. "═" indicates the location of Khone Falls.

Approximately 200 fish individuals from 7–9 populations of three fish species (*M. siamensis*, n = 240 individuals; *L. chrysophekadion*, n = 255; and *P. larnaudii*, n = 192) were sampled (Table 2). Muscle tissue (~50 mg) from each individual was taken and placed in collection tubes containing 95 % molecular-grade ethanol and then transported to the Molecular Biology Laboratory at Nha Trang University, Vietnam.

DNA extraction

Genomic DNA was extracted from preserved tissue samples using the Wizard® SV Genomic DNA Purification System kit (Promega, USA) following the manufacturer's instructions. A minor modification was made in the elution step; the extracted DNA was eluted three separate times with 100 µL of AE buffer used each time instead of 250 µL of nuclease-free water. All three elutions for each sample were inspected by gel electrophoresis (1.5 % agarose), and the elution that contained high-quality genomic DNA with the least amount of smaller, degraded DNA fragments was selected. The concentrations of the best elution for each sample were measured using a Qubit® 2.0 fluorometer (Thermo Fisher Scientific, USA) with dsDNA HS kit (Invitrogen, USA). Selected DNA samples (100 ng, ≥3 ng.µL^{-1}) were bound to AMPureXP beads (Beckman Coulter, USA) of 2:1 template/bead ratio, and then ethanol contaminants removed. Bead-water elutions (21.5 µL) were used for library preparation.

EzRAD library preparation and sequencing

Library preparation followed the previous ezRAD protocol (Toonen et al., 2013; Dang et al., 2019). The implementation process included randomly fragmenting the genomic DNA with the isoschizomeric restriction enzymes *MboI* and *Sau3AI* (New England Biolabs, USA) followed by performing end-repair, fragment size selection, poly-A tailing, ligation of dual-indexed Illumina adapters, and PCR amplification using the Illumina TruSeq Nano DNA library Prep kit. The libraries were sent to the Genomics Core Laboratory (Texas A&M University, Corpus Christi, USA) for paired-end 150 bp sequencing on the Illumina HiSeq 4000 platform.

SNP and outlier loci detection

Data processing, including sequence quality trimming, *de novo* reference assembly, mapping, and variant calling, was carried out using dDocentHPC pipeline (https://github.com/cbirdlab/dDocentHPC, Bird (2023)), a modified version of the Docent pipeline (Puritz et al., 2014). For more details on the assembly, mapping, and genotyping refer to Dang et al. (2019). The variant call format (VCF) file resulting from genotyping was filtered using the fltrVCF script (https://github.com/cbirdlab/fltrVCF, Bird and Selwyn (2023)), following Dang et al. (2019) and Biesack et al.

Table 2. Number of fish individuals collected in the Lower Mekong Basin.

Country	Sampling sites (population code)	LMB location	Geographic coordinates		Number of individuals		
					Macrognathus siamensis	*Labeo chrysophekadion*	*Pangasius larnaudii*
Lao PDR	Luang Prabang (LP)	Khan river	19°53'39.2"N	102°08'28.6"E	32	22	-
	Paksan (PA)	Mainstem	18°21'11.3"N	103°57'21.7"E	34	32	21
	Pakse (PE)	Mainstem	15°07'30.0"N	105°48'47.8"E	27	32	-
Thailand	Ubon Ratchathani (UB-MK)	Mainstem	15°18'48.2"N	105°29'52.6"E	-	28	29
	Ubon Ratchathani (UB-MR)	Mun River	15°13'26.0"N	104°51'09.7"E	32	-	24
	Roi Et (RE)	Chi River	15°57'39.8"N	103°59'31.5"E	-	30	-
Cambodia	Stung Treng (ST)	Confluence	13°31'46.7"N	105°57'06.6"E	32	28	32
	Siem Reap (SR)	Tonlé Sap Lake	13°0'N	104°3'E	-	-	30
	Kratié (KT)	Mainstem	12°49'31"N	106°01'71.5"E	28	27	-
Vietnam	Dong Thap (DT)	Mainstem	10°57'04.1"N	105°38'11.0"E	25	32	32
	An Giang (AG)	Mainstem	11°07'37.7"N	105°10'50.8"E	30	24	24
Total					240	255	192

(2023) with modifications. Briefly, samples were filtered for biallelic SNP loci, base call quality score (≥40), mean depth of coverage (<10×), minor allele count (<3), minor allele frequency (>0.05), missing data (≤0.8), allele balance (0.25–0.75), and Hardy-Weinberg equilibrium. Finally, one SNP from each contig was selected to obtain the validated SNP panel.

Outlier loci potentially under selection were identified using two different methods as implemented in LOSITAN (Antao et al., 2008) and BayeScan v2.1 (Foll and Gaggiotti, 2008). LOSITAN was run using the infinite alleles mutation model with parameter settings for "neutral", 500,000 simulations, and a confidence interval of 0.95. BayeScan was run with default parameters, and a false discovery rate (FDR) correction of 0.1 %. Any outliers identified by both methods were removed from the dataset to generate a panel of neutral SNPs. An additional dataset of putatively divergent loci shared among the two methods was used for further analysis.

Population genetic differentiation

All analyses were performed on both putatively neutral and divergent datasets.

To determine the overall population genetic differences of the studied fish species, GenAlEx v6.5.2 (Peakall and Smouse, 2012) was applied to calculate G_{ST} values (an analogue of F_{ST}, adjusted for bias) based on the genetic differences of all SNP loci. To measure the degree of genetic differentiation between pairs of populations based on allele frequencies, pairwise F_{ST} values were estimated using ARLEQUIN v3.5 (Excoffier and Lischer, 2010). The significance of these values was calculated using 1,000 bootstrap replicates of the data, and all *P*-values underwent FDR correction to avoid false positives resulting from multiple comparisons (Benjamini and Hochberg, 1995). Heatmaps of pairwise F_{ST} values between sample populations were visualised using seaborn v0.13.2 (Waskom, 2021).

Isolation by distance

To test for correlations between genetic differentiation and geographic distance, analysis of isolation by distance (IBD) was conducted by comparing the matrices of genetic distance log10(F_{ST}/(1-F_{ST})) versus geographic distance. Geographic distance was estimated in kilometres by river distance using the shortest water path among populations, determined with Google Maps. The relationship between matrices was assessed by the Mantel test with the 'vegan' function (Oksanen et al., 2016) implemented in R. The statistical significance of the parameter estimates was obtained via 9,999 permutations and 1,000 bootstraps resamplings.

As the Mantel test does not have strong power to detect correlation (Legendre et al., 2015), the alternative distance-based Moran's eigenvector maps (dbMEM) analysis was performed using the function 'dbmem' in the R package "adespatial" (Dray et al., 2020). dbMEM analysis was performed following these steps: (1) construct the Euclidean distance matrix among all populations; (2) determine a truncation threshold, *thresh*, the maximum value of the minimum spanning tree of the Euclidean distance matrix; (3) modify the distance matrix by changing all distances larger than the truncation distance to 4**thresh*, as advised in Borcard and Legendre (2002); (4) compute the principal coordinates of the modified matrix; (5) The dbMEMs having modelled positive spatial correlation (i.e., Moran's I larger than the expected value of I) based on the eigenfunctions were retained as spatial variables. The forward selection method was applied to identify significant dbMEMs, which were tested using 9,999 permutations. Additionally, adjusted R^2 values were calculated using the mem.select function in default format (only MEMs associated to positive eigenvalue are considered).

Results

RAD sequencing, SNP genotyping and outlier detecting

Following the analytical pipelines, 356,972; 103,408; and 59,769 raw SNPs were genotyped for *M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*, respectively. After filtering, 1,517 SNPs were retained for *M. siamensis*, 825 SNPs for *L. chrysophekadion*, and 1,270 SNPs for *P. larnaudii*. Information on individuals removed and SNPs retained at each step of filtering and data analysis is presented in Supplementary Table 1.

Using LOSITAN and BayeScan (Supplementary Fig. 1), a total of 273, 65, and 94 loci were identified as outlier loci and removed to produce a panel of 1244, 760, and 1176 neutral SNPs in *M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*, respectively. In addition, putatively divergent loci shared by two the methods (186, 23, and 11 loci) were used for further analysis.

Population genetic differentiation

When the neutral loci dataset was applied, overall population genetic difference was highest in *M. siamensis* (G_{ST} = 0.303, F_{ST} range from 0.003 to 0.477), lowest in *P. larnaudii* (G_{ST} = 0.003; F_{ST} from <0.001 to 0.088), and moderate in *L. chrysophekadion* (G_{ST} = 0.038, F_{ST} from 0.007 to 0.150). Meanwhile, *L. chrysophekadion* populations had the lowest value (G_{ST} = 0.023, F_{ST} from <0.001 to 0.16), followed by *P. larnaudii* (G_{ST} = 0.237, F_{ST} from 0.019 to 0.477), and *M. siamensis* (G_{ST} = 0.805, F_{ST} from 0.01 to 0.963) using divergent loci. Highly significant support was observed in all G-statistic values ($P < 0.001$).

Heatmaps showing the degree of spatial genetic differentiation based on neutral and divergent loci were presented in Figure 2 and Supplementary Figure 2. While *L. chrysophekadion* displayed all pairwise significant F_{ST} values (Fig. 2B), one non-significant F_{ST}

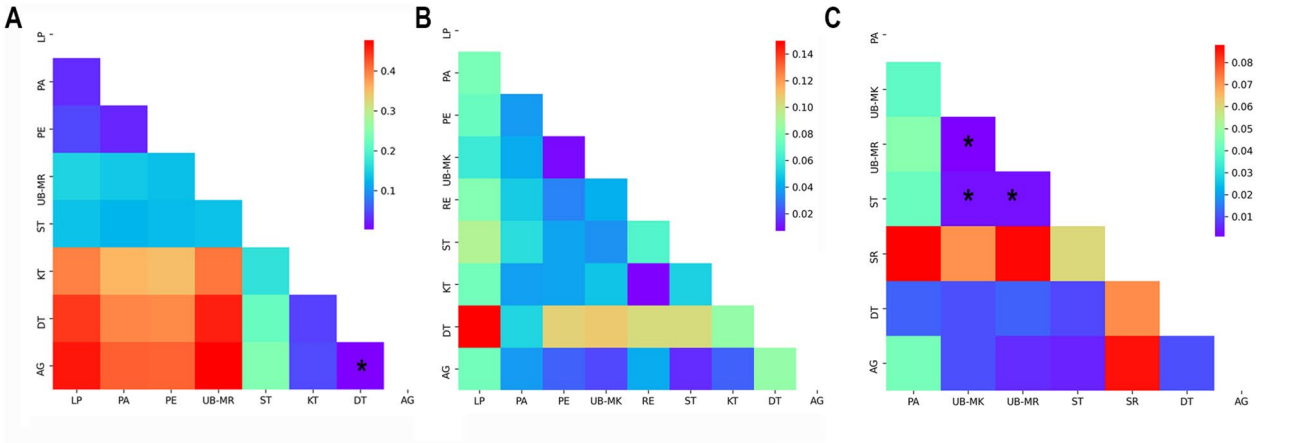


Fig. 2. Pairwise F_{ST} heatmaps of *Macrognathus siamensis* (A), *Labeo chrysophekadion* (B), *Pangasius larnaudii* (C) using neutral dataset. The colour intensity represents the magnitude of F_{ST}, with warmer colours (e.g., red, orange) indicating higher differentiation and cooler colours (e.g., blue, purple) indicating lower differentiation, according to the respective colour scales. Asterisks (*) indicate non-significant difference after FDR correction.

was observed between DT and AG in *M. siamensis* (Fig. 2A). In *P. larnaudii*, population admixture occurred between the Mun tributary (UB-MR) and 3S confluent (ST) to Mekong mainstem (UB-MK) ($P > 0.05$) (Fig. 2C).

When considering genetic differences of each species, *M. siamensis* populations located below the Khone Falls (KT, ST, DT, and AG) exhibited strong genetic differentiation to those (LP, PA, PE, and UB-MR) above the falls (Fig. 2A). For *L. chrysophekadion*, the greatest degrees of differentiation were detected between Dong Thap (F_{ST} range from 0.052–0.15) and Luang Prabang (F_{ST} from 0.06–0.15) to the remaining populations (Fig. 2B). The population from Tonlé Sap Lake (SR) displayed higher differentiation (F_{ST} range from 0.06–0.088), followed by the PA (F_{ST} from 0.012–0.088) when compared to other populations in *P. larnaudii* (Fig. 2C).

Isolation by distance

Based on the neutral dataset, a significant pattern of IBD was detected in *M. siamensis* (adjusted $R^2 = 0.2487$, $P = 0.004$; Fig. 3A), and *L. chrysophekadion* (adjusted $R^2 = 0.1627$, $P = 0.008$; Fig. 3B) using Mantel tests, whereas no significant pattern was observed in *P. larnaudii* (adjusted $R^2 = -0.0163$, $P = 0.4201$; Fig. 3C). dbMEM analysis showed the spatial structure for all three species, but only at MEM-1 (Obs. = 0.7379, 0.1449 and 0.1776, respectively, all $P = 0.01$) (Figs. 4A, B, and C). Result from adjusted $R^2 = 0.742$ in *M. siamensis* ($P = 0.001$) was concordant with strong IBD signals observed in MEM-1. However, in *L. chrysophekadion*, a modest IBD pattern was also revealed in MEM-1 (adjusted $R^2 = 0.826$, $P = 0.012$). For other MEMs, no positive spatial genetics were detected ($P > 0.05$) (Supplementary Table 2).

In the divergent loci dataset, a clear relationship between pairwise F_{ST} and geographic distance was shown in *M. siamensis* (adjusted $R^2 = 0.353$, $P = 0.0005$), while no significant correlation was observed in *L. chrysophekadion* and *P. larnaudii* (adjusted $R^2 = -0.0134$, $P = 0.4764$, and adjusted $R^2 = -0.009$, $P = 0.3804$, respectively) (Fig. 3). dbMEM analysis showed similar patterns to neutral loci, with Obs./adjusted R^2 = 0.632/0.399, 0.145/0.437, and 0.535/0.616 in *M. siamensis*, *L. chrysophekadion* and *P. larnaudii*, respectively; all $P < 0.05$) (Figs. 4D, E, and F).

Discussion

The genetic differentiation indexes and distance-based analyses supported spatial population structure and limited gene flow for the three fish species studied across their distribution ranges in the Lower Mekong Basin (LMB). An insignificant IBD pattern was observed only in *P. larnaudii*, suggesting that both dispersal and migratory distances might play crucial roles. The admixture populations found from the Ubon Ratchathani watershed to the Mun tributary and the 3S confluent likely indicate the lateral migratory capabilities of this species.

Larger genetic differences were detected between populations below and above the Khone Falls in *M. siamensis*, a phenomenon commonly observed in several short- and non-migratory fish species based on mitochondrial loci (Adamson et al., 2009; Jamaluddin et al., 2019) and microsatellite markers (Nguyen and Sunnucks, 2012). The limited upstream migration due to the physical barrier of the Khone Falls (Poulsen and Valbo-Jørgensen, 2001), along with the adaptive strategy of dispersing eggs and larvae to downstream floodplains at the beginning of the rainy season, likely explains these differences (Hughes, 2024).

In the southernmost Mekong Delta floodplain in Vietnam, the Tien River, where the DT population was sampled, experiences high water speeds of up to 25,500 $m^3.s^{-1}$ during the flood season, which is four times that of the Hau River (Vo, 2012). This significant difference could, in some cases, affect the larval

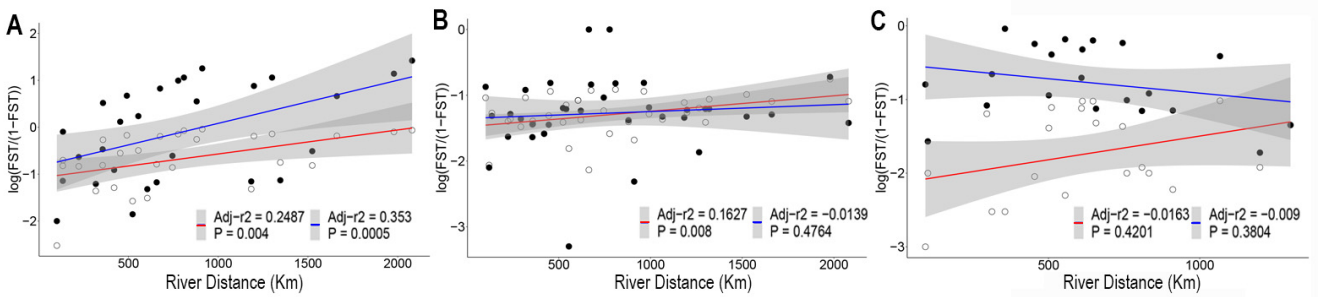


Fig. 3. The Mantel test plotting of pairwise log10(F_{ST}/1-F_{ST}) against pairwise geographic distances of *Macrognathus siamensis* (A), *Labeo chrysophekadion* (B), *Pangasius larnaudii* (C) based on neutral (red line) and divergent loci (blue line) datasets.

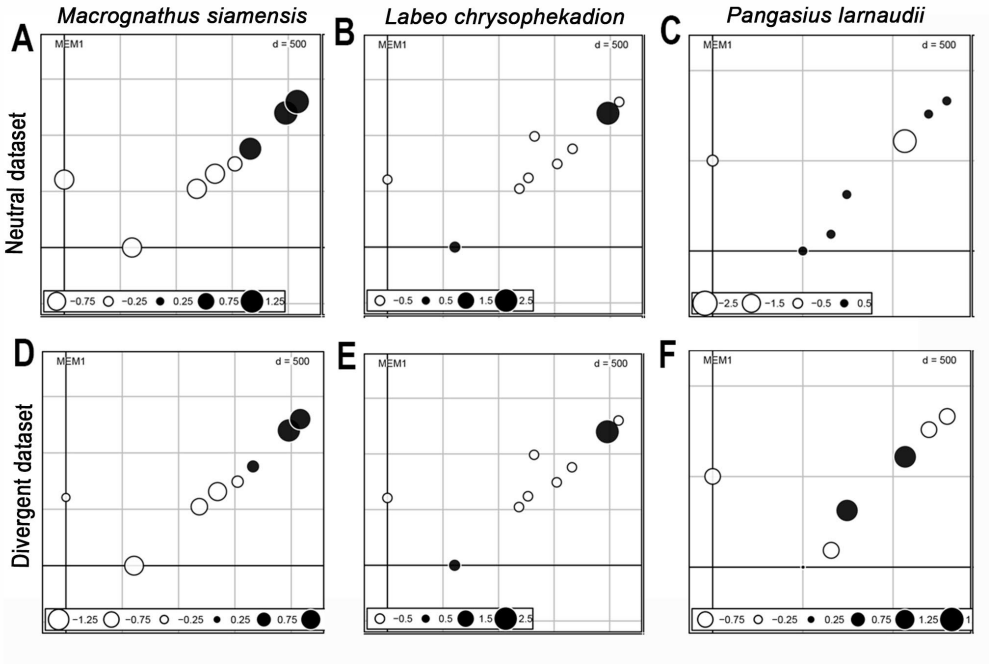


Fig. 4. The significant dbMEM variables (F_{ST} against pairwise geographic distances) using neutral (A, B, C) and divergent loci (D, E, F). The circles indicated the position and dbMEM values of each quadrat. Black dots represent positive values and white dots negative value.

dispersal (e.g., from Tonlé Sap), the feeding movements of juveniles along the river, and the upstream spawning of adult fish. *L. chrysophekadion* is known to respond to a wide range of water levels for migration (1,500 to 26,000 $m^3.s^{-1}$), and is most abundant at 1,500 to 5,000 $m^3.s^{-1}$ (Baran, 2006). These water level adaptations and its facultative migration could explain the genetic differences detected in the DT population.

Similar phenomena were also observed in the amphidromous *Polynemus melanochir* distributed in the Mekong Delta using SNP dataset (Dang et al., 2019). Besides DT, significant genetic differences in *P. larnaudii* were also found in the SR population. Our study is limited by the lack of an SR sampling site for both *M. siamensis* and *L. chrysophekadion*, which may render comparisons or interpretations misleading. Acting as a large and important natural floodwater reservoir for the Mekong system and serving as a major spawning site for Mekong fish (Poulsen et al., 2001), Tonlé Sap can simultaneously be considered a source and sink for fish populations. Formed around 6,000 years ago through geographical events of fragmentation and amalgamation (Okawara and Tsukawaki, 2002), the lake has significantly impacted the intraspecific diversity of fishes and aquatic fauna (So et al., 2006; Duong et al., 2019). Further studies on other fish species collected from this habitat are needed to better understand population connectivity and historical evolutionary processes.

The IBD patterns aligned with the migratory traits of the three fish species (non-, short-facultative, and long-migratory fishes, respectively). Previous genetic studies on Mekong fish using typically few specific mitochondrial or nuclear loci were also consistent with our results (Hurwood et al., 2007; Adamson et al., 2009; Nguyen and Sunnucks, 2012; Duong et al., 2019; Jamaluddin et al., 2019). Although the three fish species exhibited almost similar dispersal times (Table

1), factors such as maturity age, adult size, generation spans, and others (hatching time, egg and larval size) can influence dispersal ability and distance (Bradbury and Benzen, 2007). In fact, non-linear IBD patterns (IBD slopes, intercepts, and R^2 value) and their relationship with the dispersal factors have been demonstrated in many marine fish species (Bradbury and Benzen, 2007), which also need to be further tested on freshwater fish.

Meanwhile, dbMEM analysis (at MEM-1) supported a positive correlation between spatial structure and distance for all three species, with *L. chrysophekadion* showing the strongest pattern. This variation could be attributed to the Mantel test primarily explaining linear genetic variation along river distance, potentially overlooking the spatial structure of riverine networks. Previous studies that integrated these two methods have reported either consistent (Jensen et al., 2019; Maduna et al., 2023) or conflicting (Graham et al., 2022) results. Hence, relying solely on isolation by distance is inadequate to explain the spatial genetic differentiation among populations.

Based on local ecological knowledge data, the hypothesised population structures (multiple populations above and below the Khone Falls of *L. chrysophekadion*, and a homogenised population from above Khone Falls to downstream floodplains for *P. larnaudii*) (Poulsen et al., 2004), partly reflect the migratory patterns (facultative migration and bypass Khone Falls for spawning, respectively), and are further reinforced based on genetic differentiation indexes and comparative IBD models. Microsatellite data, however, suggested population connectivity of *L. chrysophekadion* between the Vietnamese Mekong Delta and Paksan (Lao PDR) (Mashyaka and Duong, 2021). Meanwhile, the subdivision of populations above and below the Khone Falls, as well as those in tributaries such as the Mun River (Ubon Ratchathani) and the Khan River (Luang Prabang), compared to the mainstem has been strongly supported in *M. siamensis* (Truong et al., 2025), and *L. chrysophekadion* (unpublished data) using SNP datasets. Further analyses, such as isolation by resistance, isolation by distance using least-cost path distance, and cluster and demographic analyses, need to be conducted to better explain the population connectivity and sub-population structures of Mekong fish species.

Conclusion

This study reveals significant genetic differentiation and varying IBD signals corresponding to different migratory patterns of three fish species (*M. siamensis*, *L. chrysophekadion*, and *P. larnaudii*) in the LMB, highlighting the crucial role of dispersal capability and migration distance of these species. The highest genetic differentiation in *M. siamensis* populations, especially above and below the Khone Falls, reflects the significant impact of physical barriers on non-migratory species. This pattern is supported by the strong IBD detected through Mantel tests and dbMEM analysis. *L. chrysophekadion* shows moderate genetic differentiation with a significant IBD pattern and positive spatial structure at MEM-1, indicating the influence of its facultative migration. In contrast, *P. larnaudii* presents low genetic differentiation and no significant IBD, reflecting its high connectivity and extensive migratory behaviour. Future research should expand on sampling sites and conduct detailed ecological assessments to inform sustainable management practices and conservation efforts in the Mekong River Basin.

Acknowledgements

We would like to thank Mekong project partners who helped us collect tissues from fish markets, and team members of Biodiversity and Conservation, Institute for Biotechnology and Environment, Nha Trang University for project support.

This study was funded by the United States Agency for International Development supported Partnerships for Enhanced Research Project 6-435 under USAID Cooperative Agreement AID-OAA-A-11-00012. PhD of Truong Thi Oanh was funded by Vingroup Joint Stock Company and supported by the Domestic Master/PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Vingroup Big Data Institute (VINBIGDATA), code VINIF.2020.TS.34, VINIF.2021.TS091, and VINIF.2022.TS.091.

Conflict of interest: The authors declare that they have no conflict of interest.

Author contributions: Oanh Thi Truong: Conceptualisation, methodology, investigation, formal analysis, visualisation, writing – original draft. Sang Quang Tran: Formal analysis, writing – review and editing. Quyen Vu Dang Ha: Methodology, investigation, writing – review and editing; Van Ngo Thai Bich: Supervision, methodology, writing – review and editing. Binh Thuy Dang: Supervision, project administration, conceptualisation, methodology, visualisation, writing – review and editing.

References

Adamson, E.A.S., Hurwood, D.A., Baker, A.M., Mather, P.B. 2009. Population subdivision in Siamese mud carp *Henicorhynchus siamensis* in the Mekong River basin: implications for management. Journal of Fish Biology 75:1371–1392. https://doi.org/10.1111/j.1095-8649.2009.02369.x

Adamson, P.T., Rutherfurd, I.D., Peel, M.C., Conlan, I.A. 2009. The hydrology of the Mekong River. In: The Mekong biophysical environment of an international river basin, Campbell, I.C. (Ed.). Academic Press, San Diego,USA, pp. 53–76.

Antao, T., Lopes, A., Lopes, R.J., Beja-Pereira, A., Luikart, G. 2008. LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323. https://doi.org/10.1186/1471-2105-9-323

Baird, I.G., Flaherty, M.S., Phylavanh, B. 2004. Mekong River Pangasiidae

catfish migrations and the Khone Falls wing trap fishery in southern Laos. Natural History Bulletin of the Siam Society 52:81–109.

Baran, E. 2006. Fish migration triggers in the Lower Mekong Basin and other tropical freshwater systems. MRC Technical Paper 14. 56 pp.

Baran, E., Baird, I.G., Cans, G. 2005. Fisheries bioecology at the Khone Falls (Mekong River, Southern Laos). WorldFish Center, Penang, Malaysia. 84 pp.

Benjamini, Y., Hochberg, Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Biesack, E.E., Dang, B.T., Ackiss, A.S., Bird, C.E., Chheng, P., Phounvisouk, L., Truong, O.T., Carpenter, K.E. 2020. Evidence for population genetic structure in two exploited Mekong River fishes across a natural riverine barrier. Journal of Fish Biology 97:696–707. https://doi.org/10.1111/jfb.14424

Bird, C.E. 2023. dDocentHPC. A pipeline to trim, assemble, map, and genotype reduced representation genomic data. https://github.com/cbirdlab/dDocentHPC

Bird, C.E., Selwyn, J.S. 2023. fltrVCF v4.4: A script to filter VCF files and simplify reporting and reproduction of filters. https://github.com/cbirdlab/fltrVCF

Borcard, D., Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51–68. https://doi.org/10.1016/S0304-3800(01)00501-4

Bradbury, R.I., Benzen, P. 2007. Non-linear genetic isolation by distance: implications for dispersal estimation in anadromous and marine fish populations. Marine Ecology Progress Series 340:245–257. https://doi.org/10.3354/meps340245

Dang, B.T., Vu, Q.H.D., Biesack, E.E., Doan, T. V, Truong, O.T., Tran, T.L., Ackiss, A.S., Stockwell, B.L., Carpenter, K.E. 2019. Population genomics of the peripheral freshwater fish *Polynemus melanochir* (Perciformes, Polynemidae) in a changing Mekong Delta. Conservation Genetics 20:961–972. https://doi.org/10.1007/s10592-019-01189-x

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guénard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N., Wagner, H.H. 2020. Adespatial: Multivariate multiscale spatial analysis. R package version 0.3-3. https://cran.r-project.org/package=adespatial

Duong, T.Y., Uy, S., Chheng, P., So, N., Tran, T.H.T., Thi, N.T., Pomeroy, R., Egna, H. 2019. Genetic diversity and structure of striped snakehead (*Channa striata*) in the Lower Mekong Basin: Implications for aquaculture and fisheries management. Fisheries Research 218:166–173. https://doi.org/10.1016/j.fishres.2019.05.014

Excoffier, L., Lischer, H.E.L. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Foll, M., Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180:977–993. https://doi.org/10.1534/genetics.108.092221

Graham, C.F., Boreham, D.R., Manzon, R.G., Wilson, J.Y., Somers, C.M. 2022. Population structure of lake whitefish (*Coregonus clupeaformis*) from the Mississippian lineage in North America. Facets 7:853–874. https://doi.org/10.1139/FACETS-2021-0191

Grill, G., Ouellet Dallaire, C., Fluet Chouinard, E., Sindorf, N., Lehner, B. 2014. Development of new indicators to evaluate river fragmentation and flow regulation at large scales: A case study for the Mekong River basin. Ecological Indicators 45:148–159. https://doi.org/10.1016/j.ecolind.2014.03.026

Hendricks, S., Anderson, E.C., Antao, T., Bernatchez, L., Forester, B.R., Garner, B., Hand, B.K., Hohenlohe, P.A., Kardos, M., Koop, B., Sethuraman, A., Waples, R.S., Luikart, G. 2018. Recent advances in conservation and population genomics data analysis. Evolutionary Applications 11:1197–1211. https://doi.org/10.1111/eva.12659

Hogan, Z., Baird, I.G., Radtke, R., Vander Zanden, M.J. 2007. Long distance migration and marine habitation in the tropical Asian catfish, *Pangasius krempfi*. Journal of Fish Biology 71:818–832. https://doi.org/10.1111/j.1095-8649.2007.01549.x

Hortle, K.G., Khounsavanh, O., Chanthasone, P., Phommachanh, P., Singhanouvong, D., Viravong, S. 2015. Larval and juvenile fish in the Mekong River in Northern Lao PDR, MRC Technical Paper. MRC Technical Paper No.46. Mekong River Commission, Phnom Penh, Cambodia. 87 pp.

Hughes, K. 2024. The Mekong's forgotten fishes and the emergency recovery plan to save them, WWF, Gland, Switzerland. https://wwfint.awsassets.panda.org/downloads/final-mekong-forgotten-fishes-report.pdf

Hurwood, D.A., Adamson, E.A.S., Mather, P.B. 2007. Evidence for strong genetic structure in a regionally important, highly vagile cyprinid (*Henicorhynchus lobatus*) in the Mekong River Basin. Ecology of Freshwater Fish 17:273–283. https://doi.org/10.1111/j.1600-0633.2007.00278.x

Jamaluddin, J.A.F., So, N., Tam, B.M., Ahmad, A., Grudpan, C., Page, L.M., Khaironizam, M.Z., Mohd Nor, S.A. 2019. Genetic variation, demographic history and phylogeography of tire track eel, *Mastacembelus favus* (Synbranchiformes: Mastacembelidae) in Southeast Asia. Hydrobiologia 838:163–182. https://doi.org/10.1007/s10750-019-03987-3

Jensen, A.M., O'Neil, N.P., Iwaniuk, A.N., Burg, T.M. 2019. Landscape effects on the contemporary genetic structure of ruffed grouse (*Bonasa umbellus*) populations. Ecology and Evolution 9:5572–5592. https://doi.org/10.1002/ece3.5112

Kang, B., Huang, X. 2021. Mekong fishes: Biogeography, migration, resources, threats, and conservation. Reviews in Fisheries Science and Aquaculture 30:170–194. https://doi.org/10.1080/23308249.2021.1906843

Legendre, P., Fortin, M., Borcard, D. 2015. Should the Mantel test be used in spatial analysis? Methods in Ecology and Evolution 6:1239–1247. https://doi.org/10.1111/2041-210X.12425

Maduna, S.N., Jónsdóttir, Ó.D.B., Imsland, A.K.D., Gíslason, D., Reynolds, P., Kapari, L., Hangstad, T.A., Meier, K., Hagen, S.B. 2023. Genomic signatures of local adaptation under high gene flow in lumpfish—Implications for broodstock provenance sourcing and larval production. Genes 14:1870. https://doi.org/10.3390/genes14101870

Mashyaka, A., Duong, T.-Y. 2021. Genetic diversity analysis revealed possible long migration of Black sharkminnow (*Labeo chrysophekadion*) along the Mekong River. Songklanakarin Journal of Science and Technology 43:955–960.

Nguyen, T.T.T., Sunnucks, P. 2012. Strong population genetic structure and its management implications in the mud carp *Cirrhinus molitorella*, an indigenous freshwater species subject to an aquaculture and culture-based fishery. Journal of Fish Biology 80:651–68. https://doi.org/10.1111/j.1095-8649.2011.03204.x

Okawara, M., Tsukawaki, S. 2002. Composition and provenance of clay minerals in the northern part of Lake Tonle Sap, Cambodia. Journal of Geography 111:341–359. https://doi.org/10.5026/jgeography.111.3_341

Oksanen, A., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P., Hara, R., Simpson, G., Solymos, P., Stevens, M., Szoecs, E. 2016. vegan: Community ecology package. R package version 2.4.3. https://cran.r-project.org/package=vegan

Peakall, R., Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel.

Population genetic software for teaching and research - An update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Poulsen, A.F., Hortle, K.G., Valbo-Jørgensen, J., Chan, S., Chhuon, C.K., Viravong, S., Bouakhamvongsa, K., Suntornratana, U., Yoorong, N., Nguyen, T.T., Tran, B.Q. 2004. Distribution and ecology of some important riverine fish species of the Mekong River Basin, MRC Technical Paper No. 10. MRC Technical Paper No. 10. https://archive.iwlearn.net/mrcmekong.org/download/free_download/Technical_paper10.pdf

Poulsen, A.F., Ouch, P., Sintavong, V., Ubolratana, S., Nguyen, T.T. 2002. Fish migrations of the Lower Mekong River Basin: Implications for development, planning and environmental management, MRC Technical Paper. MRC Technical Paper No. 8, Mekong River Commission, Phnom Penh. 62 pp.

Poulsen, A.F., Valbo-Jørgensen, J. 2001. Fish migrations and spawning habits in the Mekong mainstream – a survey using local knowledge. AMFC Technical Report, Mekong River Commission. 149 pp.

Puritz, J.B., Hollenbeck, C.M., Gold, J.R. 2014. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431. https://doi.org/10.7717/peerj.431

Rainboth, W.J. 1996. Fishes of the Cambodian Mekong. FAO species identification field guide for fishery purposes. FAO, Rome. 265 pp.

Saowakoon, H., Saowakoon, S. 2007. Breeding and nursing of spotted spiny eel (*Macrognathus siamensis*; Gunther, 1861). In: Proceedings of the 45th Kasetsart University Annual Conference: Fisheries Section, pp. 712–721.

So, N., Van Houdt, J.K.J., Volckaert, F.A.M. 2006. Genetic diversity and population history of the migratory catfishes *Pangasianodon hypophthalmus* and *Pangasius bocourti* in the Cambodian Mekong River. Fisheries Science 72:469–476. https://doi.org/10.1111/j.1444-2906.2006.01174.x

Toonen, R.J., Puritz, J.B., Forsman, Z.H., Whitney, J.L., Fernandez-Silva, I., Andrews, K.R., Bird, C.E. 2013. ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ 1:e203:1–15. https://doi.org/10.7717/peerj.203

Truong, O.T., Tran, S.Q., Carpenter, K.E., Vu, Q.D.H., Duong, T.-Y., Kyaw, M.M., Grudpan, C., Thai Bich, V.N., Dang, B.T. 2025. Population genetics of *Macrognathus siamensis* (Synbranchiformes: Mastacembelidae): Implications for non-migratory fishery resources in the Mekong River basin. Fisheries Research 281:107210. https://doi.org/https://doi.org/10.1016/j.fishres.2024.107210

Valbo-Jørgensen, J., Coates, D., Hortle, K. 2009. Fish diversity in the Mekong River Basin. In: The Mekong biophysical environment of an international river basin, Campbell, I.C. (Ed.), Elsevier, Amsterdam, pp. 161–196. https://doi.org/10.1016/B978-0-12-374026-7.00008-5

Vo, K.T. 2012. Hydrology and hydraulic infrastructure systems in the Mekong Delta, Vietnam. In: The Mekong Delta system, Renaud, F., Kuenzer, C. (Eds.), Springer Environmental Science and Engineering. Springer, Dordrecht, pp. 49–81. https://doi.org/10.1007/978-94-007-3962-8_3

Waskom, M. 2021. Seaborn: Statistical data visualization. Journal of Open Source Software 6:3021. https://doi.org/10.21105/joss.03021

Winemiller, K.O., Nam, S., Baird, I.G., Darwall, W., Lujan, N.K., Harrison, I., Stiassny, M.L.J., Silvano, R.A.M., Fitzgerald, D.B., Pelicice, F.M., Agostinho, A.A., Gomes, L.C., Albert, J.S., Baran, E., Jr, M.P., Zarfl, C., Mulligan, M., Sullivan, J.P., Arantes, C.C., Sousa, L.M., Koning, A.A., Hoeinghaus, D.J., Sabaj, M., Lundberg, J.G., Armbruster, J., Thieme, M.L., Petry, P., Zuanon, J., Vilara, G.T., Snoeks, J., Ou, C., Rainboth, W., Pavanelli, C.S., Akama, A., Soesbergen, A. Van, Sáenz, L. 2016. Balancing hydropower and biodiversity in the Amazon, Congo, and Mekong. Science 351:128–129. https://doi.org/10.1126/science.aac708

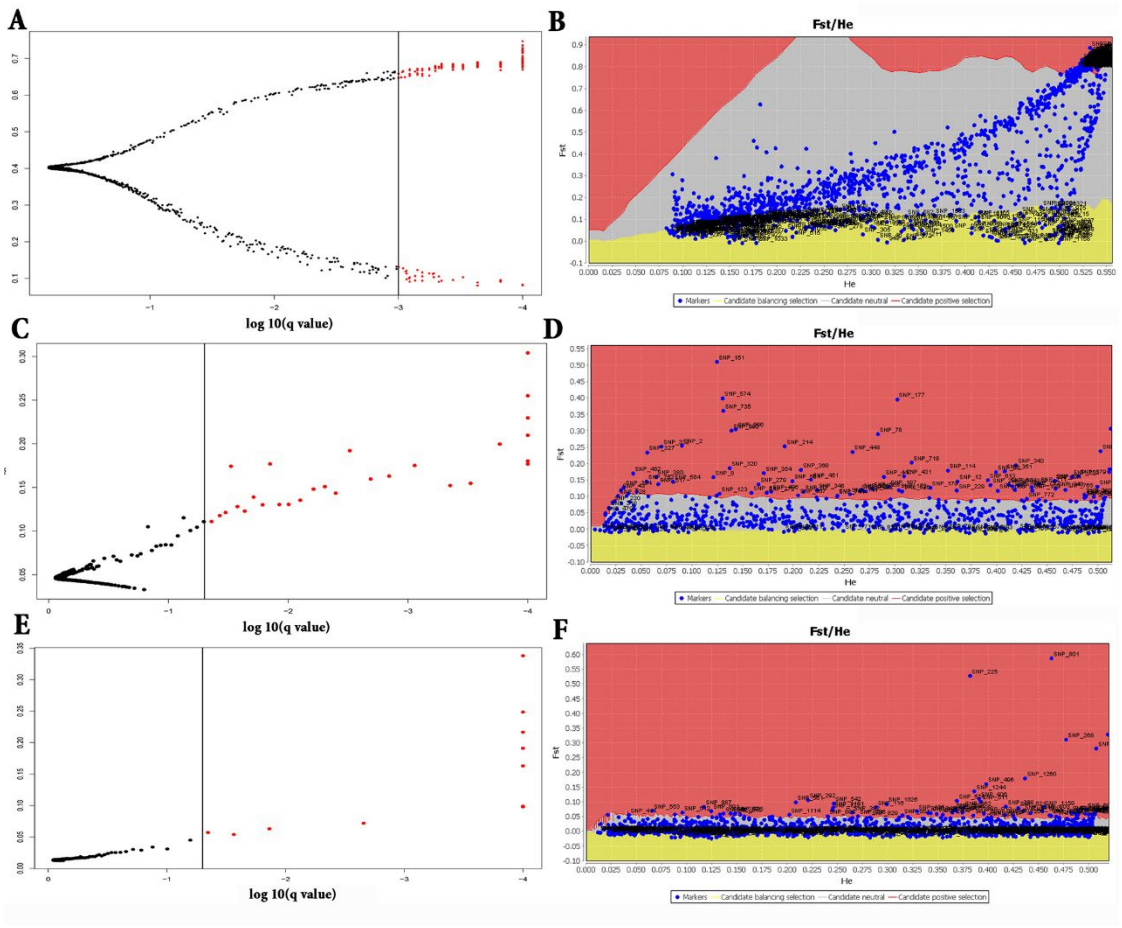
Wright, S. 1943. Isolation by distance. Genetics 28:114–138. https://doi.org/10.1093/genetics/28.2.114

Supplementary Table 1. The numbers of individuals removed and SNPs retained at the filtering steps of three fish species.

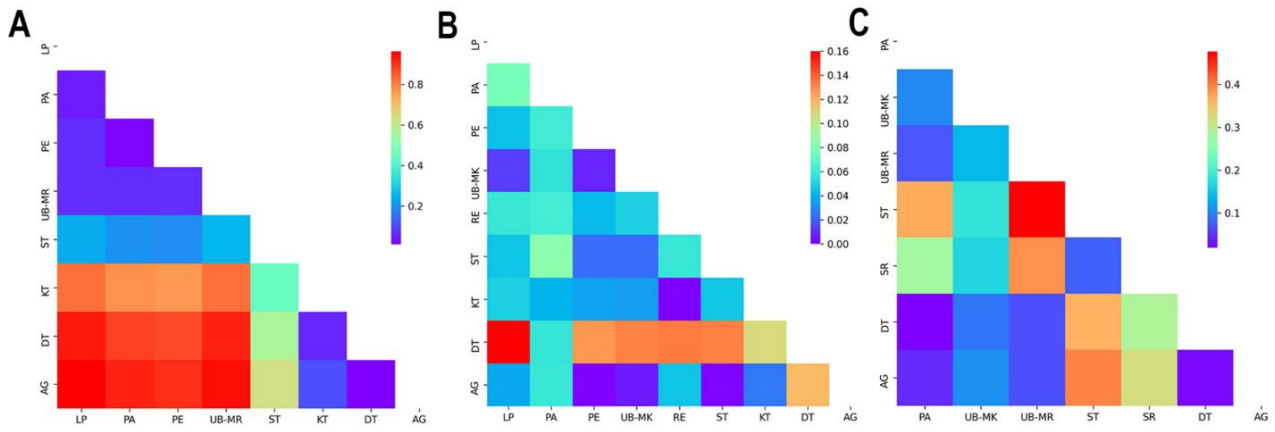
Filtering steps	*Macrognathus siamensis*		*Labeo chrysophekadion*		*Pangasius larnaudii*	
	No. of SNPs	No. of individuals	No. of SNPs	No. of individuals	No. of SNPs	No. of individuals
Raw SNPs	466,068	272	103,408	256	59,769	192
Biallelic loci	441,572	272	101,858	256	58,326	192
Remove Indel	392,771	272	89,307	256	48,526	192
MinQ ≥ 30	392,771	272	89,307	256	48,526	192
meandp ≤ 5	328,803	272	60,380	256	33,622	192
Max-missing (80%)	45,618	272	9,894	256	9,937	192
MAC	39,463	272	9,029	256	9,022	192
PAIRED	29,796	272	8,322	256	8,668	192
MAF 0.05	20,123	272	8,246	256	8,556	192
Max-missing 0.85	9,200	272	2,058	256	8,556	192
Missing-indv	9,200	239	2,058	232	1,917	160
HWE < 0.001	9,043	239	1,774	232	1,842	160
Filter_one_SNPs	4,237	239	825	232	1,270	160

Supplementary Table 2. Information of others dbMEM variables of three fish species.

dbMEMs	*Macrognathus siamensis*		*Labeo chrysophekadion*		*Pangasius larnaudii*	
	Observed	*P* value	Observed	*P* value	Observed	*P* value
Neutral dataset						
dbMEM-2	-0.083	0.334	-0.144	0.391	-0.220	0.653
dbMEM-3	-0.288	0.907	-0.156	0.796	-0.220	0.824
dbMEM-4	-0.290	0.923	-0.157	0.848	-0.220	0.838
dbMEM-5	-0.290	0.943	-0.157	0.878	-0.220	0.858
dbMEM-6	-0.291	0.953	-0.157	0.912	-0.294	1.000
dbMEM-7	-0.493	1.000	-0.157	0.900	--	--
dbMEM-8	--	--	-0.213	1.000	--	--
Divergent dataset						
dbMEM-2	0.455	0.130	-0.152	0.674	0.038	0.158
dbMEM-3	0.001	0.263	-0.156	0.781	-0.317	0.798
dbMEM-4	-0.240	0.599	-0.157	0.891	-0.323	0.873
dbMEM-5	-0.401	0.827	-0.157	0.911	-0.324	0.947
dbMEM-6	-0.705	0.999	-0.157	0.909	-0.607	1.000
dbMEM-7	-0.742	1.000	-0.157	0.860	--	--
dbMEM-8	--	--	-0.209	1.000	--	--



Supplementary Figure 1. Graphical representation of outlier detection for *Macrognathus siamensis* (A, B), *Labeo chrysophekadion* (C, D), and *Pangasius larnaudii* (E, F). (A, C, E) F_{ST} values are plotted against the log 10 of posterior odds (PO) for all loci in BayeScan; (B, D, F) Values are plotted against He (expected heterozygosity) in LOSITAN with different colours showing balancing, neutral, or positive patterns of selection (yellow, grey, and red colours, respectively).



Supplementary Figure 2. Heat map of pairwise F_{ST} value of *Macrognathus siamensis* (A), *Labeo chrysophekadion* (B), *Pangasius larnaudii* (C) using divergent dataset. The colour intensity represents the magnitude of FST, with warmer colours (e.g., red, orange) indicating higher differentiation and cooler colours (e.g., blue, purple) indicating lower differentiation, according to the respective colour scales.