BIODIVERSITAS
Volume 25, Number 11, November 2024
Pages: 4379-4388

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d251137

Genetic diversity of mud crab (*Scylla paramamosain*) in Vietnam based on *cox1* gene fragments

LE TAN PHAT[1,2,♥], HOANG TAN QUANG[3], NGUYEN VAN HUY[1], MAC NHU BINH[1], TON THAT CHAT[1,♥♥]

[1]Hue University of Agriculture and Forestry, Hue University. 102 Phung Hung Street, Hue City 49000, Thua Thien Hue, Vietnam. Tel.: +84-234-3522535, ♥email: ltphat.dhnl23@hueuni.edu.vn, ♥♥email: ttchat@hueuni.edu.vn
[2]Viet Uc Seafood Corporation. Phu My District 55300, Binh Dinh, Vietnam
[3]Institute of Biotechnology, Hue University. Nguyen Dinh Tu Street, Hue City 49000, Thua Thien Hue, Vietnam

Manuscript received: 15 August 2024. Revision accepted: 21 November 2024.

Abstract. *Phat LT, Quang HT, Huy NV, Binh MN, Chat TT. 2024. Genetic diversity of mud crab (*Scylla paramamosain*) in Vietnam based on* cox1 *gene fragments. Biodiversitas 25: 4379-4388.* Mud crab (*Scylla paramamosain* Estampador 1950) is a commonly found resource throughout Vietnam, with increasing exports each year. However, production is yet to be able to meet consumer demand, which continues to increase, leading to consistently high market prices. Accurately estimating genetic variation in mud crab populations is crucial for developing conservation management strategies for highly exploited fishery resources. In this study, genetic variations and population structure of *S. paramamosain* were examined by analyzing the nucleotide sequences of a 780 base pairs (bp) region of the mitochondrial *cox1* gene from three populations across thirteen collection sites in Vietnam, covering a diverse geographical range. A total of 45 individuals were examined, the *cox1* gene sequences were aligned, and 17 haplotypes were identified, with two of these haplotypes being shared by the populations. The haplotype diversity ranged from 0.57143 (Thua Thien Hue) to 0.80000 (Ca Mau). The haplotype network displayed that the haplotypes were divided into two clusters with Hap_3 as the center. Pairwise Fst values between populations ranged low, from 0-0.01266. The AMOVA results highlighted that within-population variation (99.85%) was higher than among-population variation (0.15%). Findings from neutral tests and mismatch analysis suggested implications for mud crab population dynamics. *S. paramamosain* in Vietnam showed high genetic diversity within the populations, as shown by low genetic variation and the significant gene flow between populations when analyzing the *cox1* gene. These results underscore the interconnectedness of the mud crab population and provide a foundation for establishing an enduring mud crab farming initiative in Vietnam.

Keywords: *cox1*, genetic variation, mud crab, population structure, *Scylla paramamosain*

INTRODUCTION

Mud crabs of the genus *Scylla* are ecologically and economically important crustacean species commonly found in the Indo-West Pacific region. They have become globally important aquaculture species and are very popular seafood in Southeast Asia (Chen and Wang 2019). In the last two decades, the global mud crab farming industry has experienced rapid development owing to its good flavor, high nutrition value, and rapid growth rate (Hongyu et al. 2016; Chen and Wang 2019. Despite the increasing consumer market demand for mud crabs, their production remains insufficient, leading to consistently high market prices (Chen and Wang 2019).

Mature mud crabs mate inshore under natural conditions. Gravid females then migrate offshore to spawn eggs, and the offspring return to their natural living environments. However, the population of mud crabs, including adults and larvae, has been rapidly decreasing due to seawater pollution and overexploitation (Hongyu et al. 2016; Shakawi et al. 2022). Mud crabs are also a model species for testing the evolutionary effects of sea level fluctuations on species from the East Asian seas (He et al. 2010). Previous research has indicated that the genus *Scylla* encompasses mud crabs that are typically confined to limited geographical regions. Mud crab (*Scylla paramamosain* Estampador 1950) species inhabit subtidal zones, while their early life stages and juvenile crabs can be found on the fringes of the mangrove ecosystem (Avianto et al. 2013). The distribution of this species is limited to the East China Sea and the surrounding areas of the South China Sea. *S. paramamosain* is the most commonly captured *Scylla* species in Vietnam, China, and Japan. However, it has also been discovered in reduced numbers along the southern coast of the South China Sea, throughout the Indonesian Seas, and along the eastern Gulf of Bengal (He et al. 2010).

In Vietnam, research on mud crab breeding began in the 1990s, leading to the emergence and development of commercial farms around 2005-2006 in some central provinces such as Khanh Hoa, Phu Yen, Binh Dinh, Thua Thien Hue, and later expanding strongly in the Mekong Delta region. In the Mekong Delta region, there are 480 breeding farms producing 0.5-12 million individuals/farm/year and 460 breeding stock centers producing 986,000 individuals/farm/year, especially in provinces like Ca Mau, Kien Giang, and Bac Lieu, which are the main green crab breeding areas. However, green mud crab breeding has mainly focused on increasing the number of farms and places less emphasis on improving breed quality. Practical production experience indicates that the size of cultured crabs at harvest tends to decrease gradually, with slow growth, low survival rates, and the prevalence of diseases.

The use of unselected parent crab sources increases the risk of inbreeding and genetic decline. These parent crab sources, largely collected from nature or cultured ponds without research investment, yield stable larval survival rates but need more control over improving seed quality.

Ca Mau (Mekong Delta), Binh Dinh (the South-Central Coast), and Thua Thien Hue (North Central Coast) are notable for their mud crab harvesting, which significantly contributes to the country's overall yield. Each region has distinct coastal ecosystems, such as mangrove forests, estuaries, and tidal flats, which affect the distribution and genetic traits of *S. paramosain*. By sampling from diverse ecosystems, researchers can gain a comprehensive understanding of the genetic diversity of the mud crabs. Although *S. paramosain* has significant economic value, knowledge of its genetic diversity still needs to be improved. Genetic diversity is the determinant of the allocation of germplasm resources, which are the genetic material used for breeding, in the genetic improvement of aquaculture species (Gao et al. 2023). Understanding the genetic diversity and population genetic structure of commercially important species is crucial, as it provides valuable insights that can guide their conservation and management (Wang et al. 2020).

Consequently, research focused on identifying the genetic diversity of *S. paramosain* in Ca Mau, Binh Dinh, and Hue is essential. This research holds the potential to provide detailed information about the genetic structure of crab populations in these regions, which in turn will assist in assessing the status of crab resources and in developing appropriate management and conservation actions.

MATERIALS AND METHODS

Field sampling

A total of 45 mud crab samples were obtained in lagoons along the coastal areas in Thua Thien Hue, Ca Mau, and Binh Dinh Provinces, Vietnam (Table 1, Figure 1). Samples were collected using collapsible traps, photographed, and identified based on morphological characteristics (Figure 2). A small amount of tissue (approximately 1 g) was collected from the walking leg of the mud crab samples and preserved in 96% ethanol. Samples were transported to the Laboratory of Gene Technology, Institute of Biotechnology, Hue University, Vietnam, for further analysis. Approval for all animal (samples) experiments was obtained from the Animal Ethics Committee of Hue University, Vietnam (Permit No. HUVN0029).

Cox1 sequencing

The muscle (100 mg) in the merus of the mud crab was used for DNA extraction. DNA was extracted using TopPURE® Genomic DNA Extraction Kit (ABT, Vietnam) following the manufacturer's guidelines. Next, to determine the quantity and quality of the extracted DNA, a 0.8% agarose gel electrophoresis was conducted, and SafeView™ Classic Nucleic Acid Stain (Applied Biological Materials Inc., Canada) was used. The isolated DNA was then stored at a temperature of -20°C. The *cox1* sequence of approximately 780 bp was amplified using the primer pairs of COI-s (5'-TTGACCCTGCTGGCGGTGG-3') and COI-a (5'-CAATTGAGGAGGGTAAAAATGGAGTAA-3') (Hongyu et al. 2016).

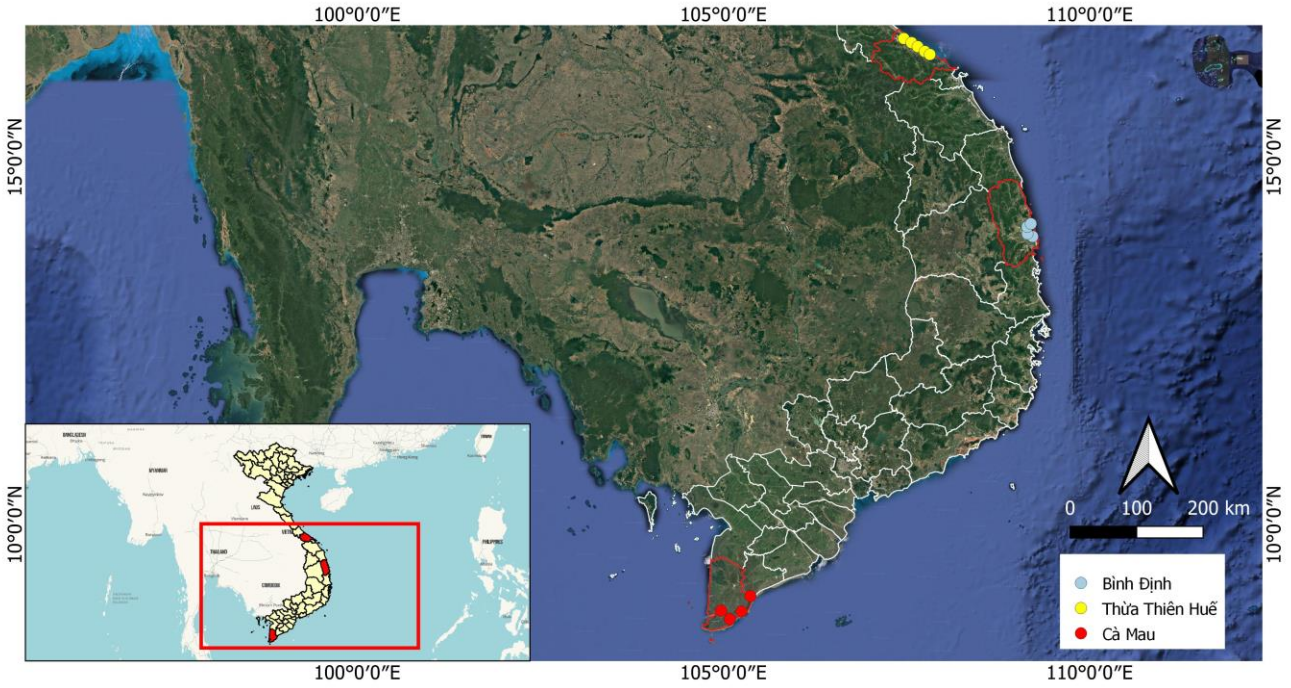

Figure 1. Map of sampling sites in province of Thua Thien Hue, Binh Dinh and Ca Mau, Vietnam

Table 1. Specimens of mud crab with locality and voucher code

Provinces	District	Sample code	Geographic coordinates
Thua Thien Hue (TTH)	Phong Dien	PDxx.HUE	16°38'49.0"N 107°28'09.3"E
	Phu Vang	PVxx.HUE	16°29'47.3"N 107°43'47.0"E
	Quang Dien	QDxx.HUE	16°36'09.5"N 107°33'22.7"E
	Huong Tra	HPxx.HUE	16°34'02.3"N 107°35'25.4"E
	Hue	THxx.HUE	16°32'18.3"N 107°38'32.2"E
Binh Dinh (BD)	Tuy Phuoc	PTxx.BD,	13°50'02.1"N 109°13'17.2"E;
		PSxx.BD,	13°51'41.5"N 109°13'41.3"E;
		PHxx.BD	13°51'45.5"N 109°14'36.6"E
	Quy Nhon	NBxx.BD	13°48'43.1"N 109°13'39.1"E
Ca Mau (CM)	Ngoc Hien	CM1-CM2, CM11-CM12	8°39'10.4"N 105°04'57.8"E
	Nam Can	CM3-CM5, CM13-CM14	8°48'31.7"N 105°14'24.5"E
	Dam Doi	CM6-CM7, CM15	9°00'39.5"N 105°22'05.2"E
	Cai Nuoc	CM8-CM10	8°51'12.3"N 105°00'51.5"E

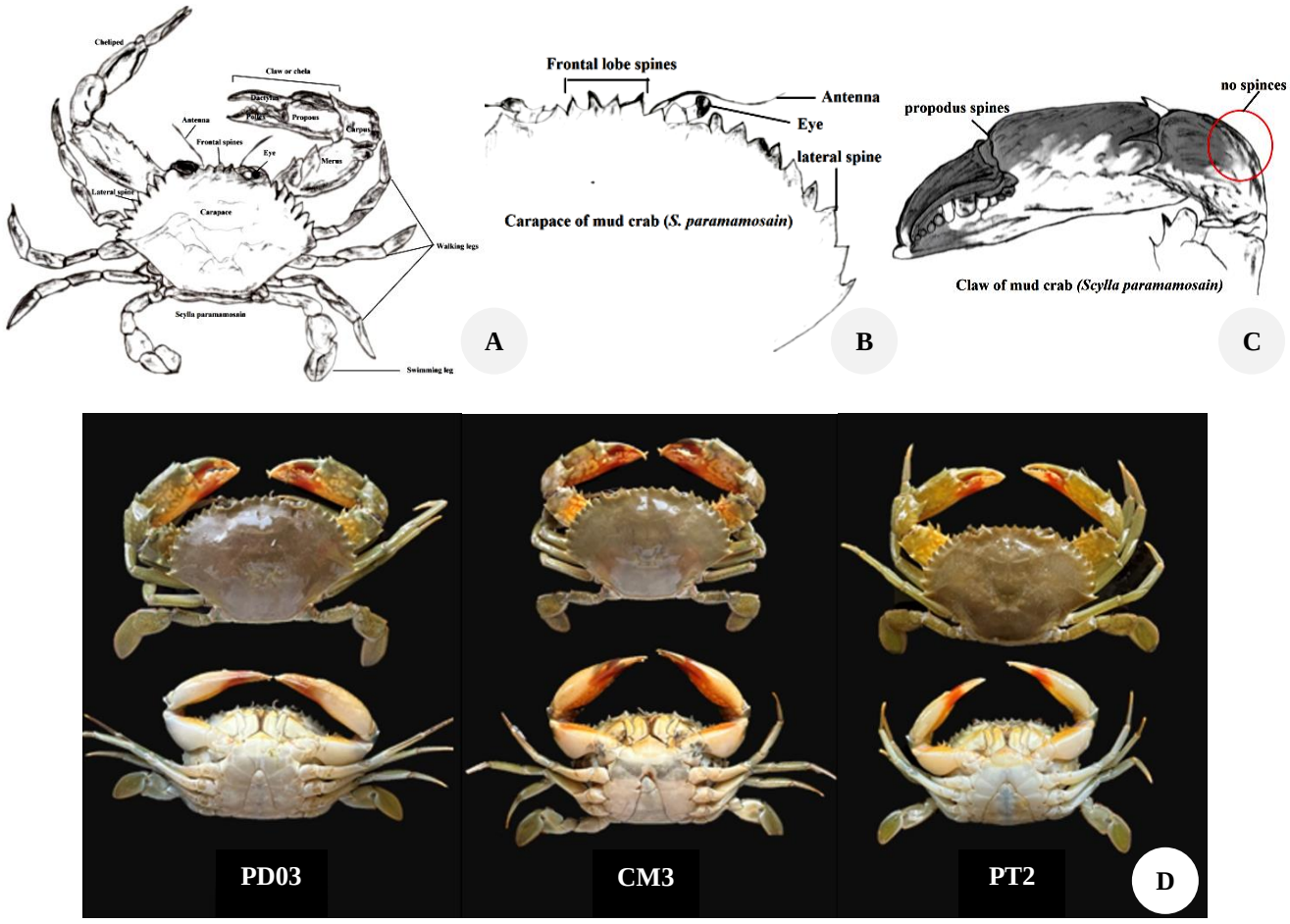

Figure 2. Morphology of collected mud crabs. A. Appearance of *Scylla paramamosain*; B. Spikes on the cephalothorax; C. Spikes on leg 1; D. Samples from 3 provinces, PD03 from Phong Dien District, Thua Thien Hue Province, CM3 from Nam Can District, Ca Mau Province and PT2 from Tuy Phuoc District, Binh Dinh Province

Polymerase chain reaction (PCR) was carried out in a total reaction volume of 60 µL, composed including 5 µL 100 ng μL^{-1} normalized DNA, 5 µL 10 nM of each primer, 30 µL 2× Go Taq® Green Master Mix (M7502, Promega, USA), and distilled water to the final reaction volume. PCR procedure included denaturation at 94°C for 10 min, followed by 35 cycles consisting of 94°C for 10s, annealing at 54°C for 20s, 72°C for 30s, and a final extension at 72°C for 10 min. PCR products are purified using the Wizard®SV Gel and PCR CleanUp System kit (Promega, USA). *Cox1* fragment sequencing was performed using the Sanger method (First BASE Laboratories, Selangor, Malaysia).

Data analysis

The DNA sequences obtained from both strands were assembled to obtain consensus sequences for each sample using BioEdit v7.0.5. All sequences produced in this research were submitted to GenBank (https://www.ncbi.nlm.nih.gov). The average base composition was determined using the GC Content Calculator (https://www.biologicscorp.com/tools/GCContent/).

Multiple sequence alignment was manually performed with the reference sequences of the South China Sea available in the NCBI database with MEGA 11 software. Genetic distances and diversity were also calculated with MEGA 11 (Tamura et al. 2021). The phylogenetic tree was constructed by the IQ-TREE v. 2.1.3 with the Maximum-likelihood (ML) method (Nguyen et al. 2015; Minh et al. 2020). The resulting trees were plotted using Interactive Tree of Life (iTOL) v.6 (Letunic and Bork 2021).

The number of haplotypes (H), nucleotide diversity (Pi), haplotypes diversity (Hd), number of segregating sites (S), number of mutations (η), the average number of nucleotide differences (K), Tajima's D and Fu's Fs were calculated using DnaSP v6.12 software (Rozas et al. 2017). Population expansion patterns in mud crab species were tested by estimating Fu's Fs and Tajima's D. The Fu's Fs statistic relies on haplotype distribution, while Tajima's D is based on allele frequency when comparing pairwise differences between sequences (Ramírez-Soriano et al. 2008).

Molecular variance analysis (AMOVA) was conducted to calculate the genetic differentiation indices (Fst) and partition genetic variation within and between populations using Arlequin 3.5.2, a population genetics software package. Gene flow (Nm) (Nei 1973) was analyzed by DnaSP v6.12 software (Rozas et al. 2017).

RESULTS AND DISCUSSION

Genetic variation and diversity

A portion of mtDNA, corresponding to the coding region of *cox1,* was successfully amplified from *S. paramamosain* individuals. The resulting 780 bp sequence was analyzed, and no insertions or deletions were detected across all samples. All *cox1* sequences were deposited at GenBank with the accession numbers form PP068926-PP068940 (Thua Thien Hue), PP068896-PP068910 (Binh Dinh) and PP068911-PP068925 (Ca Mau).

The proportion of Thiamine (T) was the highest, averaging 39.32%, while the proportion of Guanine (G) was the lowest, averaging 14.48%. All three codon positions exhibited base-compositional biases towards A+T (67.41%), which was higher than G+C (32.59%). Among the 45 sequences, 22 variable sites were observed, of which 10 were single-variable sites (site positions: 102, 201, 216, 303, 417, 486, 523, 633, 703, and 741) and 12 were parsimony-informative sites (site positions: 12, 60, 96, 210, 297, 366, 369, 507, 570, 663, 779, and 780) (Table 2). The average number of nucleotide differences (K) ranged from 0.91429 (Thua Thien Hue) to 2.55238 (Ca Mau), corresponding to a range of nucleotide diversity (Pi) ranging from 0.00117 (Thua Thien Hue) to 0.00327 (Ca Mau), respectively (Table 3).

Phylogenetic relationships

The phylogenetic tree of mud crab, compared to the other 70 sequences from GenBank, revealed that 45 samples were present in the different branches of the tree, indicating a high level of diverse gene flow among populations (Figure 3).

Table 2. Variable sites in 45 *cox1* sequences

Variable type	Variable sites	Nucleotide change	Samples
Single-variable sites	102	T>C	HP01.HUE
	201	C>A	CM13
	216	C>T	CM11
	303	T>C	CM13
	417	C>T	PD02.HUE
	486	T>C	PV01.HUE
	523	T>C	CM13
	633	C>A	TH4.HUE
	703	G>A	PH8-BD
	741	T>C	CM11
Parsimony-informative sites	12	G>A	HP02.HUE
	60	A>G	CM4, CM9, QD02.HUE, TH1.HUE
	96	G>A	CM6, CM14, CM15, TH4.HUE
	210	T>C	ND1-BD, HP02.HUE
	297	T>C	PH2-BD, TH4.HUE
	366	T>C	PH5-BD, TH2.HUE
	369	G>A	NB1-BD, CM6
	507	T>C	CM4, CM9, HP01.HUE, QD02.HUE, TH1.HUE
	570	C>T	CM4, CM9, QD02.HUE, TH1.HUE
	663	A>G	CM4, CM9, QD02.HUE, TH1.HUE
	779	T>A	CM7, TH1.HUE
	780	A>C	CM7, TH1.HUE

Table 3. Genetic diversity of mud crab populations based on *cox1* sequences

	Total	Thua Thien Hue	Binh Dinh	Ca Mau
No. of ind.	45	15	15	15
S	22	6	13	15
H	17	6	6	9
Hd	0.71414	0.57143	0.76190	0.80000
K	1.95758	0.91429	2.40000	2.55238
Pi	0.00251	0.00117	0.00308	0.00327

Notes: H: Number of haplotypes; S: Number of segregating sites; K: Average number of nucleotide differences; Hd: Haplotypes diversity; Pi: Nucleotide diversity

A total of 17 haplotypes were identified in 45 (17/45; 37.78%) mud crab individuals (Table 4). The haplotype diversity value (Hd) increased from 0.57143 (Thua Thien Hue) to 0.80000 (Ca Mau) (Table 3), corresponding to the percentage of the haplotype from 40% (Thua Thien Hue and Binh Dinh) to 60% (Ca Mau). The results shown in Table 4 indicated that only four haplotypes were abundant, with more than two individuals each, including Hap_3 (n=24), Hap_2, Hap_7, and Hap_8 (n=3). Three haplotypes occurred in more than one population, with a frequency of 66.67% (30/45 samples), while 13 haplotypes had only one sequence per haplotype.

The genealogic network analysis revealed the presence of two distinct clusters of haplotypes, with 3 haplotypes belonging to Cluster I and 13 to Cluster II. Hap_3 was the central of the haplotype network and all samples of this haplotype were obtained from mud crab populations in Vietnam (Table 4, Figure 4). Cluster I consisted of Hap_1 (PH2-BD), Hap_8 (CM6, CM14, CM15), and Hap_17 (TH4.HUE); these samples were collected from all three different regions and related to the nucleotide changes in site 96 (G>A) and 297 (T>C). The nucleotide changes of Hap_16 (TH1.HUE) were A>G (site 60), T>C (site 507), T>A (site 779), A>C (site 780), while that of Hap_13 (PV01.HUE) was T>C (site 486), thus, Hap_13 and Hap_16 located in the different branches and clusters (Table 2, Figure 4).

Population genetic structure

The analysis of molecular variance (AMOVA) showed that 0.15% of the variation occurred among mud crab populations in Vietnam, while 99.85% occurred within populations. The fixation index (Fst) was 0.00151, with a p-value of 0.38238 (Table 5).

The genetic distance among populations was assessed using the Fst, and in this investigation, a single negative Fst value (BD/CM) suggested limited genetic differentiation among these populations (Table 6). Across all populations, the mud crab population's Fst values were consistently lower than 0.039, with no significant differences in gene flow (Nm) between populations - an overall value of over 60.00 was observed.

Table 4. Number of haplotypes of three mud crab populations

Haplotypes	Number	Sample(s)
Hap_1	1	PH2-BD
Hap_2	2	PH5-BD and TH2.HUE
Hap_3	24	PH6-BD, PH7-BD, PS3-BD, PS5-BD, PS7-BD, PS8-BD, NB-BD, NB7-BD, PT6-BD, SP1-BD, CM1, CM2, CM3, CM5, CM8, CM10, CM12, PD01.HUE, PD03.HUE, PV02.HUE, PV03.HUE, QD01.HUE, TH3.HUE, and HUE15.HUE
Hap_4	1	PH8-BD
Hap_5	1	PS2-BD
Hap_6	1	NB1-BD
Hap_7	3	CM4, CM9, and QD02.HUE
Hap_8	3	CM6, CM14, CM15
Hap_9	1	CM7
Hap_10	1	CM11
Hap_11	1	CM13
Hap_12	1	PD02.HUE
Hap_13	1	PV01.HUE
Hap_14	1	HP01.HUE
Hap_15	1	HP02.HUE
Hap_16	1	TH1.HUE
Hap_17	1	TH4.HUE

Table 5. Analysis of Molecular Variance (AMOVA) of the four populations

Source	Degree of freedom	Sum of squares	Variance components	Percentage of total variance (%)	P-value
Among populations	2	2.000	0.00148 Va	0.15	0.38238
Within populations	42	41.467	0.97778 Vb	99.85	
Total	44	43.067	0.97926	100	
Fixation index (Fst)	0.00151	Remark: Va and Fst p-value=0.45663 ± 0.00530			

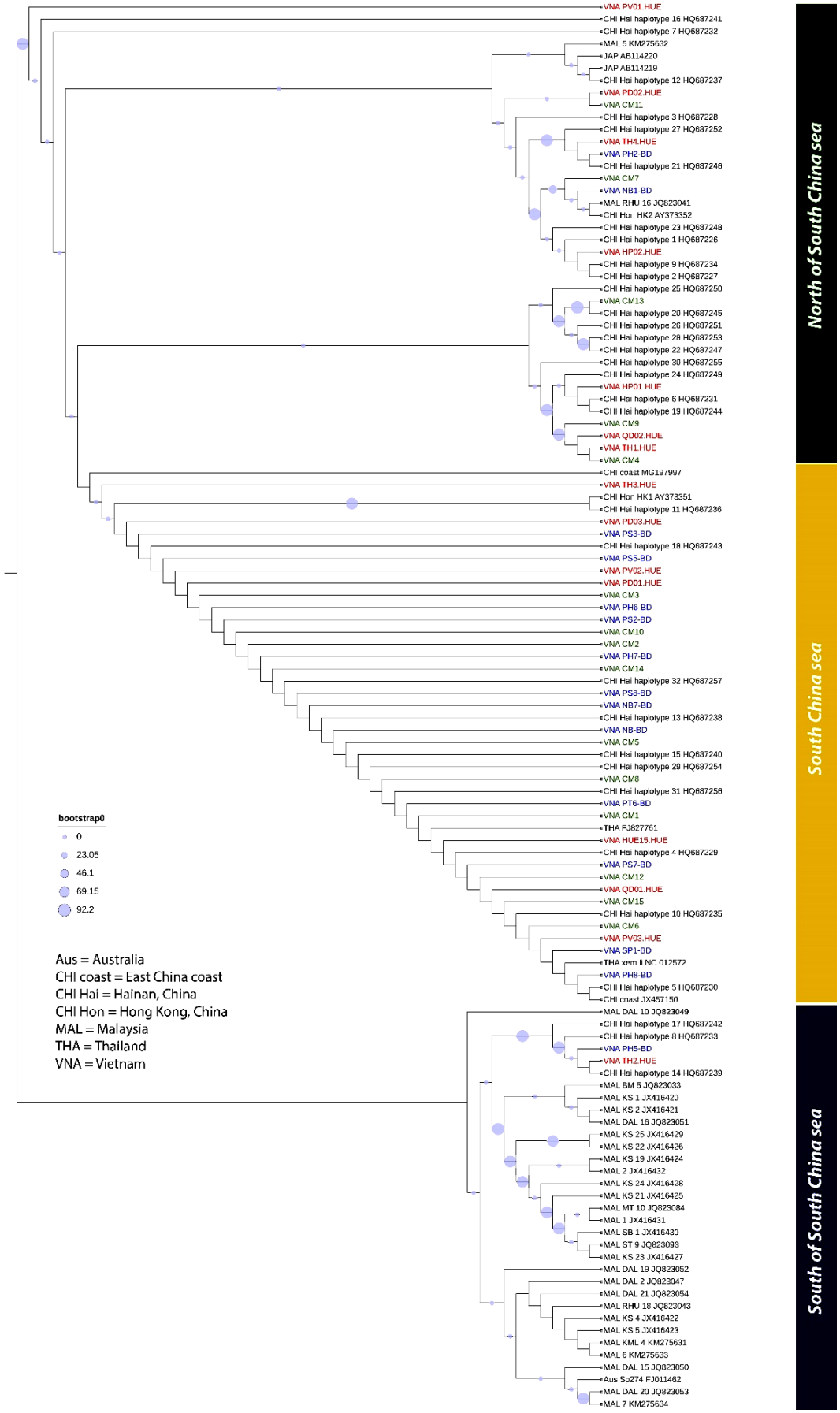

Figure 3. Phylogenetic tree of mud crab from different locations in the South China Sea. The sample names were Country + specimen name + NCBI accession numbers

Table 6. Fst values (above diagonal) and probability values (below diagonal) among different 3 populations

	Thua Thien Hue	Binh Dinh	Ca Mau
Thua Thien Hue	-	0.03903	0.01266
Binh Dinh	0.08108	-	-0.03366
Ca Mau	0.44144	0.89189	-

Note: No significant diffirence

Neutrality test and mismatch analysis

According to the DnaSP software analysis (Table 7), all the Tajima's D and Fu's Fs values were negative, indicating the presence of many low-frequency mutations in populations; this also suggests that the population is expanding. The values of Fu's Fs were strongly negative (P<0.02) in each population, showing the sensitive increase in detecting deviations from neutrality. The negative value of Fu's Fs provides evidence of an excess number of alleles compared to what is expected from a recent population expansion (Mehraj et al. 2017).

One common method used in population genetics is the analysis of mismatch distributions. Mismatch distributions are used to investigate the demographic history of populations, including factors such as population growth or decline, migration, and selection pressures. The result showed smooth and main unimodal curve peaks (Figure 5), which align with the population expansion model.

Overall, the findings suggested limited genetic diversity and significant gene migration between various populations, similar to previous discoveries in *S. paramamosain* population in China utilizing *cox1* sequences (Wang et al. 2020).

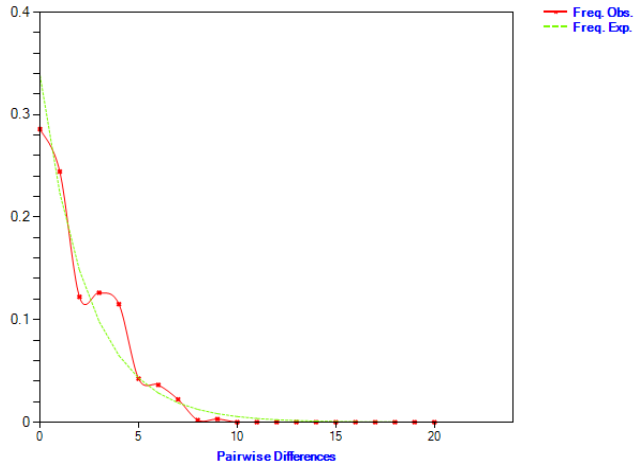

Figure 5. The unimodal pattern of mismatch distributions analysis for the *cox1* gene of mud crab

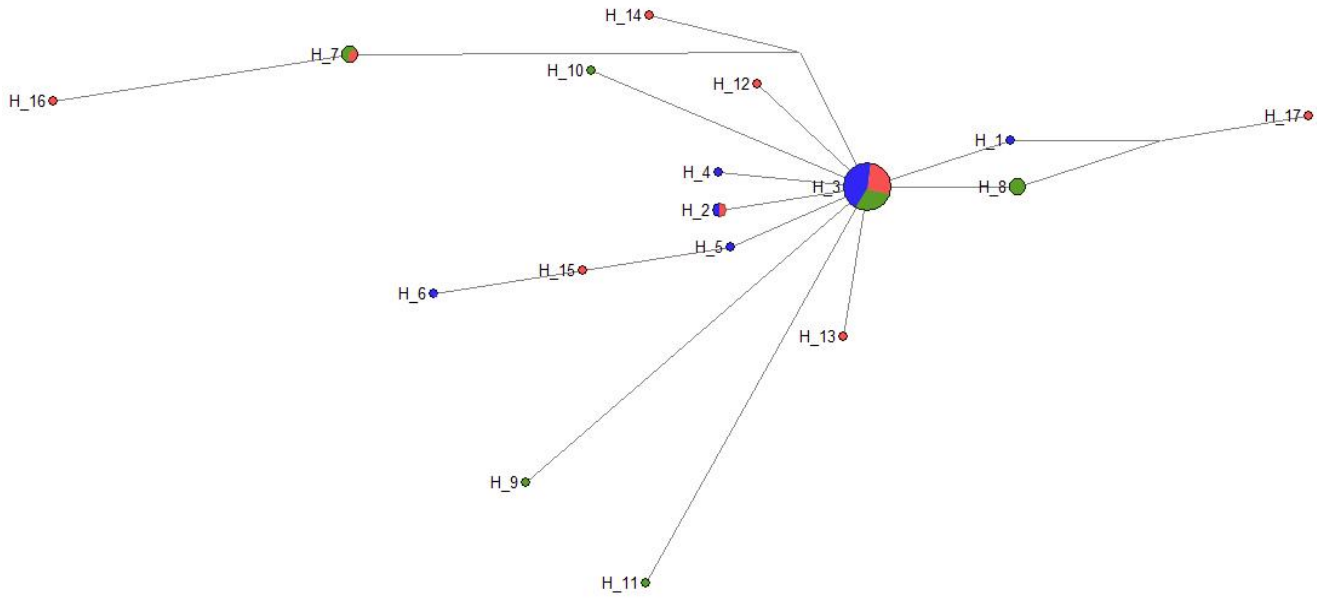

Figure 4. Median-joining haplotype network for the *cox1* dataset (red for Thua Thien Hue, blue for Binh Dinh and green for Ca Mau)

Table 7. Genetic diversity of mud crab populations based on *cox1* sequences

Test	Total	Thua Thien Hue	Binh Dinh	Ca Mau
Fu's Fs statistic	-10.821 (p=0.000)	-3.445 (p=0.024)	-3.097 (p=0.034)	-0.336 (p=0.208)
Fu and Li's D	-1.61261 (p>0.10)	-1.84715 (p>0.10)	-1.95602 (0.10>p>0.05)	-1.28236 (p>0.10)
Fu and Li's F	-2.06981 (0.10>p>0.05)	-2.10215 (0.10>p>0.05)	-2.18061 (0.10>p>0.05)	-1.56336 (p>0.10)
Tajima's D	-2.00303* (p<0.05)	-1.78340 (0.10>p>0.05)	-1.76624 (0.10>p>0.05)	-1.57170 (p>0.10)

Notes: *Significance Level=0.0500

Discussion

The most reliable barcodes for discriminating between different animal species are obtained using the mitochondrial gene coding for cytochrome c oxidase 1 (*cox1*) and cytochrome b (*cytb*). Specifically, a ~650 bp fragment of the mitochondrial gene *cox1* is the most commonly used DNA barcode for seafood identification. Numerous studies have demonstrated the effectiveness of *cox1* barcoding for accurately identifying a wide range of fish species and for detecting mislabeling. Mitochondrial (mtDNA) and nuclear (nDNA) genes play roles in various approaches, but *cox1* and *cytb* are particularly valuable for species discrimination (Filonzi et al. 2023).

Researchers have consistently found the *cox1* gene to be a highly efficient tool in DNA barcoding for identifying crustaceans. This research was conducted at various levels - order, family, genus, and species (Eischeid et al. 2016), has shown greater sequence variation in the *cox1* DNA barcoding region than in other groups. They demonstrate a high ratio of inter/intra-specific sequence divergence, referred to as the barcode gap, and average sequence divergence is 17% between different species of the same genus and 0.46% within species (Eischeid et al. 2016). The *cox1* gene, the most commonly used genetic marker in crab DNA barcoding studies, has been reported to be highly efficient for species identification and genetic diversity of the *Scylla* genus in China (Zhang et al. 2008; Wang et al. 2020) and *Cancer irroratus* Say 1817 in Iceland (Gíslason et al. 2013).

This study aimed to investigate the genetic diversity and population genetic structure of mud crab (*S. paramamosain*) in Vietnam by analyzing 45 mitochondrial *cox1* sequences. It is often observed errors occur in mtDNA nucleotide sequencing due to co-amplifying nuclear mitochondrial pseudogenes (numts) and mitochondrial genes in eukaryotes (Dung et al. 2023). Crab species, such as the Atlantic rock crab (*C. irroratus*) (Gíslason et al. 2013) or mud crab (*S. paramamosain*) (Zhang et al. 2008), have been found to possess numts. In mud crab, 14 pseudogenes have been identified and divided into two groups. Group 1 showed no insertion or deletion sites, while Group 2 exhibited 8 deletion and 5 insertion sites, resulting in frame-shift mutations (Zhang et al. 2008). However, in this study, no insertion or deletion sites were found in 45 *cox1* sequences; therefore, the *cox1* gene sequences can be used to investigate mud crab genetic diversity.

Genetic variation and diversity

Genetic variation pertains to the variances in genes within a particular species, which is crucial for the species to adapt to changes in their environment. A high level of genetic diversity indicates the species' ability to thrive and adapt (Birader 2023). Base composition analysis of *cox1* sequences from three provinces in Vietnam showed a significant bias towards A/T content (67.41%) over G/C content (32.59%). This finding is consistent with the results of other crab species, such as *S. paramamosain* in China, where the A/T content (66.9%) was significantly higher than the G/C content (34.1%) (Wang et al. 2020) or fresh crab *Geothelphusa dehaani* White 1847 in Japan, which also displayed a strong A/T bias in their *cox1* regions (Huervana et al. 2023). In a study by Pang et al. (2023), the mitogenome composition of *Diogenes edwardsii* De Hann 1849 was highly A/T biased (72.16%).

In *cox1* sequence analysis, Hd and π are important parameters for evaluating genetic variation (Dung et al. 2023). A high level of genetic diversity is indicated by Hd>0.5 and π>0.005. The mud crab populations demonstrate a significant level of genetic diversity with Hd=0.71414 (>0.5) and π=0.00251 (<0.005). These findings are consistent with results from mud crab populations in China (Hd=0.738 and π=0.00194) (Wang et al. 2020). Our results show elevated high haplotype diversity (Hd=0.71414) but low nucleotide diversity (π=0.00251), which implies a historical population bottleneck followed by rapid growth and the accumulation of new mutations (Grant and Bowen 1998).

Phylogenetic relationships

An analysis of the partial *cox1* sequences from mud crab populations in Vietnam discovered 17 different haplotypes and 22 variable sites within a sample size of 45 (Table 3). Hap_3 was the dominant haplotype, shared by 24 individuals (53.33% of all individuals) and present in all localities, while 10 unique haplotypes (22.22%) were identified. The phylogenetic tree of 115 sequences from Vietnam (this study), China (East Coast, Hong Kong, and Hainan), Japan, Malaysia, Thailand, and Australia showed three main clusters corresponding to three regions: North of South China Sea (Cluster I), the South China Sea (Cluster II), and the South of South China Sea (Cluster III). Cluster I consisted of samples from China and Japan, 13 sequences from Vietnam (including all populations), and 1 sequence from Malaysia. Cluster II had 29 sequences (64.44%) from Vietnam, 13 sequences from China, and 2 sequences from Thailand. Most sequences in Cluster 3 came from Malaysia; however, this cluster also included 2 sequences from Vietnam (PH5.BD and TH2.HUE), and 3 sequences were obtained from Hainan (China). This result indicated that the migration of mud crabs in the South China Sea is from south to north. According to He et al. (2010), the surface circulations in the South China Sea range from the Sunda Shelf (Java Sea, Indonesia) to the East Sea of Vietnam, Hainan, Hong Kong, Taiwan Island, and the East China Sea. This is the reason why the loss of mud crabs from Vietnam was present in Cluster I, and a high level of gene flow was found (He et al. 2010).

Population genetic structure

The AMOVA results suggest significant genetic differentiation among mud crab populations in Vietnam, with a Va value of 0.00148, an Fst value of 0.00151, and a p-value of 0.38238. The observed average p-value for both Va and Fst values is 0.45663±0.00530, signifying significant genetic diversity within the Vietnamese mud crab populations. This robust and reliable data underscores the importance of determining the genetic structure of natural populations, a crucial aspect of population genetics and holds significant relevance in evolutionary biology, conservation, forensics, and animal breeding, a crucial aspect of population genetics. The calculation of Fst, a

broad approach to evaluating population structure, further supports our findings. Fst values typically range from 0 to 1.0 (0: No genetic differentiation; 0-0.05: Little genetic differentiation; 0.05-0.15: Moderate genetic differentiation; 0.15-0.25: Significant genetic differentiation; 0.25-1.0: Highly significant genetic differentiation; and 1.0: Complete genetic differentiation (Khan et al. 2021). In this study, the Fst value was 0.00151 (<0.05), suggesting little genetic differentiation between the crab populations. The greatest level of differentiation was discovered to be 0.03903 between Thua Thien Hue and Binh Dinh, while the smallest Fst value of -0.03366 was observed for Binh Dinh and Ca Mau. Negative Fst values can arise when alleles from different populations are less similar to one another than those within a population, but this is not the case here. Instead, the negative Fst values should essentially be interpreted as zero values; a zero value for Fst indicates no genetic subdivision between the considered populations (Smaragdov et al. 2018). This finding is similar to a study on mud crab populations along the Southeastern coast of China based on *cox1* (Fst=0.005707, p<0.05) (Wang et al. 2020). Furthermore, the majority of genetic variation (99.85%) stemmed from differences within populations, indicating extensive gene exchange among the three populations (Table 4). This outcome was similar to those found for mud crabs in China, where the total genetic variation mainly occurred within populations (99.68%), and only 0.32% was contributed by population variation (Wang et al. 2020). In contrast, in coastal horseshoe crab (*Tachypleus gigas* O.F.Müller 1785) in Indonesia, most of the variation was found within populations (95.23%) compared to less among populations (4.77%) (Aini et al. 2021). According to Su et al. (2023), six populations of Chinese mitten crab (*Eriocheir sinensis* H.Milne Edwards 1853) showed higher genetic variability among individuals of the same population and very low genetic differentiation between populations. These findings showed that the genetic differentiation within crab populations was higher than that of different populations in each country, so investigating the genetic distance among different countries may have the same findings.

The genetic structure of a population is influenced by gene flow (Nm), which represents the number of migrants per generation. This research has significant implications for genetic exchange. Gene flow can be classified into three categories: high (Nm≥1.0), medium (between 0.250 and 0.99), and low (between 0.0 and 0.249) (Dung et al. 2023). Nm≥1.0 indicates a significant gene flow between populations. Our data suggest that there was a very high level of gene flow among the three populations studied (Nm=60). The results also indicated a high level of gene flow (Nm ranges from 4.35 to 13.68) but no significant differences between populations. The level of genetic exchange of the population between Thua Thien Hue and Binh Dinh and Binh Dinh and Ca Mau was lower than that between Thua Thien Hue and Ca Mau (as shown in Table 7). This finding suggested that the seedling mud crabs from Ca Mau may have been introduced into Thua Thien Hue and can be leaked into the natural environment, highlighting the practical relevance of our research.

Neutrality test and mismatch analysis

Mud crab populations in Vietnam exhibit an excess of rare genetic variants, which is indicative of positive selection. This was based on the negative values of Tajima's D and Fu's Fs tests conducted on each population, although the results were not statistically significant. Additionally, the mismatch distributions of the mud crab populations showed smooth unimodal peaks corresponding to the population expansion model (Figure 5). For mud crabs in Vietnam, the *cox1* data indicated that the population genetic structure was genetically homogeneous with low differentiation, similar to the mud crab population distributed along the South-eastern coasts of China (He et al. 2010; Hongyu et al. 2016). The findings show that it is possible to use seedling mud crabs from natural sources, or it is necessary to evaluate the genotypes of parent crab sources in artificial reproduction. However, the sample size is a limitation of this study, hence, a comprehensive study with more populations and sample size will be conducted to investigate the overall genetic diversity of mud crab populations in Vietnam.

In conclusion, *S. paramamosain* populations in Vietnam showed high genetic diversity within the populations, as indicated by their low genetic variation and substantial gene flow between populations when analyzing the *cox1* gene. The haplotype diversity ranged from 0.57143 (Thua Thien Hue) to 0.80000 (Ca Mau). The variation within-population (99.85%) was higher than among-population variation (0.15%), and the population is expanding.

ACKNOWLEDGEMENTS

This study was supported by a grant from the Vietnam Ministry of Education and Training (Project No. B2023-DHH-23).

REFERENCES

Aini NK, Wardiatno Y, Effendi H, Mashar A, Madduppa H. 2021. High genetic diversity and mixing of coastal horseshoe crabs (*Tachypleus gigas*) across major habitats in Sundaland, Indonesia. PeerJ 9: e11739. DOI: 10.7717/peerj.11739.

Avianto, Sulistiono, Setyobudiandi I. 2013. Habitat characteristics and potency of mud crabs *Scylla serrata*, *S. transquaberica*, and *S. olivacea* in Cibako Mangrove Forest, Garut District, West Java. Bonorowo Wetlands 3: 55-72. DOI: 10.13057/bonorowo/w030201.

Birader K. 2023. Genetic diversity and the adaptation of species to changing environments. J Biodivers Endang Species 11: 474. DOI: 10.37421/2332-2543.2023.11.474.

Chen F, Wang K. 2019. Characterization of the innate immunity in the mud crab *Scylla paramamosain*. Fish Shellfish Immunol 93: 436-448. DOI: 10.1016/j.fsi.2019.07.076.

Dung TQ, Quang HT, Nguyen PTT, Trung NT, Huy ND, Quyen BV. 2023. Assessment of genetic variation in the pond loach *Misgurnus anguillicaudatus* (Cantor, 1842), based on COI gene fragments. Appl Ecol Environ Res 22 (1): 933-947. DOI: 10.15666/aeer/2201_933947.

Eischeid AC, Stadig SR, Handy SM, Fry FS, Deeds J. 2016. Optimization and evaluation of a method for the generation of DNA barcodes for the identification of crustaceans. LWT 73: 357-367. DOI: 10.1016/j.lwt.2016.06.033.

Filonzi L, Ardenghi A, Rontani PM, Voccia A, Ferrari C, Papa R, Bellin N, Nonnis Marzano F. 2023. Molecular barcoding: A tool to guarantee

correct seafood labelling and quality and preserve the conservation of endangered species. Foods 12: 2420. DOI: 10.3390/foods12122420.

Gao W, Cui W, Wu F, Chen H, Liu S, Guan M, Saqib HS, Ye S, Ikhwanuddin M, Ma H. 2023. Genetic diversity and differences among three F1 families and two wild populations of genus *Scylla* using microsatellite markers. Fishes 8: 18. DOI: 10.3390/fishes8010018.

Gíslason ÓS, Svavarsson J, Halldórsson HP, Pálsson S. 2013. Nuclear mitochondrial DNA (numt) in the Atlantic rock crab *Cancer irroratus* Say, 1817 (Decapoda, Cancridae). Crustaceana 86 (5): 537-552. DOI: 10.1163/15685403-00003191.

Grant WAS, Bowen BW. 1998. Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. J Heredity 89 (5): 415-426. DOI: 10.1093/jhered/89.5.415.

He L, Zhang A, Weese D, Zhu C, Jiang C, Qiao Z. 2010. Late Pleistocene population expansion of *Scylla paramamosain* along the coast of China: A population dynamic response to the Last Interglacial sea level highstand. J Exp Mar Biol Ecol 385 (1-2): 20-28. DOI: 10.1016/j.jembe.2010.01.019.

Hongyu M, Chunyan M, Lingbo M, Xincang L, Yuanyou L. 2016. Microsatellite markers in the mud crab (*Scylla paramamosain*) and their application in population genetics and marker- assisted selection. In: Ibrokhim YA (eds). Microsatellite Markers. IntechOpen, Rijeka. DOI: 10.5772/65041.

Huervana JJC, Kano Y, Ando D, Onikura N, Kurita Y. 2023. Preliminary assessment of genetic variation in the Japanese endemic freshwater crab, Geothelphusadehaani, based on mitochondrial DNA sequences. Biodivers Data J 11: e97438. DOI: 10.3897/BDJ.11.e97438.

Khan MMH, Rafii MY, Ramlee SI, Jusoh M, Al Mamun M, Halidu J. 2021. DNA fingerprinting, fixation-index (Fst), and admixture mapping of selected Bambara groundnut (*Vigna subterranea* [L.] Verdc.) accessions using ISSR markers system. Sci Rep 11: 14527. DOI: 10.1038/s41598-021-93867-5.

Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49 (W1): W293-W296. DOI: 10.1093/nar/gkab301.

Mehraj H, Sharma S, Ohnishi K, Shimasaki K. 2017. Study on inter-taxon population structure and diversity variation of hosta inferring from trnG-trnS regional cpDNA. Annal Agric Sci 62 (2): 211-220. DOI: 10.1016/j.aoas.2017.12.003.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37 (5): 1530-1534. DOI: 10.1093/molbev/msaa015.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating Maximum-Likelihood Phylogenies. Mol Biol Evol 32: 268-274. DOI: 10.1093/molbev/msu300.

Pang X, Fu W, Feng J, Guo B, Lin X, Lu X. 2023. The complete mitochondrial genome of the hermit crab *Diogenes edwardsii* (Anomura: Diogenidae) and phylogenetic relationships within Infraorder Anomura. Genes 14 (2): 470. DOI: 10.3390/genes14020470.

Ramírez-Soriano A, Ramos-Onsins SE, Rozas J, Calafell F, Navarro A. 2008. Statistical power analysis of neutrality tests under demographic expansions, contractions and bottlenecks with recombination. Genetics 179: 555-567. DOI: 10.1534/genetics.107.083006.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34 (12): 3299-3302. DOI: 10.1093/molbev/msx248.

Shakawi AMHA, Hassan R, Mustapah DS. 2022. Effects of water parameters on population structure of mud crab from Buntal Mangroves, Kuching, Sarawak, Malaysia: A GLM analysis. Biodiversitas 23: 2580-2585. DOI: 10.13057/biodiv/d230539.

Smaragdov MG, Kudinov AA, Uimari P. 2018. Assessing the genetic differentiation of Holstein cattle herds in the Leningrad region using Fst statistics. Agirc Food Sci 27: 96-101. DOI: 10.23986/afsci.69777.

Su S, Nsekanabo JD, Munganga BP, He X, Li J, Yu F, Wang M, Tang Y. 2023. Genetic diversity and population structure of the Chinese mitten crab (*Eriocheir sinensis*) from six different lakes using microsatellites. Fishes 8 (5): 220. DOI: 10.3390/fishes8050220.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol Biol Evol 38 (7): 3022-3027. DOI: 10.1093/molbev/msab120.

Wang W, Ma C, Chen W, Jin Z, Zhao M, Zhang F, Liu Z, Ma L. 2020. Population genetic diversity of mud crab (*Scylla paramamosain*) from southeast coastal regions of China based on mitochondrial COI gene sequence. Gene 751: 144763. DOI: 10.1016/j.gene.2020.144763.

Zhang FY, Ma LB, Qiao ZG, Chen LQ. 2008. Multiple nuclear pseudogenes of mitochondrial cytochrome oxidase I gene in *Scylla paramamosain* (Decapoda : Portunidae). J Mol Cell Biol 41 (2): 155-161. [Chinese]