

Original Research Articles

Characterization and development of genome-wide microsatellites from mud crab *Scylla paramamosain*

Sigang Fan^{1,2,a}, Changhong Cheng¹, Hongling Ma¹, Guangxin Liu¹, Minze Liao¹, Zhixun Guo^{1,2b}

¹ Key Laboratory of South China Sea Fishery Resources Exploitation & Utilization, Ministry of Agriculture and Rural Affairs, South China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences / Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Guangzhou, 510300, China, ² Sanya Tropical Fisheries Research Institute, Sanya, Hainan, 572018, China

Keywords: microsatellite, genome-wide, genetic diversity, genomic regions, *Scylla paramamosain*

<https://doi.org/10.46989/001c.161390>

Israeli Journal of Aquaculture - Bamidgah

Vol. 78, Issue 2, 2026

Scylla paramamosain is a globally important mariculture crab species. However, its resources for simple sequence repeat (SSR) loci remain severely underdeveloped. In the present study, MISA software was employed to systematically detect and characterize SSR loci across the whole genome of *S. paramamosain*. A total of 3,184,249 SSR loci were identified in the *S. paramamosain* genome, with a cumulative length of 241,424,069 bp, accounting for 15.62% of the total genome size. The frequency and density of SSRs were 2,059.64 loci/Mb and 156,158.13 bp/Mb, respectively. A total of 259 distinct repeat motifs were detected, with repeat numbers ranging from 5 to 1,551. Mononucleotide repeats were the most abundant type, followed by dinucleotide repeats. Among all repeat motifs, motif A was the most prevalent, followed by the dinucleotide motifs AG and AC. Regarding chromosomal distribution, chromosome 3 had the most SSR loci, while chromosome 44 had the fewest. Notably, chromosome 44 exhibited the highest SSR frequency and density. In total, 2,382,227 SSR markers were developed; among these, 72.82% were located in intergenic regions, 20.15% in introns, and 7.03% in exons. Sixteen SSR markers were amplified in 32 *S. paramamosain* individuals, among which eight markers possessed high polymorphism. To the best of our knowledge, this is the first chromosome-level genome-wide SSR analysis in *S. paramamosain*. These findings provide valuable genetic resources and a theoretical basis for subsequent genetic breeding and germplasm improvement of *S. paramamosain*.

INTRODUCTION

Mud crab *Scylla paramamosain* is widely distributed across the Indo-West Pacific region.¹ Endowed with desirable traits including large body size, rapid growth rate, palatable flesh, rich nutritional value, and strong environmental adaptability, *S. paramamosain* is extensively cultured along the coastal areas of China and many Southeast Asian countries.^{2,3} In 2024, the aquaculture output of mud crabs in China reached 162 thousand tons, making it the highest-yielding mariculture crab species and underscoring its significant economic value.⁴

Simple sequence repeats (SSRs), also referred to as microsatellites, are tandemly repeated DNA sequences consisting of 1–6 base pairs (bp).⁵ SSRs originate from strand misalignment induced by DNA polymerase slippage during replication, with mutation rates ranging from 10^{-2} to 10^{-6} nucleotides per generation in eukaryotic genomes.⁶ SSRs

are characterized by high genomic abundance, codominant inheritance, high polymorphism, and random distribution across the genome.⁵ Due to these inherent advantages, SSRs have been widely utilized in population genetic analysis, species discrimination, and marker-assisted selection breeding.^{7–11} To date, only a limited number of SSR loci and markers have been developed for *S. paramamosain*,¹² which has hindered research on the genetic structure and germplasm resource conservation of mud crab.

With the advancement of next- and third-generation sequencing technologies, the whole-genome sequences of numerous aquatic organisms have been successfully sequenced and assembled, laying a robust foundation for the large-scale mining of genomic SSRs. For instance, 412,381 SSR loci were identified in *Acanthopagrus schlegelii*¹³; 417,724 SSRs were detected in *Pelteobagrus vachelli*¹⁴; and 354,613 SSR loci were characterized in *Oncorhynchus mykiss*.¹⁵ Notably, Zhao et al.² first reported the genome

a Corresponding author. Sigang Fan, e-mail: fansigang@scsfri.ac.cn

b Corresponding author, Zhixun Guo, e-mail: guozhixun1@163.com

sequence of *S. paramamosain*, providing abundant genomic resources for the exploration and utilization of its genetic resources and thereby addressing the aforementioned limitation of insufficient SSR markers for this species.

In the present study, a large number of SSRs and SSR markers were identified and characterized from the *S. paramamosain* genome. The location of SSR markers in the genome was precisely mapped. Additionally, more than 40 polymorphic and stable SSR markers were developed, and a wild population of *S. paramamosain* was initially evaluated using eight polymorphic SSR markers. Collectively, the findings of this study provide valuable support for the conservation and genetic breeding of *S. paramamosain*.

MATERIALS AND METHODS

2.1. GENOME ACQUISITION, SSR IDENTIFICATION, AND CORRELATION ANALYSIS

The whole-genome of *S. paramamosain* was retrieved from the National Genomics Data Center (NGDC; <https://ngdc.cncb.ac.cn/search/?dbId=gwh&q=GWHALOH000000000&page=1>) under the Accession number GWHALOH000000000. As previously reported, the genome of *S. paramamosain* spans 1.55 Gb, encompassing 49 chromosomal sequences and additional unanchored contigs that cannot be mapped to specific chromosomes.²

SSR loci within the genome were systematically identified and characterized using the MicroSatellite (MISA) software. The search parameters were established as follows: mononucleotide repeats with a minimum of 10 tandem copies; dinucleotide repeats with no fewer than 6 copies; tri-, tetra-, penta-, and hexanucleotide repeats with a minimum of 5 copies; and compound SSRs with an inter-motif interval of < 100 nucleotides. Motifs differing solely due to base complementarity or nucleotide arrangement were categorized as the same type of SSR locus for unified statistical analysis.

SPSS 19 software was employed to perform correlation analyses between chromosome length and four SSR-related metrics: locus number, total length, density, and frequency.

2.2. SSR PRIMER DESIGN AND SYNTHESIS

Primers targeting the identified SSR loci were designed using Primer3 software (<http://primer3.org/>) with the following optimized parameters: primer length ranging from 18 to 30 bp; annealing temperature between 50 and 60°C; GC content of 40%–60%; and expected PCR product size of 80–350 bp. The genomic positions of SSR loci were annotated based on the GFF format annotation file of the *S. paramamosain* genome.²

2.3. GENOMIC DNA EXTRACTION

Thirty-two wild *S. paramamosain* individuals were collected from Dapeng Bay, Shenzhen (22°26'N, 114°15'E). Muscle tissue were dissected from each individual, immediately

placed into 2 mL cryovials, and stored at –80°C until DNA extraction.

Genomic DNA was isolated from the muscle using the Marine Animal Tissue Genomic DNA Extraction Kit (Tiangen Biochemical Technology Co., Ltd., Beijing, China) following the manufacturer's standard protocol. The integrity of the extracted DNA was assessed by 1% agarose gel electrophoresis, and its concentration was quantified using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA).

2.4. SSR MARKER VALIDATION

A total of 125 pairs of SSR primers were selected and synthesized (Table S1). The strategy for SSR marker was listed as follows: only 2–6 bp SSR motifs were selected. Amplicon lengths were set to 100–350 bp. For dinucleotide repeats, the repeat number was 8–50; for tri- to hexanucleotide repeats, the repeat number was 5–25. Initial validation of primer amplification efficiency was conducted using genomic DNA from 3 *S. paramamosain* individuals. The PCR reaction was performed in a 20 µL volume containing 0.2 µL ExTaq enzyme (Takara, Japan), 2 µL 10×ExTaq Buffer (with Mg²⁺), 1.6 µL dNTP Mix (2.5 mM), 0.8 µL of each forward and reverse primer (10 µmol/L), 0.5 µL genomic DNA (50 ng/µL), and 14.1 µL deionized water. The PCR cycling program was set as follows: pre-denaturation at 95°C for 5 min; 35 cycles of denaturation at 95°C for 30 s, annealing at 53°C for 30 s, and extension at 72°C for 30 s; followed by a final extension at 72°C for 5 min. PCR products were separated by 1% agarose gel electrophoresis to verify the generation of single target bands. Primers that successfully amplified distinct, single bands were selected for subsequent experiments, and the 5' end of each forward primer was labeled with Fluorescein Amidite (FAM) (Table S1). Primer synthesis and fluorescence labeling were commissioned to Sangon Biotech (Shanghai) Co., Ltd (China).

2.5. POPULATION GENETIC ANALYSIS

Eight SSR markers were randomly chosen to evaluate the genetic diversity of the wild *S. paramamosain* population from Dapeng Bay, Shenzhen. The PCR reaction system and cycling conditions were identical to those described in section 2.4. Fluorescently labeled PCR products were analyzed by capillary electrophoresis using an ABI 3730XL automatic sequencer (Thermo Fisher Scientific, USA) to determine the genotype of each individual. Popgene 1.32 software (Yeh and Boyle, 2000) was utilized to calculate genetic diversity parameters, including the number of alleles (*Na*), number of effective alleles (*Ne*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), Shannon's Information index (*I*), Hardy-Weinberg equilibrium (HWE), and fixation index (*Fis*). The polymorphism information content (PIC) of each SSR locus was computed using Cervus software (Marshall et al. 1998). Linkage disequilibrium (LD) was analysed using Arlequin 3 (Excoffier et al., 2007).

RESULTS

3.1. GENOME-WIDE DISTRIBUTION PROFILES OF SSR IN *S. PARAMAMOSAIN*

A total of 3,184,249 SSR loci, including 1,185,523 compound SSRs and 1,998,726 perfect SSRs, were comprehensively identified across the *S. paramamosain* genome (Table S2). The average distance between microsatellite loci was 486 bp. The total length of all SSRs was 241,424,069 bp, accounting for 15.62% of the genome. The frequency and density of SSRs were 2,059.64 loci/Mb and 156,158.13 bp/Mb, respectively (Table S2).

A total of 259 repeat motifs were detected, with repeat numbers ranging from 5 to 1,551 (Table S3). The abundance of SSRs was significantly negatively correlated with repeat number.

The number of SSRs with repeated 10 was the domain, followed by those with repeated 6 (Figure 1). Furthermore, a negative correlation was detected between motif length and SSR count, with longer repeat motifs corresponding to fewer total loci (Figure 1).

Among perfect SSRs, a total of 1,037,673 mononucleotide repeats were identified, accounting for 51.92% of all perfect SSRs, followed by dinucleotide repeats (643,927 loci, 32.22%) and trinucleotide repeats (257,651 loci, 12.89%) (Table 1). Only 457 hexanucleotide repeats were detected (Table 1). Mononucleotide repeats showed the highest density (4,298.13 loci/Mb), followed by dinucleotide (2,667.20 loci/Mb) and trinucleotide repeats (1,067.21 loci/Mb) (Table 1).

The top 10 most abundant SSR motifs were as followed: A, AG, AC, AT, AGG, AAT, ACT, ACC, C, AAC, and AAAT (Table 2). A total of 1,928,296 A was detected, accounting for 94.47% of mononucleotide repeats, whereas C accounted for only 5.53% (Table 2, Figure 2a). The frequency of A was also highest (1,247.26 loci/Mb).

Among all motifs, the dinucleotide motif AG had the longest length (52,100,292 bp) and highest density (33,699.62 bp/Mb), followed by AC motif (34,301,090 bp, 22,186.71 bp/Mb) (Table 2). AG and AC were accounted for 46.45% and 35.71% in dinucleotide, respectively (Figure 2b).

The most abundant and frequent trinucleotide was AGG (246,396, 19.16%), followed by AAT (239,814, 18.65%) (Table 2, Figure 2c). The AAT motif displayed the longest total length (7,940,709 bp) and highest density (5,136.23 bp/Mb), followed by ACT (7,916,031 bp, 5,120.26 bp/Mb) (Table 2).

AAAT was the dominant motif in tetranucleotides, accounting for 30.89%, followed by ACAG (11.58%) (Figure 2d). AACCT was the primary motif in pentanucleotide, accounting for 67.40% (Figure 2e). AAGGAGA was the most prevalent in hexanucleotide, accounting for 11.49%, followed by AGAGGG (11.45%) and AAGAGG (11.11%) (Figure 2f).

3.2. CHROMOSOMAL DISTRIBUTION CHARACTERISTICS OF SSRs

Among the 49 chromosomes of *S. paramamosain*, chromosome 41 exhibited the shortest total SSR length (1,099,300 bp), followed by chromosome 44 (1,117,260 bp) (Table S2, Figure 3a). Chromosome 3 harbored the largest number of SSR loci (124,426) and the longest total SSR length (9,715,090 bp), whereas chromosome 44 contained the fewest SSR loci (only 14,458) (Table S2, Figure 3b). With regard to SSR frequency, the highest value was observed on chromosome 44 (2,496.67 loci/Mb), while the lowest was found on chromosome 1 (1,565.01 loci/Mb) (Table S2, Figure 3c). For SSR density, chromosome 44 also showed the highest level (192,933.04 bp/Mb), and chromosome 1 exhibited the lowest density (119,971.87 bp/Mb) (Table S2, Figure 3d).

Correlation analysis revealed that chromosome length was significantly positively correlated with SSR locus number, total SSR length, and SSR frequency ($P < 0.01$), with Pearson's correlation coefficients (r) of 0.974, 0.976, and 0.968, respectively. In contrast, no significant correlation was observed between chromosome length and SSR density. Among the four SSR-related indices (locus number, total length, frequency, and density) within individual chromosomes, only SSR locus number was extremely significantly positively correlated with total SSR length ($P < 0.01$), with a Pearson's correlation coefficient (r) of 0.942.

3.3. ANALYSIS OF SSR MARKERS IN *S. PARAMAMOSAIN*

A total of 2,575,599 SSR markers were developed, including 2,382,227 markers mapped to the 49 chromosomes and 193,372 unanchored markers that could not be assigned to specific chromosomes. On average, each chromosome contained 48,617 SSR markers, with a mean frequency of 1,687.72 loci/Mb (Table S4). Chromosome 3 harbored the largest number of SSR markers (99,894) (Figure 4, Table S4), whereas chromosome 44 contained the fewest (11,487) but exhibited the highest frequency (1,998.74 loci/Mb). Chromosome 1 displayed the lowest SSR marker frequency (1,270.07 loci/Mb) (Figure 4, Table S4). Among all developed SSR markers, 1,638,028 (63.60%) were classified as perfect SSR markers. For perfect SSR markers across the 49 chromosomes, mononucleotide were the most abundant, accounting for 30.25%–46.93%, followed by dinucleotide (12.71%–21.49%). The proportion of compound SSR markers varied from 33.65% to 37.90% across all chromosomes (Table S4).

A total of 167,441, 479,767, and 1,733,881 SSR markers were localized to exons, introns, and intergenic regions, accounting for 7.03%, 20.15%, and 72.82% of the total markers, respectively (Table S5). Across the 49 chromosomes, the proportion of SSR markers in exons ranged from 3.93% to 9.45%, while the proportions of intronic and intergenic SSR markers were 7.05%–39.99% and 55.26%–87.12%, respectively (Figure 5, Table S5). Chromosome 2 contained the largest number of exon-localized SSR markers (8,004), whereas chromosome 44 had the fewest (451) (Figure 5, Table S5).

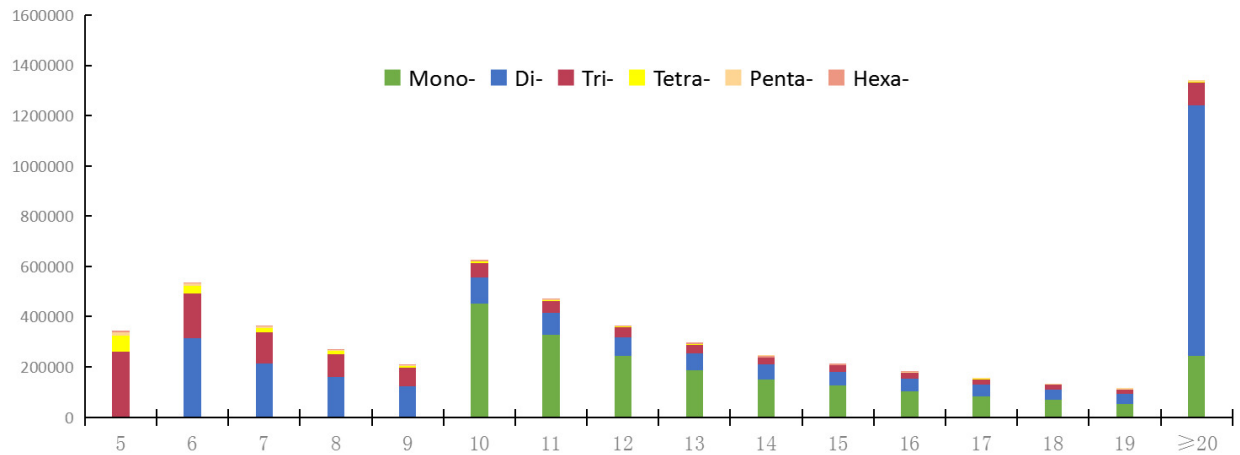


Figure 1. Distribution of all SSR motif repeat numbers in *S. paramamosain* genome

Table 1. Distribution of perfect SSRs across the *S. paramamosain* genome

Repeat type	Number	Proportion (%)	Frequency (Loci/Mb)
Mono-	1,037,673	51.92%	4298.13
Di-	643,927	32.22%	2667.20
Tri-	257,651	12.89%	1067.21
Tetra-	47,981	2.40%	198.74
Penta-	11,037	0.55%	45.72
Hexa-	457	0.02%	1.89
Total	1,998,726	100.00%	8278.90

Table 2. The most abundant SSR motifs in the *S. paramamosain* genome

Repeat type	Number/Loci	Frequency/(Loci/Mb)	Length (bp)	Density/(bp/Mb)
A	1,928,296	1,247.26	27,703,095	17,918.98
AG	1,129,336	730.48	52,100,292	33,699.62
AC	868,253	561.61	34,301,090	22,186.71
AT	427,746	276.68	24,490,982	15,841.31
AGG	246,396	159.37	6,872,001	4,444.96
AAT	239,814	155.12	7,940,709	5,136.23
ACT	218,900	141.59	7,916,031	5,120.26
ACC	143,467	92.80	3,198,900	2,069.12
C	112,871	73.01	1,446,066	935.35
AAC	109,905	71.09	2,970,861	1,921.62
AAAT	95,576	61.82	1,195,196	773.08

3.4. POLYMORPHISM INITIALLY EVALUATION OF SSR MARKERS

A total of 125 pairs of SSR primers were selected and synthesized. More than 40 pairs yielded clear and reproducible target bands, and 85 pairs produced non-specific band or failed to amplify (Table S1). The genetic variation of 32 wild *S. paramamosain* individuals was analyzed by 8 poly-

morphic SSR markers. A total of 78 alleles were detected. The *Na* ranged from 4 to 16 (mean = 9.75). The *Ne* varied from 1.8788 to 9.0660, with an average of 4.7069. The *Ho* ranged from 0.2903 to 0.9355 (mean = 0.6532), while the *He* varied from 0.3421 to 0.9043 (mean = 0.70). The *I* ranged from 0.6999 to 2.2917 (mean = 1.658), and the PIC ranged from 0.3195 to 0.88 (mean = 0.6638) (Table 3). Of the eight SSR loci analyzed, five loci were in HWE ($P > 0.05$), while

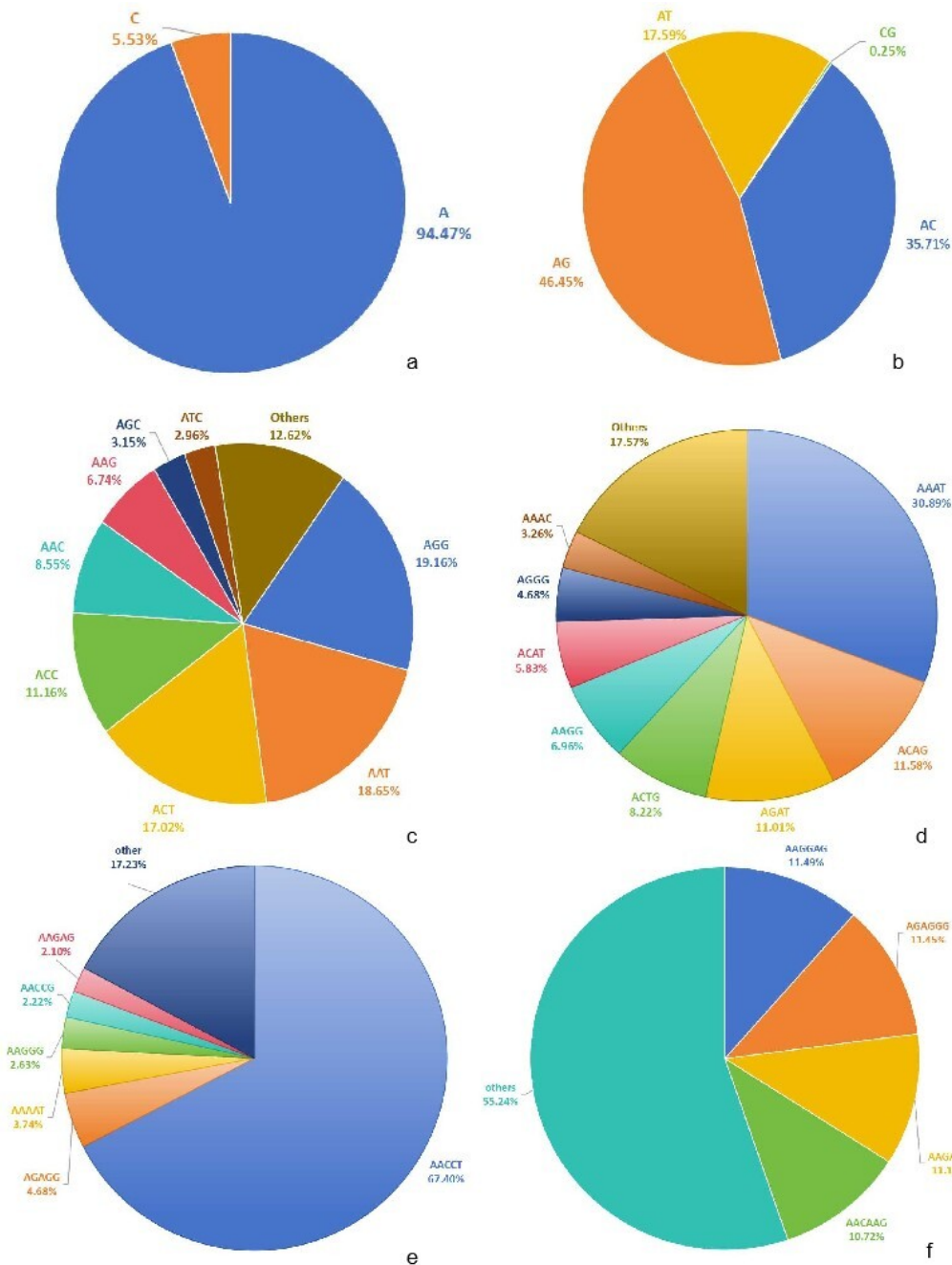


Figure 2. Distribution of SSR motifs across different repeat types in the *S. paramamosain* genome

three loci (LG1-1, LG1-5, and LG16-2) deviated significantly from HWE ($P < 0.05$). Only six pairs of locus showed significant LD: LG1-1 & qxLG3-50, LG1-3 & LG1-5, qxLG16-4 & qxLG16-2, LG1-3 & LG1-1, qxLG3-50 & qxLG3-67, and LG1-5 & LG1-1. The inbreeding coefficient (FIS) ranged from -0.2384 (LG3-50) to 0.3793 (LG1-5) (Table 3).

DISCUSSION

With the advancement of sequencing technologies and the reduction in their costs, genome sequencing and assembly have been successfully completed for numerous aquatic animal species.^{16,17} Leveraging genomic data, researchers can efficiently screen and develop a large number of SSR loci

and molecular markers for genetic research.^{13,18-20} However, SSR research on *S. paramamosain* remains relatively scarce, with only a limited number of SSR markers identified to date.²¹⁻²³ Zhao et al.² integrated PacBio long-read sequencing and Illumina paired-end sequencing data to generate the first high-quality, chromosome-level genome assembly of *S. paramamosain*. In the present study, comprehensive analysis of this chromosome-level genome assembly yielded extensive SSR data, which significantly expands the genomic resources available for this species. SSR markers have been widely used in marker-assisted selection and breeding. Specifically, their applications include: 1) precise mapping and screening of functional gene loci associated with economic traits; 2) improving the accuracy of selective breeding and reducing errors in phenotypic

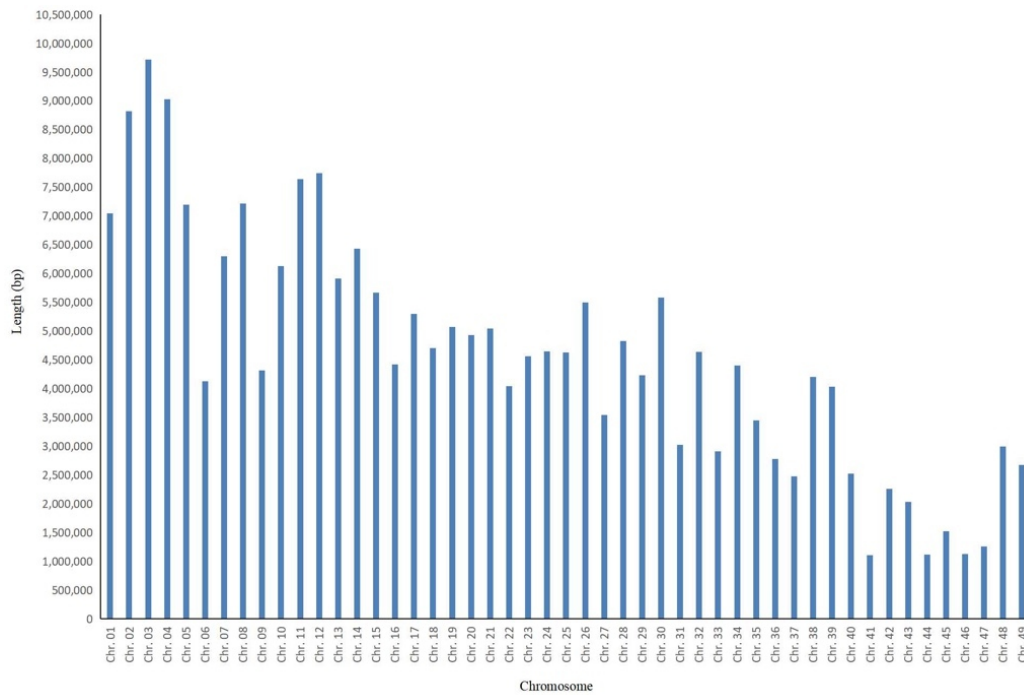


Figure 3. Chromosome-wide distribution of SSRs in the *S. paramamosain* genome: (a) total SSR length; (b) number of SSR loci; (c) SSR frequency; (d) SSR density

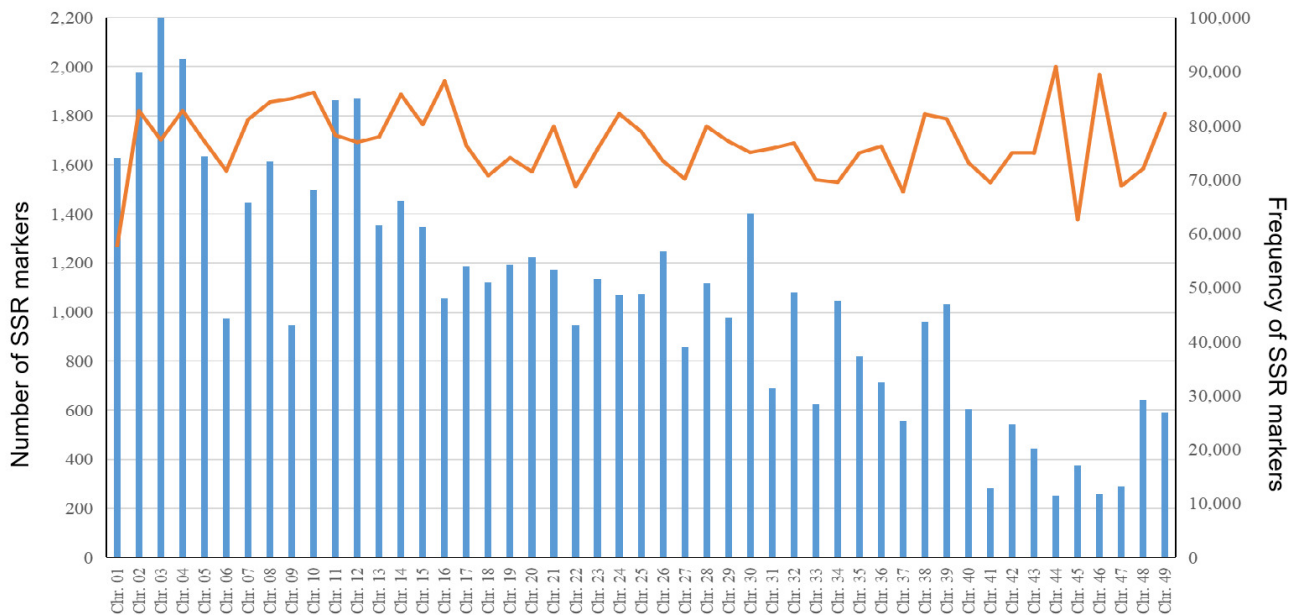


Figure 4. Distribution of the number (column chart) and frequency (line chart) of SSR markers on each chromosome of *S. paramamosain*

identification; 3) accurately constructing parentage determination systems and pedigree records to avoid inbreeding depression; and 4) analyzing the genetic structure of breeding populations and optimizing breeding strategies.²⁴

SSR markers are widely recognized as reliable molecular genetic tools, which have been extensively applied in studies of genetic diversity, population structure, and genetic breeding.⁵ Genome-wide SSR identification has been car-

ried out across a wide range of plant and animal species.^{11, 25} In this study, a total of 3,184,249 SSR loci were detected in the *S. paramamosain* genome, which was more than that in *Eriocheir sinensis* (886,782 perfect SSR)¹⁹ and less than that in *Litopenaeus vannamei* (9,123,781).²⁶ The total length of SSRs in *S. paramamosain* was 241.42 Mb, accounting for 15.62% of the entire genome. For context, SSRs constitute 23.93% of the *Penaeus vannamei* genome and over 10%

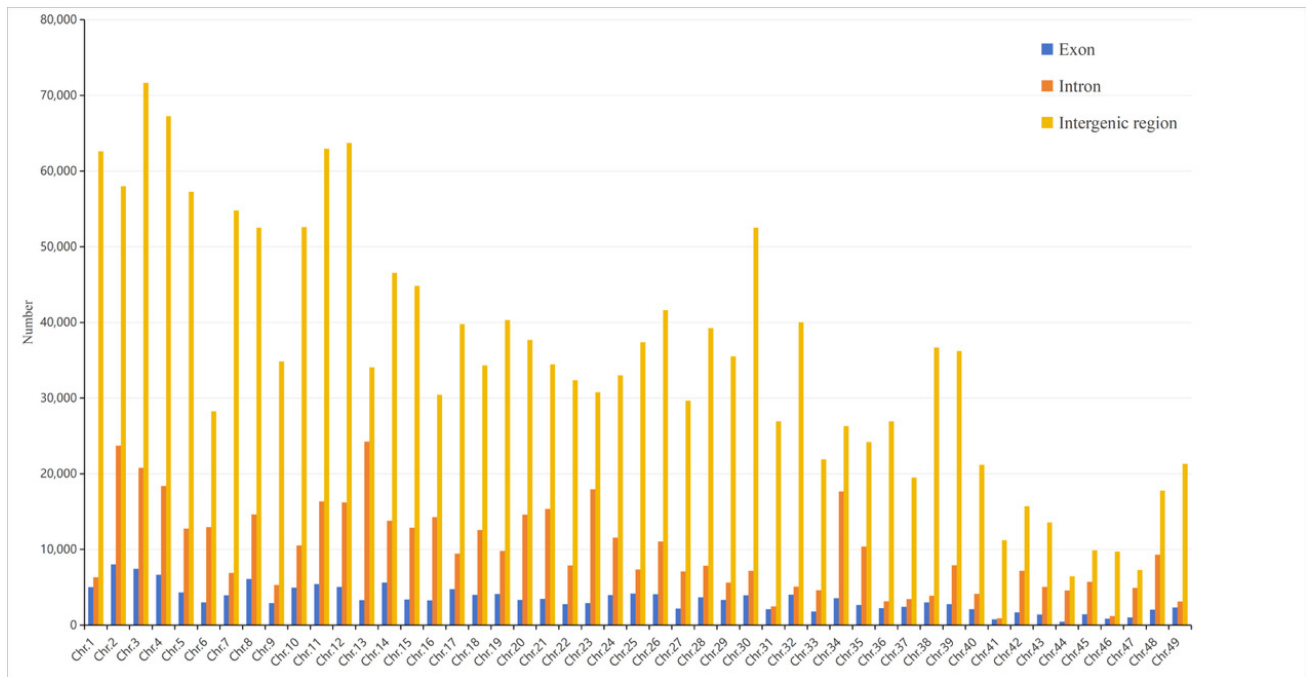


Figure 5. Distribution of SSR markers across different genomic regions on each chromosome of *S. paramamosain*

Table 3. Characteristics of 8 polymorphic SSR markers in *S. paramamosain*

Locus	<i>N_a</i>	<i>N_e</i>	<i>H_o</i>	<i>I</i>	PIC	FIS
LG1-3	13	8.0418	0.8387	2.2917	0.8640	0.0422
LG1-5	5	1.8788	0.2903	0.9577	0.4386	0.3793
LG1-1	4	1.5075	0.3226	0.6999	0.3195	0.0417
LG3-50	14	5.1946	0.9298	2.0717	0.7902	-0.2384
LG16-4	4	2.2401	0.4839	0.9682	0.4789	0.1259
LG8-22	15	9.0660	0.7097	2.4056	0.8800	0.2023
LG16-2	16	5.6199	0.9355	2.222	0.8097	-0.1380
LG3-67	7	4.1068	0.6452	1.6469	0.7292	0.1472
Mean	9.75	4.7069	0.6532	1.658	0.6638	
SD	5.2304	2.8253	0.2692	0.6901	0.2177	

N_a, observed number of alleles; *N_e*, effective number of alleles; *H_o*, observed heterozygosity; *H_e*, expected heterozygosity; *I*, Shannon's information index; *PIC*, polymorphism information content; *FIS*, fixation index; *SD*, standard deviation.

of the genomes of *P. monodon* and *Marsupenaeus japonicus*,²⁷ indicating that crustaceans generally harbor a high proportion of SSRs in their genomes. The SSR frequency and density in *S. paramamosain* were 2,059.64 loci/Mb and 156,158.13 bp/Mb, respectively, with an average inter-locus distance of 486 bp (Table S2). Collectively, these findings demonstrate that SSR loci are highly abundant in the *S. paramamosain* genome, highlighting the great potential of this species for the development of robust genetic markers.

Among the SSR repeat types identified in *S. paramamosain*, mononucleotide repeats were the most prevalent, followed by di-, tri-, tetra-, penta-, and hexanucleotide repeats. This distribution pattern is consistent with previous observations in *E. sinensis*,¹⁹ mosquitoes,²⁸ and several marine fish species.^{29,30} In contrast, dinucleotide repeats were more abundant than mononucleotide repeats in *Pelteobagrus vachelli*,¹⁴ *Scatophagus argus*,³¹ palmae,³² and

peanuts,²⁰ suggesting taxon-specific differences in SSR repeat type preferences. A motif was overwhelmingly dominant (94.47%), which may be attributed to the abundance of poly(A) tails.³³ The AG motif was the most common (46.45%), followed by AC (35.71%) in dinucleotide repeats. In *E. sinensis*, AC was the dominant dinucleotide, followed by AG.¹⁹ Among trinucleotide, AGG was the most prevalent in *S. paramamosain* and *E. sinensis*¹⁹; however, AAT was the dominant trinucleotide motif in some fish species and mammals.^{18,29,34} These interspecific differences are likely driven by genomic sequence variations, evolution, and selective pressures.³⁵

In this research, chromosome length was significantly positively correlated with the number, frequency, and length of SSR within each chromosome ($P < 0.01$). This result is consistent with previous findings in *Lateolabrax maculatus*,¹⁸ birds,³⁶ and mammals.^{34,37,38} This correlation

may be attributed to the “genome-scale dosage effect” of SSR distribution.³⁹ First, SSR loci are not randomly distributed across chromosomes but instead exhibit a “proportional distribution” relative to chromosome length; second, longer chromosomes provide a larger physical space and are associated with higher frequencies of mutational events (e.g., DNA replication errors, unequal crossing-over), which facilitate the accumulation of SSR loci and the extension of their repeat lengths.^{34,37} Notably, no significant correlation was observed between chromosome length and SSR density, suggesting that SSR distribution is not only independent of chromosome size but is also influenced by regional genomic stability and repeat tolerance.

In this study, most of SSR markers (90.97%) were located in non-coding regions (introns and intergenic regions), with only 7.03% in exons. This result is also found in mosquitoes,²⁸ birds,³⁶ and mammals.^{34,40} SSRs within exons are subject to intense evolutionary selective pressure. Expansion or contraction of SSR repeats may induce frameshift mutations or amino acid substitutions, which can subsequently impair gene function or result in complete loss of function.³⁵ By contrast, non-coding regions are subject to relatively weak functional constraints, as alterations in SSR copy number do not directly affect protein translation.⁶ Moreover, the fidelity of DNA repair processes is comparatively lower in non-coding regions, thereby increasing the likelihood that replication slippage events arise and become fixed within the genome.⁶ Certain non-coding SSRs may also confer adaptive advantages by modulating transcriptional activity, facilitating their evolutionary maintenance through positive or neutral selection.³³ For instance, a (CT)_n SSR located in the 5' UTR of the IRF gene has been shown to regulate the expression level of IRF in *L. vannamei*.⁴¹ Similarly, a CCTTT microsatellite near exon 3 of the glyA locus may affect the transcriptional activity of glyA in *Fenneropenaeus chinensis*.⁴²

The polymorphism of these SSR markers was preliminarily verified in the present study. Of the 125 SSR markers assessed, 40 (32%) produced clear and reproducible target amplicons. This proportion is consistent with that reported in *Xiphopenaeus kroyeri* (31.11%)⁴³ and higher than that observed in *Gammarus pulex* (15.71%).⁴⁴ The relatively low amplification rate observed in this study may be attributed to suboptimal primer design and quality. Due to length constraints imposed on target PCR products, the SSR-flanking sequences available for primer design were considerably restricted. Moreover, sequences containing poly runs, stable hairpin structures, or extreme AT/GC compositional bias tend to reduce the efficiency and success rate of PCR amplification. In addition, primers that produced nonspecific amplification products were excluded from further analysis.

Polymorphism rates of SSR markers in *S. paramamosain* were reported to be 38% and 22.91% based on polyacrylamide gel electrophoresis.^{22,23} In this study, capillary fluorescence electrophoresis was employed for genotyping, which provides high accuracy and sensitivity in discriminating allele size differences, thereby guaranteeing the reliability of the identified polymorphic markers.⁴⁵

Our PIC results indicated that five of the SSR markers displayed high polymorphism, while the remaining three exhibited moderate polymorphism. In addition, three SSRs showed significant deviations from HWE, which may be attributed to unusually high frequencies of particular alleles or an excess of homozygotes. LD analysis further suggested that these loci may be located within the same linkage group in *S. paramamosain*, or influenced by evolutionary forces including natural selection, population admixture, and inbreeding.

Most SSRs of $FIS > 0$, indicating a slight heterozygote deficit within the population. By contrast, FIS of two loci (LG3-50 and LG16-2) < 0 , implying heterozygote excess. All of results showed that no significant genetic subdivision was identified. Overall, the population exhibited a relatively homogeneous genetic structure, with only a small number of loci displaying minor genetic deviations, which may be attributed to inbreeding or selection effects.

In conclusion, this study identified and characterized a large number of genome-wide SSRs in *S. paramamosain*, which enriches the available genomic resources for this species. These findings will facilitate the resource management and genetic breeding of *S. paramamosain*.

ACKNOWLEDGMENTS

This research was supported by the Shrimp and Crab Industry Innovation Team Project of Guangdong Province (2024CXTD25), Basic and Applied Basic Research fund of Guangdong Province (2024A1515240032), Central Public-interest Scientific Institution Basal Research Fund, South China Sea Fisheries Research Institute, CAFS (NO.2024RC10), Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (SML2024SP003), the earmarked fund for CARS-48-21, Science and Technology Plan Project of Guangzhou (2024E04J0004 and 2025D04J0030), Science and Technology Planning Project of Guangdong Province, China (2023B1212060047), research on breeding technology of candidate species for Guangdong modern marine ranching (2024-MRB-00-001), and the Central PublicInterest Scientific Institution Basal Research Fund, CAFS (nos. 2023TD44 and 2021SD05).

AUTHORS' CONTRIBUTION

Conceptualization: Sigang Fan (Lead). Methodology: Sigang Fan (Equal), Changhong Cheng (Equal), Minze Liao (Equal). Formal Analysis: Sigang Fan (Equal), Hongling Ma (Equal), Guangxin Liu (Equal). Writing – original draft: Sigang Fan (Lead). Writing – review & editing: Zhixun Guo (Lead). Funding acquisition: Zhixun Guo (Lead).

SUPPLEMENTARY MATERIALS

Table S1 Microsatellite markers of *S. paramamosain* used in this study; Table S2 The character of SSRs in different chromosomes and unmapped chromosome sequences; Table S3 Summary of SSR motifs and repeats; Table S4 SSR markers

counts and proportion of five motifs in each chromosome of *S. paramamosain*; Table S5 The distribution of SSR markers in the different regions of the *S. paramamosain* genome.

COMPETING OF INTEREST – COPE

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

ETHICAL CONDUCT APPROVAL – IACUC

The material of this paper does not involve ethical conflicts. We got permission to take this sample. All experimental protocols and methods in this study were approved by the

Animal Care and Use Ethics Committee in the South China Sea Fisheries Research Institute (202100321m).

INFORMED CONSENT STATEMENT

All authors and institutions have confirmed this manuscript for publication.

DATA AVAILABILITY STATEMENT

Data will be made available on request.

Submitted: December 30, 2025 CDT. Accepted: April 15, 2026 CDT. Published: June 02, 2026 CDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

REFERENCES

1. Syafaat MN, Azra MN, Waiho K, et al. A Review of the Nursery Culture of Mud Crabs, Genus *Scylla*: Current Progress and Future Directions. *Animals*. 2021;11(7):2034. doi:[10.3390/ani11072034](https://doi.org/10.3390/ani11072034)
2. Zhao M, Wang W, Zhang F, et al. A chromosome-level genome of the mud crab (*Scylla paramamosain estampador*) provides insights into the evolution of chemical and light perception in this crustacean. *Mol Ecol Resour*. 2021;21(4):1299-1317. doi:[10.1111/1755-0998.13332](https://doi.org/10.1111/1755-0998.13332)
3. Tian Y, Jie YK, Zeng XB, et al. Effect of density on antioxidant and nonspecific immunity of mud crab (*Scylla paramamosain*). *South China Fish Sci*. 2023;19(3):60-67. doi:[10.12131/20220312](https://doi.org/10.12131/20220312)
4. Fishery administration of Ministry of Agriculture and Rural Affairs. *China Fishery Statistics Yearbook*. China Agricultural Press; 2025:22.
5. Wenne R. Microsatellites as Molecular Markers with Applications in Exploitation and Conservation of Aquatic Animal Populations. *Genes*. 2023;14(4):808. doi:[10.3390/genes14040808](https://doi.org/10.3390/genes14040808)
6. Fan H, Chu JY. A brief review of short tandem repeat mutation. *Genom Proteom Bioinform*. 2007;5(1):7-14. doi:[10.1016/S1672-0229\(07\)60009-6](https://doi.org/10.1016/S1672-0229(07)60009-6)
7. Huang H, Fan S, Wang P, et al. Genetic diversity analysis of six geographical populations of *Lateolabrax maculatus* based on microsatellite markers. *South China Fish Sci*. 2022;18(1):99-106. doi:[10.12131/20210126](https://doi.org/10.12131/20210126)
8. Lu HY, Tian HF, Hu QM, Li Z. Parentage assignment of *Monopterus albus* using multiplex PCR of microsatellites. *South China Fish Sci*. 2023;19(6):97-106. doi:[10.12131/20230081](https://doi.org/10.12131/20230081)
9. Singh RB, Mahenderakar MD, Jugran AK, Singh RK, Srivastava RK. Assessing genetic diversity and population structure of sugarcane cultivars, progenitor species and genera using SSR markers. *Gene*. 2020;753:144800. doi:[10.1016/j.gene.2020.144800](https://doi.org/10.1016/j.gene.2020.144800)
10. Kawakatsu K, Yagi M, Harada T, et al. Development of an SSR marker-based genetic linkage map and identification of a QTL associated with flowering time in *Eustoma*. *Breed Sci*. 2021;71(3):344-353. doi:[10.1270/jsbbs.20100](https://doi.org/10.1270/jsbbs.20100)
11. Yang ZL, Tang MJ, Zhang DD, Chen XW, Wu XG. Genetic diversity analysis of four consecutive selective breeding generations in even years of white-carapaced strain of *Eriocheir sinensis* based on SSR markers. *South China Fish Sci*. Published online 2025. doi:[10.12131/20250054](https://doi.org/10.12131/20250054)
12. Cui H, Ma H, Ma L, Ma C, Ma Q. Development of eighteen polymorphic microsatellite markers in *Scylla paramamosain* by 5' anchored PCR technique. *Mol Biol Rep*. 2011;38(8):4999-5002. doi:[10.1007/s11033-010-0645-6](https://doi.org/10.1007/s11033-010-0645-6)
13. Luo X, Zhang L, Chen S. Microsatellite Genome-Wide Database Development for the Commercial Blackhead Seabream (*Acanthopagrus schlegelii*). *Genes (Basel)*. 2023;14(3):620. doi:[10.3390/genes14030620](https://doi.org/10.3390/genes14030620)
14. Peng Y, Li J, Wang T, et al. Preliminary study on distribution characteristics and positioning of microsatellites in whole genome of *Pelteobagrus vachelli*. *South China Fish Sci*. 2022;18(1):90-98. doi:[10.12131/20210168](https://doi.org/10.12131/20210168)
15. Ma H, Huang T, Liu E, Wang G, Gu W, Xu G. Development of polymorphic simple sequences repeats markers from whole gene resequencing data comparison of 68 *Oncorhynchus mykiss*. *Front Mar Sci*. 2024;11:1375524. doi:[10.3389/fmars.2024.1375524](https://doi.org/10.3389/fmars.2024.1375524)
16. Zhou Q, Wang J, Li J, et al. Decoding the fish genome opens a new era in important trait research and molecular breeding in China. *Sci China Life Sci*. 2024;67(10):2064-2083. doi:[10.1007/s11427-023-2670-5](https://doi.org/10.1007/s11427-023-2670-5)
17. Ouyang Y, Pan JM, Xian L, et al. Chromosome-level genome and characteristic analysis of *Platax teira*. *South China Fish Sci*. 2024;20(6):31-42. doi:[10.12131/20240112](https://doi.org/10.12131/20240112)
18. Fan SG, Huang H, Liu Y, et al. Genome-wide development of polymorphic microsatellite markers and their application in spotted sea bass (*Lateolabrax maculatus*). *Aquac Rep*. 2021;20:100677. doi:[10.1016/j.aqrep.2021.100677](https://doi.org/10.1016/j.aqrep.2021.100677)
19. Xu JJ, Bi YH, Cheng JH, et al. Study on distribution characteristics of whole genome microsatellite of *Eriocheir sinensis*. *Genom Appl Biol*. 2021;40(7-8 combined issue):2422-2429. doi:[10.13417/j.gab.040.002422](https://doi.org/10.13417/j.gab.040.002422)

20. Ma J, Zhao Y, Chen H, et al. Genome-wide development of polymorphic microsatellite markers and their application in peanut breeding program. *Electron J Biotechnol.* 2020;44:25-32. doi:[10.1016/j.ejbt.2020.01.004](https://doi.org/10.1016/j.ejbt.2020.01.004)
21. Xu XJ, Wang GZ, Wang KJ, Li SJ. Isolation and characterization of ten new polymorphic microsatellite loci in the mud crab, *Scylla paramamosain*. *Conserv Genet.* 2009;10:1877-1878. doi:[10.1007/s10592-009-9843-y](https://doi.org/10.1007/s10592-009-9843-y)
22. Ma H, Ma C, Li S, et al. Transcriptome analysis of the mud crab (*Scylla paramamosain*) by 454 deep sequencing: assembly, annotation, and marker discovery. *PLoS One.* 2014;9(7):e102668. doi:[10.1371/journal.pone.0102668](https://doi.org/10.1371/journal.pone.0102668)
23. Ma H, Jiang W, Liu P, et al. Identification of transcriptome-derived microsatellite markers and their association with the growth performance of the mud crab (*Scylla paramamosain*). *PLoS One.* 2014;9(2):e89134. doi:[10.1371/journal.pone.0089134](https://doi.org/10.1371/journal.pone.0089134)
24. Sharma K, Yashveer S, Singh V, Sharma S, Redhu M, Rahimi M. Microsatellite marker-assisted backcross breeding for improvement of wheat salt tolerance using Kharchia 65. *BMC Genomics.* 2024;25(1):550. doi:[10.1186/s12864-024-10468-y](https://doi.org/10.1186/s12864-024-10468-y)
25. Xie YJ, Su MX, Gao H, et al. SSR marker-based genetic diversity and structure analyses of *Camellia nitidissima* var. *phaeopubisperma* from different populations. *PeerJ.* 2025;13:e18845. doi:[10.7717/peerj.18845](https://doi.org/10.7717/peerj.18845)
26. Mae D, Noh E, Lee M, Bae S, Hwang I, Kang J. Development of Genome-Based Microsatellite Markers for Whiteleg Shrimp *Litopenaeus vannamei*. *Korean J Fish Aquat Sci.* 2025;58(4):273-279. doi:[10.5657/KFAS.2025.0273](https://doi.org/10.5657/KFAS.2025.0273)
27. Yuan JB, Zhang XJ, Liu CZ, et al. Genomic resources and comparative analyses of two economical penaeid shrimp species, *Marsupenaeus japonicus* and *Penaeus monodon*. *Mari Genom.* 2018;39:22-25. doi:[10.1016/j.margen.2017.12.006](https://doi.org/10.1016/j.margen.2017.12.006)
28. Wang XT, Zhang YJ, Qiao L, Chen B. Comparative analyses of simple sequence repeats (SSRs) in 23 mosquito species genomes: Identification, characterization and distribution (Diptera: Culicidae). *Insect Sci.* 2019;26(4):607-619. doi:[10.1111/1744-7917.12577](https://doi.org/10.1111/1744-7917.12577)
29. Lei Y, Zhou Y, Price M, Song Z. Genome-wide characterization of microsatellite DNA in fishes: survey and analysis of their abundance and frequency in genome-specific regions. *BMC genom.* 2021;22(1):421. doi:[10.1186/s12864-021-07752-6](https://doi.org/10.1186/s12864-021-07752-6)
30. Tian HF, Hu QM, Li Z. Genome-wide identification of simple sequence repeats and development of polymorphic SSR markers in swamp eel (*Monopterus albus*). *Sci Prog.* 2021;104(3):368504211035597. doi:[10.1177/00368504211035597](https://doi.org/10.1177/00368504211035597)
31. Yang YR, Yang W, Ren XL, et al. Distribution Patterns of Microsatellites and Development of Polymorphic Markers from *Scatophagus argus* Genome. *Journal of Guangdong Ocean University.* 2020;40(04):7-14. doi:[10.3969/j.issn.1673-9159.2020.04.002](https://doi.org/10.3969/j.issn.1673-9159.2020.04.002)
32. Manee MM, Al-Shomrani BM, Al-Fageeh MB. Genome-wide characterization of simple sequence repeats in Palmae genomes. *Genes Genom.* 2020;42(5):597-608. doi:[10.1007/s13258-020-00924-w](https://doi.org/10.1007/s13258-020-00924-w)
33. Bagshaw ATM. Functional Mechanisms of Microsatellite DNA in Eukaryotic Genomes. *Genome Biol Evol.* 2017;9(9):2428-2443. doi:[10.1093/gbe/evx164](https://doi.org/10.1093/gbe/evx164)
34. Qi WH, Jiang XM, Du LM, et al. Genome-Wide Survey and Analysis of Microsatellite Sequences in Bovid Species. *PLoS One.* 2015;10(7):e0133667. doi:[10.1371/journal.pone.0133667](https://doi.org/10.1371/journal.pone.0133667)
35. Sjakste T, Paramonova N, Sjakste N. Functional significance of microsatellite markers. *Medicina (Kaunas).* 2013;49(12):505-509.
36. Huang J, Li W, Jian Z, Yue B, Yan Y. Genome-wide distribution and organization of microsatellites in six species of birds. *Biochem Syst Ecol.* 2016;67:95-102. doi:[10.1016/j.bse.2016.05.023](https://doi.org/10.1016/j.bse.2016.05.023)
37. Subramanian S, Mishra RK, Singh L. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol.* 2003;4(2):R13. doi:[10.1186/gb-2003-4-2-r13](https://doi.org/10.1186/gb-2003-4-2-r13)
38. Xu Y, Hu Z, Wang C, Zhang X, Li J, Yue B. Characterization of perfect microsatellite based on genome-wide and chromosome level in Rhesus monkey (*Macaca mulatta*). *Gene.* 2016;592(2):269-275. doi:[10.1016/j.gene.2016.07.016](https://doi.org/10.1016/j.gene.2016.07.016)
39. Garner TW. Genome size and microsatellites: the effect of nuclear size on amplification potential. *Genome.* 2002;45(1):212-215. doi:[10.1139/g01-113](https://doi.org/10.1139/g01-113)
40. Manee MM, Algarni AT, Alharbi SN, et al. Genome-wide characterization and analysis of microsatellite sequences in camelid species. *Mamm Res.* 2020;65(2):359-373. doi:[10.1007/s13364-019-00458-x](https://doi.org/10.1007/s13364-019-00458-x)

41. Yin B, Wang H, Zhu P, Weng S, He J, Li C. A Polymorphic (CT)_n SSR Influences the Activity of the *Litopenaeus vannamei* IRF Gene Implicated in Viral Resistance. *Front Genet.* 2019;10:1257. doi:[10.3389/fgene.2019.01257](https://doi.org/10.3389/fgene.2019.01257). PMID:31921300
42. Yuan J, Zhang X, Wang M, et al. Simple sequence repeats drive genome plasticity and promote adaptive evolution in penaeid shrimp. *Commun Biol.* 2021;4(1):186. doi:[10.1038/s42003-021-01716-y](https://doi.org/10.1038/s42003-021-01716-y)
43. Kerkhove TRH, Hellemans B, De Troch M, De Backer A, Volckaert FAM. Isolation and characterisation of 14 novel microsatellite markers through Next Generation Sequencing for the commercial Atlantic seabob shrimp *Xiphopenaeus kroyeri*. *Mol Biol Rep.* 2019;46(6):6565-6569. doi:[10.1007/s11033-019-05026-9](https://doi.org/10.1007/s11033-019-05026-9)
44. Švara V, Norf H, Luckenbach T, Brack W, Michalski SG. Isolation and characterization of eleven novel microsatellite markers for fine-scale population genetic analyses of *Gammarus pulex* (Crustacea: Amphipoda). *Mol Biol Rep.* 2019;46(6):6609-6615. doi:[10.1007/s11033-019-05077-y](https://doi.org/10.1007/s11033-019-05077-y)
45. Navals P, Rangaswamy AMM, Kasyanchyk P, Berezovski MV, Keillor JW. Conformational Modulation of Tissue Transglutaminase via Active Site Thiol Alkylating Agents: Size Does Not Matter. *Biomolecules.* 2024;14(4):496. doi:[10.3390/biom14040496](https://doi.org/10.3390/biom14040496)

SUPPLEMENTARY MATERIALS

Supplementary Materials

Download: <https://ija.scholasticahq.com/article/161390-characterization-and-development-of-genome-wide-microsatellites-from-mud-crab-scylla-paramamosain/attachment/341917.zip>
