We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,200
Open access books available

116,000
International authors and editors

125M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the top 1% most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Microsatellite Markers in the Mud Crab (*Scylla paramamosain*) and their Application in Population Genetics and Marker-Assisted Selection

Hongyu Ma, Chunyan Ma, Lingbo Ma, Xincang Li and Yuanyou Li

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/65041

**Abstract**

The mud crab (*Scylla paramamosain*) is a commercially important species for aquaculture and fisheries in China. In this study, a total of 302 polymorphic microsatellite markers have been isolated and characterized. The observed and expected heterozygosity ranged from 0.04 to 1.00 and from 0.04 to 0.96 per locus. The wild populations distributed along South-eastern China coasts showed high genetic diversity (*H* \(O\) ranged from 0.62 to 0.77) and low genetic differentiation (*F* \(st\) = 0.018). Meanwhile, a significant association (\(r^2 = 0.11\)) was identified between genetic and geographic distance of 11 locations. Furthermore, a PCR-based parentage assignment method was successfully developed using seven polymorphic microsatellite loci that could correctly assign 95% of the progeny to their parents. Moreover, three polymorphic microsatellite loci were identified to be significantly associated with 12 growth traits of *S. paramamosain*, and four genotypes were considered to be great potential for marker-assisted selection. Finally, a first preliminary genetic linkage map with 65 linkage groups and 212 molecular markers was constructed using microsatellite and AFLP markers for *S. paramamosain*. This map was 2746 cM in length, and covered approximately 50% of the estimated genome. This study provides novel insights into genome biology and molecular marker-assisted selection for *S. paramamosain*.

**Keywords:** genetic linkage map, growth traits linked loci, microsatellite marker, parentage assignment, population genetic diversity
1. Introduction

The mud crab (*Scylla paramamosain*), a big size crustacean, is one of the most important aquaculture species and marine fisheries in China. It is naturally distributed along the coasts of South-eastern China, as well as other East and Southeast Asian countries. Due to the good flavor, high nutrition value, and fast growth speed, *S. paramamosain* is now becoming more and more popular in the above countries. In China, the artificial culture of this crab can date back more than 100 years [1], and in recent years, the culture production is usually above 100,000 tons per year [2]. However, the current culture capacity cannot meet the demands of market. Under the natural conditions, mature *S. paramamosain* mates inshore, then the gravid females migrate offshore for spawning eggs, and finally the offspring return to inshore. Nowadays, the wild resource of this crab including adults and larvae has been decreasing quickly due to seawater pollution and overexploitation. Therefore, for conservation and sustainable utilization of this valuable marine resource, we first need to better understand its population genetic diversity and improve economic traits. Molecular marker-assisted selection (MAS) is thought to be a good method for genetic improvement, because it can shorten the selection period, and increase the accuracy of improvement. Of the many known molecular markers, microsatellite marker is an ideal genetic tool for helping to fulfill this purpose.

Microsatellites, normally known as simple sequence repeats (SSR), are widely used for population structure analysis [3, 4], parentage assignment [5], genetic map construction [6, 7], and marker-assisted selection [8, 9] because they are abundantly distributed throughout genome, codominant, and hyper-variable in most eukaryotic organism genomes. Before this study, there were only 15 polymorphic microsatellite loci available for *S. paramamosain* [10, 11]. The lack of efficient microsatellite markers has severely blocked the genetic studies in *S. paramamosain*.

The purpose of this study is to massively develop polymorphic microsatellite markers, uncover the population genetic diversity, create molecular parentage assignment technique, identify growth performance-associated markers, and construct a genetic linkage map, so as to provide novel insights into population genetic diversity and genetic improvement of economic traits for *S. paramamosain*.

2. Microsatellites and their application in population genetics and MAS

2.1. Material and methods

For microsatellite loci isolation, a total of six different strategies based on PIMA [12], FIASCO [13], GenBank-derived genes [14], 5’ anchored PCR [15], cDNA library [16], and 454 sequencing transcriptome [17, 18] have been employed to isolate microsatellite markers. The polymorphisms of microsatellite loci were evaluated by using a wild population with approximately 30 individuals.
For population genetics analysis, a total of 397 wild individuals were sampled from 11 locations (Sanmen, Ningde, Zhangzhou, Shantou, Shenzhen, Zhanjiang, Haikou, Wenchang, Wanning, Dongfang, and Danzhou) of South-eastern coasts of China. Nine polymorphic microsatellite markers were genotyped in these specimens [19].

For development of parentage assignment technique, four G1 families were collected, with 46 progenies in each family. Family 1 lost both parents information, and families 2, 3, and 4 only had maternal information. Ten polymorphic microsatellite loci were selected for genotyping the above crabs [20].

For trait-marker association analysis, a total of 96 three-month-old full-sib specimens were randomly sampled from a G1 family. Sixteen growth traits including carapace length (CL), internal carapace width (ICW), carapace width (CW), body height (BH), carapace frontal width (CFW), carapace width at spine 8 (CWS8), abdomen width (AW), fixed finger length of the claw (FFLC), fixed finger width of the claw (FFWC), fixed finger height of the claw (FFHC), distance between lateral spine 1 (DLS1), distance between lateral spine 2 (DLS2), meropodite length of pereiopod 1 (MLP1), meropodite length of pereiopod 2 (MLP2), meropodite length of pereiopod 3 (MLP3), and body weight (BW) were measured. Moreover, 129 transcriptome-derived polymorphic microsatellite loci were genotyped in these animals [17].

For genetic linkage map construction, a G1 family with 95 individuals was selected as mapping population. Microsatellite and AFLP markers were employed for linkage analysis. A total of 337 polymorphic microsatellite markers and 64 AFLP selective primer combinations were used in this study [21].

For data analysis, the observed (\(N_a\)) and effective (\(N_e\)) number of alleles, the observed (\(H_O\)) and expected (\(H_E\)) heterozygosity, Hardy-Weinberg equilibrium (HWE), and linkage disequilibrium (LD) were calculated by softwares POPGENE 1.31 [22] and ARLEQUIN 3.01 [23]. Significance values for multiple tests were corrected by sequential Bonferroni procedure [24]. The null allele frequency was predicted by MICRO-CHECKER 2.2.3 [25]. Genetic differentiation was tested by AMOVA analysis using software GENAIEX 6.41 [26]. The UPGMA tree among 11 locations was constructed by software MEGA 4.0 [27]. The links between genetic distance and geographic distance were evaluated by Mantel test [28]. The exclusion probability of microsatellite loci and parentage assignment were carried out using software CERVUS 3.0 [29]. Double-blind test was performed using seven most informative microsatellites and 40 specimens. The UPGMA tree of these 40 individuals was created by software MEGA 4.0. The general linear model (GLM) was used to identify the association between microsatellite loci and growth performance. A linear animal model with the fixed effects was used as follows: 

\[ Y_{ijk} = \mu + G_i + S_j + e_{ijk}, \]

where \(Y_{ijk}\) is the observed value of the \(ijth\) trait; \(\mu\) is the mean value of the trait; \(G_i\) is the effect of the \(ith\) genotype; \(S_j\) is the effect of the \(jth\) sex; and \(e_{ijk}\) is the random error effect. Significant differences in growth traits among the different genotypes were calculated through multiple comparison analysis using the S-N-K method. The software JoinMap 3.0 [30] was used for linkage analysis of microsatellite and AFLP genotypes. The population type was defined as cross-pollination (CP). A critical logarithm of odds (LOD) score threshold ≥2.5 was referenced for markers assignment. Linkage groups were drawn by MapChart 2.1 software [31]. The expected genome size (\(G_e\)) was estimated using the formula: 

\[ G_e = \frac{(G_1 + G_2)}{2}. \]

The
expected genome size is the sum of the revised lengths of all linkage groups [33]. The observed map length \( G_{oa} \) was the total length of groups, triplets, and doublets. The estimated coverage of the genome \( C_{oa} \) was calculated as: \( G_{oa}/G_e \), accordingly.

2.2. Results and discussion

2.2.1. Isolation and characterization of microsatellite loci

In this study, a total of 302 polymorphic microsatellite markers were successfully developed by using six different strategies. For methods based on PIMA, FIASCO, GenBank-derived genes, 5’ anchored PCR, cDNA library, and 454 sequencing transcriptome, a total of 12, 54, 18, 18, 36, and 164 polymorphic microsatellite loci were identified, respectively. A total of 1858 alleles were detected with an average of 6.15 alleles per locus from these microsatellite markers. The observed and expected heterozygosity ranged from 0.04 to 1.00, and from 0.04 to 0.96 per locus, respectively. The genotype proportions at 45 microsatellite loci significantly deviated from Hardy-Weinberg equilibrium expectations after Bonferroni correction; this could be due to the small sample size or the presence of null alleles, but cannot be attributed to technical or statistical artifacts. No significant linkage disequilibrium was detected between these loci pairs. According to the utilities for comparative mapping, molecular markers are classified into two types: type I markers are linked with genes of known functions, while type II markers are linked with anonymous genomic fragments. Among these 302 microsatellite loci, 218 may be associated with functional genes, which were classified as type I markers that are usually considered to have lower polymorphic level than type II markers. In this study, the genetic diversity level of type I loci was slightly lower than that of genome-derived loci too. Moreover, the polymorphisms of type II loci isolated in this study were lower than those described in previous references [10, 11].

2.2.2. Population genetic diversity and differentiation

The population genetic diversity of \( S. \) paramamosain distributed along South-eastern coasts of China was found to be high by nine microsatellite markers. A total of 104 alleles were observed at these nine microsatellite loci, with an average of 11.6 alleles per locus. The \( H_o \) values ranged from 0.32 to 1.00 per locus-location combination, and from 0.62 to 0.77 per location. This result was in accordance with the previous study by mtDNA marker [34]. Three factors are usually thought to be associated with high genetic variation of marine animals: environmental heterogeneity, the life history characteristics, and large population size [35, 36]. Further, we determined that the genetic diversity level of \( S. \) paramamosain population was gradually increased from the Northern location to the South. This interesting trend was also found in previous study by mtDNA marker [34].

Approximately 98.2% of variance was within locations and 1.8% of that was among locations, which indicated that the population genetic variation mainly existed within locations, and the genetic differentiation level was very low among locations (Table 1).
### Table 1. AMOVA design and results in 11 locations of *Scylla paramamosain*.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>( F_{ST} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among locations</td>
<td>10</td>
<td>83.706</td>
<td>0.064</td>
<td>1.83</td>
<td>0.0183</td>
</tr>
<tr>
<td>Among individuals within locations</td>
<td>386</td>
<td>1465.290</td>
<td>0.388</td>
<td>11.18</td>
<td></td>
</tr>
<tr>
<td>Within individuals</td>
<td>397</td>
<td>1199.000</td>
<td>3.020</td>
<td>86.99</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>793</td>
<td>2747.996</td>
<td>3.472</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

The mtDNA data also indicated that the population genetic structure of this crab was genetically homogeneous [37]. Therefore, we concluded that *S. paramamosain* population distributed along South-eastern coasts of China is a single genetically homogeneous population with low differentiation. Moreover, from the UPGMA tree (Figure 1), we observed that there were totally

![Figure 1](http://dx.doi.org/10.5772/65041)

**Figure 1.** The UPGMA tree constructed among 11 locations of *Scylla paramamosain*. SM, Sanmen; ND, Ningde; ZZ, Zhangzhou; ST, Shantou; SZ, Shenzhen; ZJ, Zhanjiang; HK, Haikou; WC, Wenchang; WN, Wanning; DF, Dongfang; DZ, Danzhou.
two groups: one consisted of 10 locations and the other one contained only one location (Sanmen). Mantel tests showed a significantly positive link between pairwise $F_{ST}/(1 - F_{ST})$ and natural logarithm of geographic distance (km) (Figure 2).

![Figure 2](image)

**Figure 2.** Relationship estimated between genetic differentiation and geographic distance (km) among 11 locations of *Scylla paramamosain*.

### 2.2.3. Parentage assignment technique based on PCR

In this study, 10 polymorphic microsatellite loci produced 1870 genotypes in 184 offspring and three parents. The genetic diversity indexes showed a relative high variation of these individuals, with $H_O$ and PIC values ranging from 0.38 to 0.99 and from 0.44 to 0.75, respectively. Two loci deviated from HWE in family 1 and 4, four loci in family 2, and six loci in family 3. According to Mendelian inheritance principle, the genotypes of all parents were successfully deduced based on genotypes of offspring, suggesting microsatellites are ideal molecular markers for evaluating genetic relationship among different individuals. In panda and tiger, microsatellite loci were also successfully used to identify the relationship among different specimens [38, 39].

Furthermore, the exclusion probability was used to distinguish the pedigree relationship of *S. paramamosain* individuals. The PIC value was found to be associated with the exclusion probability: when the PIC value went up, the exclusion probability increased accordingly. Moreover, the combined exclusion probability was observed higher than single locus in this study (Figure 3). Ten microsatellite loci had 97% exclusion probability under without parent information conditions. The assignment success rate reached to 100% when seven microsatellite markers were combined together under the condition of no any parent information. Practical application showed that seven microsatellite loci combination could accurately assign 95% of the offspring to right parents (Figure 4).
2.2.4. Identification of growth performance related microsatellite markers

In aquatic animals, a set of microsatellite markers were identified to link to growth traits [40, 41]. In this study, of 129 polymorphic microsatellite loci, 30 showed polymorphisms in the
experimental G family. Statistical analysis indicated that three markers (Scpa36, Scpa75, and Spm30) were significantly linked to 12 growth traits (CL, BH, ICW, AW, FFWC, FFLC, FFHC, CWS8, MLP2, MLP3, DLS2, and BW) in *S. paramamosain*. Microsatellite marker Scpa36 was significantly associated with growth traits CL, FFLC, AW, MLP2, BH, FFHC, and MLP3. Out of four genotypes AB, BB, BC, and AC at this locus, the genotype BC had the highest potential for artificial selection (Table 2). Microsatellite marker Scpa75 was significantly linked to growth traits ICW, DLS2, MLP3, AW, and CWS8. Among four genotypes AD, AC, BC, and BD at this locus, the genotypes BC and BD showed the highest correlation rate with growth traits in *S. paramamosain* (Table 3). Locus Spm30 was significantly associated with three traits BH, BW, and DLS2. Multiple comparison analysis indicated that genotypes AC and BD at this locus had the highest association degree with growth traits (Table 4). Moreover, growth traits CW, CFW, DLS1, and MLP1 were not found to link to any microsatellite marker in this study.

### Table 2. Association analysis performed between microsatellite locus Scpa36 and eight growth traits of *Scylla paramamosain*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genotype</th>
<th>Number</th>
<th>Growth trait (Means ± SD, mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>CL</td>
</tr>
<tr>
<td>Scpa36</td>
<td>AB</td>
<td>19</td>
<td>47.87 ± 7.88a</td>
</tr>
<tr>
<td></td>
<td>BB</td>
<td>21</td>
<td>49.80 ± 7.85ab</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>23</td>
<td>52.50 ± 8.83ab</td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>21</td>
<td>54.46 ± 5.01bc</td>
</tr>
</tbody>
</table>

CL, carapace length; AW, abdomen width; BH, body height; FFLC, fixed finger length of the claw; FFHC, fixed finger height of the claw; MLP2, meropodite length of pereiopod 2; MLP3, meropodite length of pereiopod 3.

### Table 3. Association analysis performed between microsatellite locus Scpa75 and five growth traits of *Scylla paramamosain*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genotype</th>
<th>Number</th>
<th>Growth trait (Means ± SD, mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>ICW</td>
</tr>
<tr>
<td>Scpa75</td>
<td>AD</td>
<td>32</td>
<td>69.79 ± 11.83b</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>17</td>
<td>74.70 ± 11.75b</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>24</td>
<td>76.82 ± 9.99b</td>
</tr>
<tr>
<td></td>
<td>BC</td>
<td>14</td>
<td>77.59 ± 7.34b</td>
</tr>
</tbody>
</table>

ICW, internal carapace width; AW, abdomen width; CWS8, carapace width at spine 8; DLS2, distance between lateral spine 2; MLP3, meropodite length of pereiopod 3.

a, b, and c; the superscripts in different genotypes without a common one are significant different (*P* < 0.05).
Among 16 growth traits tested in this study, traits AW and MLP3 were associated with two loci Scpa36 and Scpa75, and traits BH and DLS2 were associated with two loci Scpa75 and Spm30. Meanwhile, traits CL, FFWC, FFLC, ICW, MLP2, BW, and CWS8 were associated with only one microsatellite locus. It is considered to be a common event that one locus contributes to several quantitative traits and/or several different loci influence a same quantitative trait [41, 42]. In the next artificial breeding program, these three microsatellite markers should be first considered for marker-assisted selection of *S. paramamosain*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Genotype</th>
<th>Number</th>
<th>Growth trait (Means ± SD, mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>BH</td>
</tr>
<tr>
<td>Spm30</td>
<td>CD</td>
<td>18</td>
<td>27.34 ± 4.47&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>AB</td>
<td>24</td>
<td>29.52 ± 5.24&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>BD</td>
<td>33</td>
<td>30.88 ± 4.22&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>21</td>
<td>31.08 ± 3.58&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

BH, body height; DLS2, distance between lateral spine 2; BW, body weight. <sup>a</sup> and <sup>b</sup>, the superscripts in different genotypes without a common one are significant different (*P* < 0.05).

Table 4. Association analysis performed between microsatellite locus Spm30 and three growth traits of *Scylla paramamosain*.

2.2.5. Construction of genetic linkage map

Of 337 microsatellite markers, 118 segregated from parents to offspring of *S. paramamosain* with a rate of 35%. Meanwhile, 64 AFLP selective primer combinations produced 574 segregated bands. After chi-square test according to Mendelian ratio, a total of 470 molecular markers were suitable for genetic map construction. The phenomenon that markers deviated from Mendelian ratio could be caused by small population size, scoring errors, selection pressure, nonrandom segregation, and gametes competition [43].

A first preliminary genetic linkage map was developed for *S. paramamosain* (Figure 5). This map contained 65 linkage groups and 212 molecular markers (60 microsatellites and 152 AFLPs). In theory, the number of linkage group is equal to the number of haploid chromosome. In this study, the haploid chromosome number (*N* = 49) [44] was much lower than the total group number (*N* = 65), which indicated that this genetic map may still be preliminary with small number and low resolution of markers. The number of markers per genetic group ranged from 2 to 14, with an average of 3.3. All markers were evenly distributed in genetic groups and no clustering was found. This linkage map was 2746 cM in length with an average resolution of 18.7 cM. The expected genome was estimated to be approximately 5540 cM, which covered about 50% by our preliminary genetic map. In the next step, a high density genetic linkage map needs to be constructed in order to facilitate QTL mapping and marker-assisted selection for *S. paramamosain*. 
Figure 5. Genetic linkage map constructed for the mud crab (*Scylla paramamosain*). Genetic distances are listed on the left in Kosambi units (cM). Markers are shown on the right of each linkage group. Amplified fragment length polymorphism (AFLP) markers are named by the primer combination and fragment size. Microsatellite markers are in bold.

3. Conclusions

This study isolated and characterized 302 polymorphic microsatellite markers for the mud crab (*S. paramamosain*), and uncovered the high polymorphism and low genetic differentiation of wild populations distributed along South-eastern China coasts. Furthermore, a PCR-based parentage assignment method was well developed which could correctly assign 95% of the offspring to right parents. Moreover, three microsatellite loci were identified to link with growth performance of *S. paramamosain*. Finally, a first preliminary genetic linkage map was created for *S. paramamosain* using microsatellite and AFLP markers. These findings should provide novel insights into genome biology, wild resource background, and molecular marker-assisted selection in *S. paramamosain*.

Acknowledgements

This work was supported by the Top-Notch Young Talents Program of China, the National Natural Science Foundation of China (Grant No. 31001106), and the Fund of Key Laboratory
of Sustainable Development of Marine Fisheries, Ministry of Agriculture, China (Grant No. 2013-SDMFMA-KF-5).

Author details

Hongyu Ma1*, Chunyan Ma2, Lingbo Ma2, Xincang Li2 and Yuanyou Li1

*Address all correspondence to: mahongyuhome@163.com

1 Guangdong Provincial Key Laboratory of Marine Biology, Shantou University, Shantou, China

2 East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai, China

References


Ma HY, Ma CY, Li SJ, Jiang W, Li XC, Liu YX, Ma LB. Transcriptome analysis of the mud crab (*Scylla paramamosain*) by 454 deep sequencing: assembly, annotation, and marker discovery. Plos One. 2014; 9: e102668. DOI: 10.1371/journal.pone.0102668


