First genetic linkage map for the mud crab (*Scylla paramamosain*) constructed using microsatellite and AFLP markers

H.Y. Ma\(^1,2\), S.J. Li\(^1,2\), N.N. Feng\(^1,2\), C.Y. Ma\(^1,2\), W. Wang\(^1,2\), W. Chen\(^1,2\) and L.B. Ma\(^1,2\)

\(^1\)East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai, China
\(^2\)Key Laboratory of East China Sea and Oceanic Fishery Resources Exploitation, Ministry of Agriculture, Shanghai, China

Corresponding author: L.B. Ma
Email: malingbo@vip.sina.com

Received June 2, 2015
Accepted January 8, 2016
Published May 20, 2016
DOI http://dx.doi.org/10.4238/gmr.15026929

**ABSTRACT.** The mud crab (*Scylla paramamosain*) is of economic importance for the fisheries and aquaculture industry in China. In this study, we constructed the first genetic linkage map for this species using microsatellite and amplified fragment length polymorphism (AFLP) markers. The map consisted of 65 linkage groups, including 34 triplets and 9 doublets. A total of 212 molecular markers were mapped, including 60 microsatellites and 152 AFLP markers. The linkage groups ranged from 7 to 102.5 cM and covered 2746.4 cM in length. The mean length was 42.3 cM per linkage group, and the mean spacing was 18.68 cM. The genome size was estimated to be 5539.62 cM, with 50% coverage by the present map. Moreover, we reported 5 transcriptome-derived polymorphic microsatellite markers and characterized their polymorphism in a first-generation family. This study will facilitate
studies on high-density maps and molecular marker-assisted selection in *S. paramamosain* and related crustacean species.

**Key words:** Mud crab; Microsatellites; AFLP; Polymorphism; Genetic linkage map

**INTRODUCTION**

Mud crabs (genus *Scylla*) are widely distributed over a vast geographic area, ranging from southeast Africa to southeast Asia and Indo-Pacific regions (Fuseya and Watanabe, 1996). *Scylla paramamosain*, one of the above species, mainly inhabits the coasts of China, Vietnam, and Japan (Takano et al., 2005). In China, *S. paramamosain* is one of a number of traditionally important fishery resources and aquaculture species. In recent years, the total culture yields of *S. paramamosain* have remained stable at approximately 110,000 tons per year (Fishery Bureau of Ministry of Agriculture of China, 2012). Because of its high nutrition value, good flavor, and fast growth rate, market demands for *S. paramamosain* are quickly increasing. However, current culture capacity cannot meet these market demands.

In natural environments, sexual maturity and spawning of *S. paramamosain* occur continuously throughout the year, with seasonal peaks in certain areas (Le Vay, 2001). Individuals mate inshore and spawn eggs offshore, and then the offspring return to inshore. The genetic population diversity of wild *S. paramamosain* is high, and significant correlations have been found between genetic distance and geographic distance among different populations (Ma et al., 2012). Moreover, significant associations between each pair of 20 growth-related morphologic traits have been tested (Ma et al., 2013a). Under artificial conditions, hatchery stocks show significantly faster growth rates than wild populations (Utet et al., 2007), and genetic differences among artificial families are much higher than among wild populations (Cui et al., 2011a). In order to boost molecular marker-assisted selection (MAS) for *S. paramamosain*, a microsatellite-PCR based parentage assignment technique was created successfully, with a success rate of 95% (Ma et al., 2013b). Growth performance-related microsatellite markers were developed in *S. paramamosain*, and the genotypes with the most potential were also identified (Ma et al., 2014a). In addition, a large number of transcriptome resources, including functional genes, single nucleotide polymorphisms (SNPs), and microsatellites have been developed (Gao et al., 2014; Ma et al., 2014b). However, aside from the above-mentioned studies, little genetic reference is available for *S. paramamosain*.

Genetic linkage maps are considered to be powerful tools for MAS and comparative genomics research. Over the last decade, genetic linkage maps have been mainly constructed for a few economically important aquaculture crustaceans, such as *Penaeus monodon* (Staelens et al., 2008), *Portunus trituberculatus* (Liu et al., 2012), and *Litopenaeus vannamei* (Andriantahina et al., 2013). To date, no genetic linkage map is available for *S. paramamosain*, which has severely limited the development of MAS.

The purpose of this study is to construct a first-generation genetic linkage map for *S. paramamosain* using microsatellite and amplified fragment length polymorphism (AFLP) markers, which should be useful in the development of high-density genetic maps, comparative genomics, and MAS for *S. paramamosain* and related species.
MATERIAL AND METHODS

Mapping family construction and DNA extraction

A first-generation (G1) family of *S. paramamosain* was created at Qionghai Research Center in Hainan Province in May 2012. Both parents were wild individuals, and all progeny were cultured in a pond. A total of 95 one-month-old offspring were randomly collected for map construction. All individuals were stored in 95% ethanol. Genomic DNA was extracted from the whole body using the phenol/chloroform method, as described previously (Ma et al., 2009).

Microsatellite markers and genotyping

In order to acquire as many markers as possible, we first developed polymorphic microsatellite markers from the transcriptome, using a previously described method (Ma et al., 2014a). Then, we collected all published microsatellite markers for *S. paramamosain* (Xu et al., 2009; Ma et al., 2011a,b,c; Cui et al., 2011b; Yao et al., 2012; Song et al., 2013; Ma et al., 2014a,b). Finally, a total of 337 microsatellite markers (S1 Table) were genotyped in the mapping family.

PCRs were performed on a Peltier Thermal Cycler (PTC-200) in a 12.5-µL volume, which included 0.2 µM each primer, 0.2 mM each dNTP, 1XPCR buffer, 1.5 mM MgCl2, 0.3 U Taq polymerase, and approximately 50 ng DNA. PCR conditions were as follows: one cycle of denaturation at 94°C for 4 min; 30 cycles of 30s at 94°C, 50s at a primer-specific annealing temperature (S1 Table), and 50s at 72°C, with a final extension step at 72°C for 7 min. PCR products were separated on 6% denaturing polyacrylamide gels and visualized using silver-staining. Allele sizes were estimated according to the pBR322/MspI marker.

AFLP genotyping

In brief, 100 ng of genomic DNA was digested with 1 µL mixture EcoRI and MseI (1.25 U/µL; Li-Cor, USA) for approximately 6 h at 37°C, then heated at 70°C for 15 min to inactivate the enzymes. Ligation of specific adapters to restriction fragments was performed by adding 12 µL adapter mixture and incubating for 20 h at 20°C, and then the ligation reaction products were diluted 10 times. Pre-amplification PCR was performed in a 25.25 µL volume with 20 µL pre-amplification primer mixture, 2.5 µL10X PCR reaction buffer, 0.05 U Taq DNA polymerase, and 2.5 µL diluted ligation mixture. The PCR conditions were as follows: pre-denaturation at 94°C for 5 min, 20 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 1 min, and extension at 72°C for 1 min. The pre-amplified PCR products were diluted 45 times.

Sixty-four selective primer combinations (Table 1) were used for selective amplification. PCRs were performed in a 10.6 µL volume that contained 6.0 µL buffer mixture, 2.0 µL MseI selective amplification primer, 0.5 µL EcoRI selective amplification primer, 0.05 U Taq DNA polymerase, and 2.0 µL diluted pre-amplified PCR product. The PCR was performed using a “touchdown” program: one cycle of denaturation at 94°C for 30 s, annealing at 65°C for 30 s, and extension at 72°C for 1 min; 12 cycles of subsequently lowering the annealing temperature (65°C) by 0.7°C per cycle while maintaining denaturation at 94°C for 30 s and extension at 72°C for 1 min; and 23 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min. The PCR products were separated on 6% denaturing polyacrylamide gels on a 4300 DNA Analyzer (Li-Cor, USA). Allele sizes were calculated according to the 50 to 700 bp size standard.
Table 1. AFLP selective primer sequences used in the creation of a genetic linkage map for *Scylla paramamosain*.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Mse series</th>
<th>Primer name</th>
<th>EcoRI series</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>GATGAGTCCTGAGTAA−CAA</td>
<td>E1</td>
<td>GACTGCGTACCAATTC+AAAC</td>
</tr>
<tr>
<td>M2</td>
<td>GATGAGTCCTGAGTAA−CAC</td>
<td>E2</td>
<td>GACTGCGTACCAATTC+AAC</td>
</tr>
<tr>
<td>M3</td>
<td>GATGAGTCCTGAGTAA−CAG</td>
<td>E3</td>
<td>GACTGCGTACCAATTC+ACA</td>
</tr>
<tr>
<td>M4</td>
<td>GATGAGTCCTGAGTAA−CAT</td>
<td>E4</td>
<td>GACTGCGTACCAATTC+ACC</td>
</tr>
<tr>
<td>M5</td>
<td>GATGAGTCCTGAGTAA−CAG</td>
<td>E5</td>
<td>GACTGCGTACCAATTC+ACT</td>
</tr>
<tr>
<td>M6</td>
<td>GATGAGTCCTGAGTAA−CTA</td>
<td>E6</td>
<td>GACTGCGTACCAATTC+AGG</td>
</tr>
<tr>
<td>M7</td>
<td>GATGAGTCCTGAGTAA−CTG</td>
<td>E7</td>
<td>GACTGCGTACCAATTC+AGC</td>
</tr>
<tr>
<td>M8</td>
<td>GATGAGTCCTGAGTAA−CTT</td>
<td>E8</td>
<td>GACTGCGTACCAATTC+AGG</td>
</tr>
</tbody>
</table>

Linkage analysis and map construction

All microsatellite and AFLP genotypes were converted into the data format of the software JoinMap 3.0 (Van Ooijen, 2001), and the population type was set to cross-pollination (CP). The deviation of allelic segregation from expected Mendelian ratios (1:1 or 1:2:1) was tested using chi-square ($\chi^2$) values for all individual markers. Only markers with a P value >0.01 were used for further linkage analysis. Markers were assigned to different linkage groups with a critical logarithm of odds (LOD) score threshold ≥ 2.5. Genetic distance (cM) between markers was estimated using the Kosambi mapping function. Linkage maps were drawn using MapChart 2.1 software (Voorrips, 2002) and rearranged using Office Visio 2007 software.

The expected genome size ($G_e$) was estimated using the formula:

$$G_e = \frac{(G_{c1} + G_{c2})}{2}$$  \hspace{1cm} (Equation 1)

where $G_{c1}$ was calculated by adding 2s (where s is the mean maker space of the linkage map) to the observed length of each linkage group (Fishman et al., 2001), and $G_{c2}$ was estimated by multiplying the observed length of each linkage group by $(m + 1)/(m - 1)$ (where m is the number of markers in each linkage group). The expected genome size is the sum of the revised lengths of all linkage groups (Chakravarti et al., 1991). The observed map length ($\phi_{oa}$) was calculated as the total length of the map including triplets and doublets. The estimated coverage of the genome ($C_{oa}$) was calculated as:

$$\phi_{oa}/\phi_e$$  \hspace{1cm} (Equation 2)

RESULTS

Characterization of five novel polymorphic microsatellite loci

In this study, five novel microsatellite loci were isolated and confirmed to be polymorphic in a G1 family, and their sequences were submitted to the GenBank database (accession No. KJ000478-KJ000482; Table 2). A total of 10 alleles were detected, with each locus containing two alleles. The observed and expected heterozygosity per locus ranged from 0.35 to 0.58, and from 0.42 to 0.51, with a mean of 0.50 and 0.49, respectively. All
five loci showed medium information, with polymorphism information content (PIC) values ranging from 0.33 to 0.37. One locus (Scpa117) significantly deviated from Hardy-Weinberg equilibrium (HWE, P < 0.05), which may be the result of small sample size or the presence of null alleles, but no evidence for stuttering and allelic dropout was found in any locus and it cannot be attributed to technical or statistical artifacts. Furthermore, no significant genotypic linkage disequilibrium was detected between any pairs of the loci.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Repeat sequence</th>
<th>Primer sequence (5’–3’)</th>
<th>T_a (°C)</th>
<th>N_a / N_e</th>
<th>H_o</th>
<th>H_e</th>
<th>PIC</th>
<th>P</th>
<th>GenBank accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scpa117</td>
<td>(TAT)_14</td>
<td>CCTCTTTTTTCAATGACCC</td>
<td>35</td>
<td>2.0/1.7</td>
<td>0.78</td>
<td>0.42</td>
<td>0.33</td>
<td>0.000^</td>
<td>KJ000478</td>
</tr>
<tr>
<td>Scpa118</td>
<td>(TG)_9</td>
<td>CCTATGGGATGCTCTGTTGTTCAAGGATGTAATGAAC</td>
<td>52</td>
<td>2.0/2.0</td>
<td>0.55</td>
<td>0.50</td>
<td>0.37</td>
<td>0.287</td>
<td>KJ000479</td>
</tr>
<tr>
<td>Scpa119</td>
<td>(TTA)_13</td>
<td>AGTCGCCGTCGTTAGGTCATTTCACGCCCTTTCGTCATTGAC</td>
<td>50</td>
<td>2.0/2.0</td>
<td>0.35</td>
<td>0.50</td>
<td>0.37</td>
<td>0.000^</td>
<td>KJ000480</td>
</tr>
<tr>
<td>Scpa120</td>
<td>(GT)_15</td>
<td>GATTTGCTCAGGCTCTACGTTGAGGACAGGCACGAGGCTTG</td>
<td>50</td>
<td>2.0/2.0</td>
<td>0.47</td>
<td>0.50</td>
<td>0.37</td>
<td>0.591</td>
<td>KJ000481</td>
</tr>
<tr>
<td>Scpa121</td>
<td>(TG)_10</td>
<td>TTATATTTGTCCGCTGCTGTTGTTCAAGGATGTAATGAAC</td>
<td>50</td>
<td>2.0/1.9</td>
<td>0.54</td>
<td>0.51</td>
<td>0.36</td>
<td>0.410</td>
<td>KJ000482</td>
</tr>
<tr>
<td>Mean</td>
<td>-</td>
<td>-</td>
<td>2.0/1.9</td>
<td>0.50</td>
<td>0.49</td>
<td>0.36</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

T_a = annealing temperature; N_a = observed number of alleles; N_e = effective number of alleles; H_o = observed heterozygosity; H_e = expected heterozygosity; PIC = polymorphism information content. *Significant deviation from Hardy-Weinberg equilibrium.

Microsatellite and AFLP marker evaluation

A total of 337 polymorphic microsatellite markers were employed to construct a genetic linkage map for S. paramamosain, of which 118 showed clear segregation in the G_1 mapping family. Twenty-one markers segregated in the male parent, 34 segregated in the female parent, and 63 segregated in both parents. Sixty-four AFLP selective primer combinations produced a total of 574 segregating bands, of which 203 were passed on from the male parent to the progeny, 219 were passed on from the female parent to the progeny, and 152 were shared by both parents. All 692 segregated markers (118 microsatellites and 574 AFLPs) obtained in this study were tested using a chi-square (χ^2) test for a 1:1 or 1:2:1 segregation pattern based on expected Mendelian ratios at a P value of 0.01. As a result, 222 markers showed significant deviation from Mendelian ratios.

Genetic linkage map

All 470 segregated markers which conformed to Mendelian ratios were used to create a genetic linkage map for S. paramamosain. The map consisted of 65 linkage groups, including 9 triplets and 34 doublets (Figure 1). A total of 212 molecular markers (60 microsatellites and 152 AFLPs) were mapped onto this map (details of these markers are shown in S2 Table). Thirty-one of the linkage groups contained more than two markers. The mean number of markers per linkage group was 3.3, ranging from 2 to 14. The distribution of markers on this map was reasonably even and no clusters were observed. The genetic linkage map spanned 2746.4 cM, with a mean resolution of 18.68 cM. The length of linkage groups varied from 7 to 102.5 cM, with an average of 42.3 cM. The highest marker resolution of this map was 54.9 cM, and the genome size of S. paramamosain was estimated to be 5539.62 cM, with 50% coverage by the present map.
Figure 1. Genetic linkage map 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. Continued on next page
First genetic linkage map for the mud crab

Figure 1. Continued.
DISCUSSION

AFLPs are considered to be an ideal marker system for rapidly constructing a genetic linkage map for non-model organisms because they can produce a large number of polymorphic markers without any prior genomic sequence information. This technique has been successfully used to develop genetic maps for *Verasper moseri* (Ma et al., 2011d), *Artiemia franciscana* (De Vos et al., 2013), and *Torreya grandis* (Zeng et al., 2014). Compared with AFLPs, the use of microsatellite markers requires prior knowledge of DNA sequences and PCR primers. Nevertheless, this type of marker system has a number of advantages, including a co-dominant inheritance pattern and good transferability across different families and species. Thus, microsatellites are more commonly used in the construction of genetic maps, and they have been used in many animals, such as *Portunus tribuberculatus* (Liu et al., 2012), *Litopenaeus vannamei* (Andriantahina et al., 2013), and *Scophthalmus rhombus* (Hermida et al., 2014). In the present study, we constructed a genetic linkage map for *S. paramamosain* using microsatellite and AFLP markers based on a G1 family. To the best of our knowledge, this is the first report of a genetic linkage map for the genus *Scylla*.

To construct a genetic linkage map, we first isolated and characterized a set of type I polymorphic microsatellite markers from the transcriptome of *S. paramamosain*. Type I markers usually associate with functional genes, so they are thought to be more conserved than type II markers. In *V. moseri*, the variation of type I microsatellite loci (mean $H_O = 0.42$) was less than that of type II loci (mean $H_O = 0.61$; Ma and Chen, 2009). In the present study, a total of five type I microsatellite loci were found to be polymorphic using a G1 family. The genetic diversity (mean $H_O = 0.50$) was much lower than that of type II loci (mean $H_O = 0.76$; Ma et al., 2010), and even lower than that of type I loci developed in a previous study (mean $H_O = 0.68$; Ma et al., 2014a). Aside from the major reason of microsatellite loci themselves, this difference may be due to the types and sizes of populations (wild, cultured, or a single family) used. However, in some particular cases, the polymorphism level of type I microsatellite markers has been found to be equivalent or even higher than that of type II markers (Yu and Li, 2008).

Sixty-four AFLP selective primer combinations generated 574 segregating bands, with a mean of 9.0 bands per primer combination. The ratio of segregating bands was similar to that in *T. grandis* (11 per primer combination; Zeng et al., 2014), higher than that in *A. franciscana* (8.2 per primer combination; De Vos et al., 2013), but lower than that in *P. monodon* (10.6 per primer combination; Staelens et al., 2008). These differences may relate to the genetic diversity level of different species. Furthermore, a number of segregating bands with unclear genotypes were neglected in order to maintain the precise results, which reduced the total number of the segregating bands. Out of 337 microsatellite markers, 118 were polymorphic in the mapping family, with a ratio (35.0%) similar to that observed in *V. moseri* and *V. variegatus* (33.9%; Ma et al., 2011d).

Of the 692 segregated markers scored, 222 (32.1%) showed significant distortion at the $P=0.01$ level. Segregation distortion can be caused by several factors, such as small population size, scoring errors during genotyping, genetic selection pressure, non-random segregation, and competition of gametes (Liebhard et al., 2003). In addition, the deletion of genes or chromosomal regions can affect segregation of alleles (Lytte, 1991). The percentage of distorted markers (32.1%) in this study was similar to that observed previously in *Chlamys farreri* (30%; Zhan et al., 2009), lower than that in *Ostrea edulis* (85%; Lallias et al., 2007),
and higher than that in *P. trituberculatus* (10%; Liu et al., 2012). We eliminated these distorted markers during construction of the genetic map, although they are thought to be helpful for identification of deleterious genes (Lallias et al., 2007).

If triplets and doublets are included, the number of linkage groups in this study is 65, which is higher than the number of haploid chromosomes (*N* = 49) of *S. paramamosain* (Chen et al., 2004). Theoretically, the number of linkage groups is supposed to be equal to the haploid chromosome number. However, nonequivalence is often found in animals, including *V. moseri* (Ma et al., 2011d) and *P. trituberculatus* (Liu et al., 2012). The low marker density of linkage maps may be a major reason resulting in nonequivalence because it can result in gaps on the linkage maps. Thus, in order to obtain an ideal genetic map, a large number of additional molecular markers should be mapped on to the linkage map in the future. Furthermore, although AFLPs tend to cluster together on linkage groups, this phenomenon was not observed in this study. Suppressed recombination at centromeric and telomeric regions of chromosomes is thought to affect marker distribution on linkage groups (Liu et al., 2003).

**CONCLUSION**

In this study, we constructed the first genetic linkage map for the mud crab (*Scylla paramamosain*) using microsatellite and AFLP markers, which should be useful for the development of high-density maps and molecular marker-assisted selection in *S. paramamosain* and related crustacean species.

**Conflicts of interest**

The authors declare no conflict of interest.

**ACKNOWLEDGMENTS**

Research supported by the Natural Science Foundation of Shanghai (Grant #14ZR1449700), the National Program for Support to Top-notch Young Professionals, the Program of Science and Technology Commission of Shanghai (#15391900900), and Open Research Fund Program of Fujian Provincial Key Laboratory of Marine Fishery Resources and Eco-environment (Grant #FMFRE2014003).

**REFERENCES**


First genetic linkage map for the mud crab


Supplementary material

S1 Table. 337 polymorphic microsatellite markers used for constructing a genetic map of Scylla paramamosain in this study.

S2 Table. Characterization of all 212 mapped markers in the present study. F, segregated in female parent; M, segregated in male parent; B, segregated in both parents.