

中国对虾微卫星 DNA 引物的设计及筛选

张天时^{1,2}, 刘 萍¹, 孔 杰¹, 王清印¹

(1. 中国水产科学研究院 黄海水产研究所, 山东 青岛 266071; 2. 中国海洋大学 海洋生命学院, 山东 青岛 266003)

摘要:根据建立的中国对虾(*Fenneropenaeus chinensis*)部分基因组文库,对筛选的含微卫星 DNA 序列的克隆设计了 28 对引物,筛选出 5 对微卫星多态性引物,并用这 5 对引物对中国对虾的养殖群体 20 个个体进行了遗传多样性分析。在这 5 个微卫星位点中,虽然 RS0871 位点是多态位点,但其等位基因数、杂合度和多态信息含量都较低。其余 4 个微卫星位点可产生 6~8 个等位基因,等位基因的大小分布在 142~364 bp,基本上符合引物设计时理论产物长度。这些微卫星位点的期望杂合度的范围为 0.757 7~0.806 4,表明它们都有较高的杂合度。仅有位点 RS0956 的观察杂合度与期望值有较大的出入,其他微卫星位点的这两值基本相符。有 4 个微卫星位点的 PIC 值较高,在 0.718 0~0.770 9。因此,除 RS0871 位点以外的 4 个微卫星位点均可用于中国对虾种群遗传结构分析,这将为对虾品种选育、种系评估提供更多的微卫星 DNA 信息。

关键词:中国对虾;微卫星;基因组文库;遗传多样性

中图分类号:Q959.223 **文献标识码:**A **文章编号:**1005-8737-(2004)06-0567-05

自 20 世纪 90 年代初期以来,微卫星标记已经广泛应用于人类和一些重要种类动植物连锁群的构建和物理图谱的绘制^[1]。在一些海洋生物,如大西洋鲑(*Salmo salar*)^[2]、雪蟹(*Chionoecetes opilio*)^[3]和真鲷(*Pagrosomus major*)^[4]等中也已经利用微卫星标记进行了许多工作。但中国对虾(*Fenneropenaeus chinensis*)作为一个重要的经济海产种类,微卫星标记的研究工作还鲜见报道。近年来,许多海洋生物工作者开始进行对虾微卫星标记的开发和应用等研究,取得了一定进展,但是由于微卫星 DNA 序列在甲壳动物不同种之间的通用性较差,因此就每种对虾来说,目前可用微卫星标记还不是非常丰富。就此本实验室开展了中国对虾部分基因组文库的构建工作,开发了中国对虾微卫星标记引物。在筛选了具有多态性的微卫星核心序列引物的基础上,设计并筛选出 5 对微卫星多态性引物,并对这 5 对引物的多态性信息含量进行了评估,以期为中国对虾的品种选育、种系评估以及连锁图谱的构建提供信息。

1 实验方法

1.1 微卫星序列的获得

在已建立中国对虾部分基因组文库^[5]中进行微卫星序

列的筛选。

1.2 模板 DNA 的提取

检测微卫星引物多态性使用的模板 DNA 提取自中国对虾尾部肌肉,采用常规方法酚/氯仿抽提,乙醇沉淀,具体操作参照石拓等^[6]的方法。

1.3 微卫星引物的设计

根据对文库中克隆序列的测定,选取微卫星 DNA 序列的克隆,使用引物设计软件 Oligo (Version 6.31) 和 Primer Premier (Version 5.00) 在其核心序列的两端设计了相应的 PCR 引物。设计好的引物由上海生物工程公司合成。

1.4 PCR 扩增与电泳检测

将来自 8 条中国对虾的基因组 DNA 样品混合作为模板 DNA,质量浓度为 50 ng/μL,利用合成的微卫星引物进行初步 PCR 扩增筛选。

PCR 反应体系总体积为 25 μL,包括 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 2.0 mmol/L MgCl₂, 0.2 mmol/L dNTPs, 0.8 μmol/L 引物, 1.0 U Taq 酶, 100 ng 模板 DNA。反应在 PT200 扩增仪上进行。PCR 反应程序为: 94℃ 预变性 5 min; 每一个循环包括 94℃ 40 s、退火温度(表 2) 1 min、72℃ 1 min, 进行 25 个循环; 72℃ 延伸 5 min, 4℃ 保温。

收稿日期:2003-12-16; 修订日期:2004-02-17。

基金项目:国家自然科学基金项目(30271038); 国家重点基础研究发展规划“973”课题(1999012009); 国家“863”高技术课题(2003AA603021)。

作者简介:张天时(1976-),男,硕士,从事水产养殖基础生物学和遗传育种研究。E-mail: zhtianshen@yahoo.com.cn

通讯作者:王清印。E-mail: qywang@public.sd.cn

表1 从中国对虾基因组文库中筛选所得的微卫星序列

Table 1 Microsatellite sequences isolated from genomic libraries of *F. chinensis*

克隆编号 Clone no.	5'端序列 5' Flanking sequence	微卫星核心区序列 Core sequences	3'端序列 3' Flanking sequence	重复类型 Motif	GenBank 登录号 GenBank accession no.
RS0653	TTAATAAAAG	(AC) ₁₈	ATATATACACA	Perfect	AY132779
RS0676	TACCAATCTT	(AC) ₁₅ GC(AC) ₁₀	ATATATATATA	Imperfect	AY132780
RS0676	TACCAATCTT	(AC) ₁₆ GC(AC) ₃₁ A(AC) ₆	TATATATATA	Imperfect	AF526880
RS0683	CTCTAATACC	(GCTACCC) ₁₄	CTCCCAACAAA	Perfect	AY132823
RS0704	GGGCGATTGG	(AC) ₃₀ (GCACAC) ₂ (AC) ₃ AT(AC) ₂ C(AC) ₃ GC (AC) ₆ ATAT(AC) ₃ (ATAC) ₃ (AC) ₂ GCACGC (AC) ₈ AT(AC) ₁₄ (AT) ₃ A(AC) ₃ (AT) ₈ GT(AT) ₄ TT(AT) ₄ ... (AT) ₃ (AC) ₄ ... (AT) ₃ GT	TATGTATATAC	Compound imperfect	AY132781
RS0725	CACATGCTAA	(AC) ₈ ... (AC) ₃ ... (AC) ₄ ... (AC) ₇ A(AC) ₃ ... (AC) ₃ A(AC) ₃	AAAGCTGCACT	Imperfect	AY132783
RS0744	TCCCTGCCCTC	(CA) ₃₄	TCCGGATTAG	Perfect	AY132785
RS0754	ATACCTTTAT	(AC) ₄ GC(AC) ₃	TAGCAACAGT	Imperfect	AY132786
RS0762	GCAGGCTATA	(AAT) ₃ ... (AAT) ₈	TTAACATCAAC	Imperfect	AY132787
RS0774	TGACCAAAAA	(TG) ₃ TA(TG) ₄ ... (TCTG) ₃ (TG) ₄ (TA) ₃	GGAGGTCTTAT	Compound perfect	AY132788
RS0776	TTTGTCTGTG	(TA) ₃ CATG(TA) ₂ ... (TA) ₃ TG(TA) ₃ ... (TG) ₄ (TA) ₂ TG(CA) ₂ TG(TA) ₃ ... (TA) ₃ ... TG(TA) ₄	TGTCTGTTTAT	Compound imperfect	AY132789
RS0779	AACAACAGGA	(TAA) ₂ TGA(TAA) ₃ CAA(TAA) ₄ ... (AAT) ₃ A (AAT) ₈ ... (AAT) ₃ GAT(AAT) ₇ ... (AAT) ₃ GTAT (AAT)GAT(AAT)AAC(AAT)TATACATTT(AA T)AA(AAT) ₃ AAAAC(AAT) ₂ ATA	ACAATAATATA	Compound imperfect	AY132790
RS0865	GCATACATGC	(GT) ₃ AGTATGC(GT) ₃ GC(GT) ₇ T(GT) ₂	TTTTGATGAAC	Imperfect	AY132792
RS0878	ACCTCGGAGT	(GCTTGCCT) ₄	TCAGTAATCAT	Perfect	AY132824
RS0871	CACACAAGTG	(CA) ₃ ... CACG(CA) ₄ TT(CA) ₂	CTCTCCCTCTC	Imperfect	AY132793
RS0878	ACCTCGGAGT	(GCTTGCCT) ₃ GCT	TCAGTAATCAT	Perfect	AY132824
RS0903	TTATCAAGCT	(TCTG) ₃ TC(TCTG)(TC) ₃ TTTGCAG(TC) ₉ TTAC(TC) ₃	GGATTCTCGC	Compound imperfect	AY132794
RS0908	GACACATACA	(TG) ₃ AGTT(TG) ₃ (CATAT)(TC) ₂ TTTGTG (CATAT)(TC) ₁₁ T(TC)	CTGTCTTTCTT	Compound imperfect	AY132795
RS0916	GTGGAAAGGC	(TAA)TGA(TAA) ₂ ... (TAA) ₆ ... (AAC) ₂ AGC(AAC) ₃	AGGAGTAGCAT	Compound imperfect	AY132796
RS0930	TTGTAAAGTA	(GAT) ₄ CAT(GAT) ₂ AGT(GAT) ₂ A (GAT)AT(GAT) ₄ G	TATTAATAGAA	Imperfect	AY132797
RS0942	TACGAAAAAA	(CT) ₂ C(CT)T(CT) ₄ TC(CT)CCATT(CT)	TAAATATTGCT	Imperfect	AY132798
RS0951	ACAAGTTGAT	(TAA) ₄ TCATGA(TAA) ₂₆ TGGTAG(TAA)TGA (TAA) ₂ TACAAA(TAA)AC(TAA) ₂ TAT(TAA) TGC(TAATAC)(TAA)TAT(TAATAC) ₃ (TAA) TAGTGA(TAA) ₄ AAT(TAA) ₃	TTTAGTTTAT	Compound imperfect	AY132799
RS0956	ATACCGGTTA	(AG) ₃ AA(AG) ₂₆ A(AG) ₄ A(AG) ₃₂ A(AG)AA (AG) ₃ TGGTGG(AG) ₂	AACATAAGAA	Compound imperfect	AY132800
RS0961	AAAAAGAAGA	(AT) ₃ (AC) ₃ A(TA)(AT)(AC)(AT)(TA)(TG) ₂ (TA) ₄ A(TA) ₄ (CA) ₂ (TG) ₄ (TA) ₄ GA(AT) ₃ (AC) ₃ (AT)A(AT)AA(AT) ₂₁	ATAGCTATTTA	Compound imperfect	AY132801
RS0965	CAATCTATAT	(CAAA)(CA) ₄ TTAA(CA)(CAAA) ₂ (CA) ₂ TAAA (CA)(CAAA)CAGA(CA) ₃ (CAAA)... (CA) ₄ (CA A) ₂ (CA) ₂ TAAA(CA)(CAAA)(CAAA)(CA) ₄ GT TA(CA) ₂ (CAAA) ₂ (CA) ₃₁ (GA) ₈ (CA) ₇ ... CAAA TAAA(CA) ₁₆ (GA)GT(CA) ₃ (CAAA) ₂ (CA) ₈	CGTATCTTTAA	Compound imperfect	AY132802
RS0971	GGGGGAAAAA	(AAT) ₂ (GAT) ₂ AGT(GAT)(AAT)(GAT)A (GAT)... (CAT) ₄ AAT(CAT) ₃	TTCAAGTTTAA	Compound imperfect	AY132803
RS0989	AAAGGAAAAAC	(TA) ₃ (TG) ₂ (TA) ₃ CA(TA) ₁₇ (TACATATG) ₂ (TA) ₄ (TG) ₂ (TA) ₁₈ T(TA) ₂	AGCAATACATT	Compound imperfect	AY132804
RS1080	GTTTGCCCTT	(CACACT) ₂ (CACTCT) ₃ (CACTCT) ₃	CTTTTTTTTCT	Compound perfect	AY132810

表 2 中国对虾 5 对微卫星引物序列、退火温度及其评价
Table 2 Primer sequences, optimal annealing temperature and evaluation of *F. chinensis*

克隆编号 Clone no.	引物序列(5'→3') primer sequences(5'→3')	退火温度/℃ Annealing temperature	扩增片断 长度范围/bp Range	等位基因数 Alleles number	PIC	H_e	H_s
RS0653	F: TTC AGT TGT TTC AGA ATC GC R: ACA CGG AGT AAT GGA GAC C	61	242 – 276	6.0	0.718 0	0.650 0	0.757 7
RS0676	F: ACG ATG CTT ATT AGC TGC G R: TGT GGA GCT TGA TGG TTG C	63	289 – 346	6.0	0.748 0	0.650 0	0.785 9
RS0744	F: CAT CAC CCG TAT GTG CGT CCT C R: CCG GGC CGT GCA TTG TTA	63	142 – 178	8.0	0.753 1	0.842 1	0.788 1
RS0871	F: AAT GTA GAT TGT TTC CTT R: AGT AGA AAA GTC TGG ATG T	53	264 – 269	2.0	0.133 9	0.150 0	0.142 3
RS0956	F: ACG CAT TAT GTC TCA TTC GC R: GAT TCC GCA TCA AAT ATG TCC G	66	332 – 364	6.0	0.770 9	0.500 0	0.806 4

扩增完毕后,扩增产物通过 2% 琼脂糖凝胶 EB 染色电泳或 8% (W/V) 聚丙烯酰胺凝胶电泳和硝酸银染色检验其多态性。待凝胶干燥后使用扫描仪记录电泳图谱。

1.5 数据分析

利用所获得微卫星引物对中国对虾的一个养殖群体(20 个个体)进行遗传分析。对每一个微卫星位点的等位基因的数量进行统计,计算每个等位基因的频率,并计算下列统计量:

参照 Botstein 等^[7]的方法计算多态性信息含量(Polymorphism Information Content, PIC)

$$PIC = 1 - \sum_{i=1}^m P_i^2 - \sum_{i=1}^{m-1} \sum_{j=i+1}^m 2 P_i^2 P_j^2$$

式中 P_i 、 P_j 分别为群体中第 i 和第 j 个等位基因频率, m 为等位基因数。

多态位点杂合度(观测值):
 H_o 为杂合子观察数与观察个体总数之比。
多态位点多态位点杂合度(期望值):

$$H_e = 1 - \sum P_i^2$$

P_i 为该位点上第 i 个等位基因的频率。

2 结果

2.1 引物设计及初步筛选结果

根据文库中克隆序列的测定,选取了 28 个微卫星 DNA 序列设计了引物,微卫星核心序列见表 1。

经过初步筛选,在所设计的 28 对引物中共有 14 对引物有扩增产物,下面图 1 为部分引物以混合 DNA 为模板进行 PCR 扩增的电泳结果。

2.2 多态性微卫星引物的筛选结果

对 14 对能够扩增出产物的待检引物中,经过 PCR 扩增、聚丙烯酰胺电泳和硝酸银染色,最终筛选出有多态性 PCR 产物的微卫星引物有 8 对,其中适合微卫星多态性分析的有 5 对。以下图 2 为部分多态性微卫星标记在养殖群体中的聚丙烯酰胺凝胶电泳图。

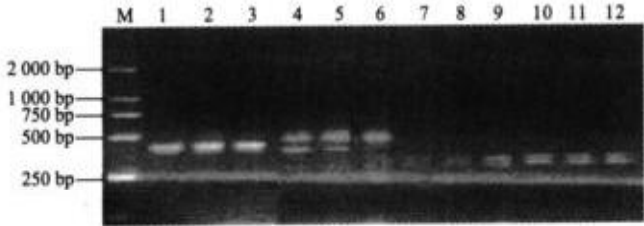


图 1 混合 DNA 模板进行 PCR 扩增的琼脂糖凝胶电泳图谱

1、2、3 为引物 RS0754 在 63 °C、61 °C、58 °C(单态);4、5、6 为引物 RS0989 在 67 °C、65 °C、63 °C(单态);7、8、9 为引物 RS0653 在 65 °C、63 °C、61 °C;10、11、12 为引物 RS0676 在 63 °C、61 °C、58 °C 3 个退火温度梯度的扩增产物。

Fig.1 Electrophoretic patterns of agarose gel of mixed DNA template

1,2,3 refer to Primer RS0754, whose annealing temperature gradient is 63 °C, 61 °C, 58 °C, respectively; 4, 5, 6 refer to Primer RS0989, 67 °C, 65 °C and 63 °C; 7, 8, 9 refer to Primer RS0754, 65 °C, 63 °C and 61 °C; 10, 11, 12 refer to Primer RS0754, 63 °C, 61 °C and 58 °C.

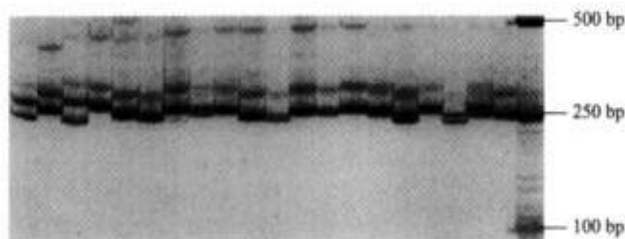


图2 引物 RS0653 的聚丙烯酰胺凝胶电泳图

Fig.2 PAGE image of Primer RS0653

2.3 数据统计与分析

利用所获得微卫星引物对中国对虾的一个养殖群体(20个个体)进行遗传分析。对每一个微卫星位点的等位基因数目(其中纯合个体的等位基因以出现两次计算)进行统计。根据各个微卫星位点各种基因型频率计算其观察杂合度,根据各个微卫星位点的等位基因频率,分别计算它们的期望杂合度和多态信息含量(PIC)。计算结果列于表2。

从表2中可以看出,在中国对虾20个样本的5个微卫星位点中,有4个微卫星位点等位基因数目从6到8不等,等位基因的大小在142~364 bp,基本上符合引物设计时理论产物长度。期望杂合度的范围从0.7577到0.8064,表明它们都有较高的杂合度。PIC值从0.7180到0.7709,指示这些微卫星位点在中国对虾中均具有较高的信息含量。位点RS0956的观察杂合度与期望值有较大的出入,其他微卫星位点的这两值均基本相符,可能是由于该位点的等位基因在所选取的20个样本中较为集中所致,如果扩大样本数可能会减小这种区别。只有RS0871位点等位基因数、杂合度和多态信息含量都较低。

3 讨论

3.1 中国对虾微卫星标记引物设计中微卫星序列的选取

在微卫星标记的引物设计中微卫星序列的选取是非常重要的和关键的。微卫星的核心序列有各种各样的类型,关于何种类型核心序列的微卫星筛选出微卫星标记的可能性较大,一直以来众说纷纭。本实验对各种类型核心序列进行了部分研究。结果表明,在扩增出PCR产物中的微卫星引物中,多态性的比例相对来说还是比较高的。在本实验中,所筛选出的微卫星标记的核心序列是以2个碱基、3个碱基或4个碱基为基本重复单位的序列。这与Pongsomboon等^[8]和Xu等^[11]在*P. monodon*中以三碱基和四碱基为基本重复序列的微卫星标记筛选效果较好的结果相符。

在微卫星核心序列的重复次数方面,Weber^[10]认为,只有在双碱基重复序列次数大于12次时,微卫星标记才有可能表现出较高PIC值,才可以进行相应的多态性分析,如位点RS0653和位点RS0676,然而在位点RS0871核心序列为以(CA)_n为基本重复单位的序列,最多重复次数只有8次,

所以等位基因数目非常少,PIC值也偏低(表2)。Xu等^[11]曾在*P. monodon*进行微卫星引物设计,最后的结果也表明,凡是核心序列重复次数较少的微卫星标记,其结果或单态的或等位基因数目非常少,PIC值也偏低。因此,在选用的微卫星序列的核心序列的重复次数上应在一个较高的水平上,避免出现由于微卫星核心序列过短,造成微卫星标记筛选中多态性引物比较低而造成浪费。

3.2 微卫星标记遗传多态性低的可能原因

在本研究中,5个中国对虾微卫星标记在20个样本中分别观测到2~8个等位基因,而Pongsomboon等^[8]、Xu等^[11]在*P. monodon*中筛选得到的多态性最高的微卫星标记其等位基因最多分别达到29和25个,大大高出本研究中所得最大等位基因数,究其原因,除了有可能是所选取中国对虾样本的亲缘关系较近,遗传多样性低以及所选取的样本较少等方面的原因外,聚丙烯酰胺电泳以及在微卫星位点的等位基因统计中出现的偏差对分析微卫星电泳结果可能有局限性。由于微卫星的较强的多态性,常常使得两条电泳带(等位基因)之间仅仅只有几个碱基的差别,反映在电泳胶上可能非常近,在区分两条带时很可能造成误判,将两条非常近的带以为1条。因此,建议使用测序仪的毛细管电泳结合荧光标记再利用其中的Genescan功能进行分析。

参考文献:

- [1] 刘志毅,相建海.微卫星DNA分子标记在海洋动物遗传分析中的应用[J].海洋科学,2001,25(6):1-13.
- [2] Ashie T N, Daniel G B, Edward P C, et al. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers[J]. Aquaculture, 2000, 182: 73-83.
- [3] Urbani N, Sevigny J M, Sainte-Marie B, et al. Identification of microsatellite markers in the snow crab *Chionoecetes opilio*[J]. Molecular Ecology, 1998, 7: 357-358.
- [4] Ricardo P E, Motohiro T, Nobuhiko Taniguchi, et al. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers[J]. Aquaculture, 1999, 173: 411-421.
- [5] 刘 萍,孟宪红,孔 杰,等.中国对虾部分基因组文库构建

- 和微卫星 DNA 的筛选 [J]. 高技术通讯, 2004, 14(2): 89 - 90.
- [6] 石拓, 孔杰, 刘萍, 等. 用 RAPD 技术对中国对虾遗传多样性分析 - 朝鲜半岛西海岸群体的 DNA 多态性 [J]. 海洋与湖泊, 1999, 30(6): 609 - 616.
- [7] Botstein D, White R L, Skolnick M, et al. Construction of genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Animal Genetics* [J], 1980, 32: 314 - 331.
- [8] Pongsomboon, Whan V, Moor S S, et al. Characterization of tri- and tetranucleotide microsatellite in the black tiger prawn, *Penaeus monodon* [J]. *Science Asia*, 2000, 26: 1 - 6.
- [9] Xu Z K, Primavera J H, Pena L D, et al. Genetic diversity of wild and cultured black tiger shrimp (*Penaeus monodon*) in the Philippines using microsatellites [J]. *Aquaculture*, 2001, 199: 13 - 40.
- [10] Weber J L. Informativeness of human (dC-dA)_n(dG-dT)_n polymorphism [J]. *Genomics*, 1990, 7: 524 - 530.
- [11] Xu Z, Dhar A K, Wyrzykowski J, et al. Identification of abundant and informative microsatellites from shrimp (*Penaeus monodon*) genome [J]. *Animal Genetics*, 1999, 30: 150 - 156.

Design and screening of microsatellite primers for *Fenneropenaeus chinensis*

ZHANG Tian-shi^{1,2}, LIU Ping¹, KONG Jie¹, WANG Qing-yin¹

(1. Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China; 2. College of Marine Life Science, Ocean University of China, Qingdao 266003, China)

Abstract: Total DNA was extracted from tail muscles of *Fenneropenaeus chinensis*. DNA quality controls were performed using agarose gel electrophoresis and only fragments at 500 - 1 000 bp were used for constructing a small-size fractionated genomic libraries. A genomic DNA library using the bacterial artificial chromosome (BAC) have been constructed in which the Genomic DNA digested by *Sau*3A I was ligated to the BAC vector and the ligation mixture was used to transform the bacterial strain DH5 α . By using the on-line bioinformatic software Repeatmasker, small-size fractionated genomic libraries of *F. chinensis* were screened. Twenty-eight microsatellite sequences were obtained, according to which 28 pairs of primers were designed using computer program Primer Premier (Version 5.00). The optimal results were obtained by optimizing quantity of reagents and reaction conditions. There were 14 primers whose PCR products can be detected by electrophoresis, among which five microsatellite loci were polymorphic in mixed templates (eight Chinese shrimp samples). In these five microsatellite loci, genetic diversity was analysed by electrophoresis analysis in a cultured population consisting of 20 Chinese shrimp *F. chinensis* samples. The allele numbers, the expected heterozygosity and the PIC values of them in the locus RS0871 were low. Except for the locus RS0871 the alleles numbers distributed from 6 to 8, and the range of these alleles were between 142 bp and 364 bp, similar to their predicted size. The expected heterozygosity and the PIC values of them distributed from 0.757 7 to 0.806 4 and 0.718 0 to 0.770 9, respectively. These values showed that 4 microsatellite loci except for the locus RS0871 could be applied for genetic analysis. This result indicated that polymorphic change of microsatellite primers was related to the core repeats motifs of the microsatellites. As a result, there will be less costing and time consuming in primer designs according to the core repeats motifs of microsatellite. Moreover, the microsatellite markers will be useful for further studing on accessions identification and breeding of *F. chinensis*, which would provide some valuable tools for marker-assisted selection breeding and gene mapping in *F. chinensis*.

Key words: *Fenneropenaeus chinensis*; microsatellite; genomic libraries; genetic diversity

Corresponding author: WANG Qing-yin. E-mail: qywang@public.sd.cn