

## 栉孔扇贝 BES-SSR 的开发及遗传多样性分析

张秀英<sup>1,2</sup>, 张晓军<sup>1</sup>, 赵 翠<sup>1,2</sup>, 常亚青<sup>3</sup>, 郇 聘<sup>1</sup>, 李富花<sup>1</sup>, 相建海<sup>1\*</sup>

(1. 中国科学院海洋研究所实验海洋生物学重点实验室, 山东 青岛 266071;

2. 中国科学院研究生院, 北京 100039;

3. 大连海洋大学水产与生命学院, 辽宁 大连 116023)

**摘要:** 通过分析栉孔扇贝 BAC 末端序列, 发现大量微卫星 DNA; 随机选择 14 个多态性 BES-SSR 标记, 在我国栉孔扇贝大连群体(DL)和青岛群体(QD)中验证标记的可用性, 同时对这两个群体的遗传结构及其分化进行研究。结果表明, 从 17 447 条 BESs 中得到微卫星 3 374 个, 以四核苷酸重复为主(26.6%), 五核苷酸重复次之(17.7%), 六核苷酸重复最少(12.0%)。BES-SSR 引物的扩增效率为 77.3%(99/128), 在作图亲本中的多态比例为 33.6%(43/128), 14 个基因座在两群体中的平均等位基因数  $N_a$  分别为 18.928 6 和 26.214 3, 平均有效等位基因数  $N_e$  为 11.750 5 和 17.089 1, 平均观察杂合度  $H_o$  为 0.510 0 和 0.420 4, 平均期望杂合度  $H_e$  为 0.915 6 和 0.945 0, 多态信息含量  $PIC$  分别为 0.894 0 和 0.930 2, 群体遗传多样性水平较高。两群体间的无偏遗传相似性系数为 0.487 9, 遗传距离为 0.717 7, 平均基因分化指数  $F_{ST}$  为 0.024 3, 基因流  $N_m$  为 10.017 9, 显示群体间遗传分化程度较弱, 遗传变异主要来自于群体内个体之间, 经 Hardy-Weinberg 平衡检验, 两群体普遍存在杂合子缺失现象。研究表明, 所开发的 BES-SSR 是高度多态位点, 用于群体遗传多样性分析效果很好, 显示 BES 是微卫星标记开发和应用的资源。

**关键词:** 栉孔扇贝; BAC 末端序列; 微卫星; 遗传多样性

**中图分类号:** Q 346; S 917

**文献标志码:** A

微卫星 DNA (microsatellite DNA), 又称短串联重复序列(short tandem repeats, STRs)或简单序列重复(sequence repeats, SSRs), 是由 1~6 个核苷酸为基本单位(核心序列)首尾相连组成的串联重复序列, 是近十几年发展起来的一种第二代分子标记。SSR 标记因其核心序列的重复数不同, 引起物种间甚至种内不同个体之间出现微卫星位点的多态性<sup>[1-3]</sup>。SSR 标记普遍存在于动植物基因组中, 数量多, 分布广泛, 遗传多态信息含量丰富, 以孟德尔共显性方式遗传, 能够区分杂合子和纯合子, 可重复性好, 因此在动植物遗传育种中具有广阔的应用前景, 可用于遗传多样性研究、家系分析、物种鉴定、遗传图谱构建、QTL

定位及分子标记辅助育种等领域<sup>[4-7]</sup>。

栉孔扇贝(*Chlamys farreri* Jones et Preston)是我国重要的经济水产养殖动物,是我国海水养殖的支柱性产业之一。但随着栉孔扇贝养殖的高密度、规模化和工厂化生产的发展, 病害不断爆发, 造成大规模死亡, 经济损失十分严重<sup>[8]</sup>。张福绥等<sup>[9]</sup>研究发现栉孔扇贝大规模死亡是由多种因素造成的, 其中养殖环境恶化、种质衰退和抗病能力下降是主要原因。许多研究也表明生物的生长速度和抗病能力等生产性状与其遗传变异水平是密切相关的<sup>[10-11]</sup>。因此, 充分了解栉孔扇贝的不同地理群体遗传多样性对于我国栉孔扇贝种质资源的保护、利用及品种培育具有重要意义。

收稿日期: 2011-11-06      修回日期: 2012-02-17  
资助项目: 国家自然科学基金(30730071/C190203)  
通讯作者: 相建海, E-mail: jhxiang@ms.qdio.ac.cn

近年来,大量的栉孔扇贝微卫星标记得到开发。最传统的筛选微卫星方法是构建小片段基因组文库,筛选阳性克隆<sup>[12-13]</sup>,但这种方法工作量大,效率低。继而一种效率较高的方法得到发展,即构建微卫星富集文库筛选 SSR 标记,有效降低了 SSR 标记开发成本,也提高了 SSR 标记的阳性得率<sup>[14-15]</sup>。随着测序技术的发展,目前已经积累了大量基因组序列和 EST 序列,因此可以利用生物信息学方法对这些序列进行微卫星标记的开发。相比之下,利用 EST 进行 SSR 开发的研究较多,Roberts 等<sup>[16]</sup>和 Zhan 等<sup>[17-18]</sup>从栉孔扇贝 EST 序列中成功开发出 EST-SSR,并用于栉孔扇贝遗传图谱的构建工作中<sup>[19]</sup>。除了 EST 序列,利用 BAC(bacterial artificial chromosome)末端序列(BAC-end sequences, BES)进行 SSR 标记筛选也是一种标记开发的重要方法。这些 BES-SSR 来自于基因组序列,与 EST-SSR 相比具有较高的多态性,加之数量多,随机分布,更适用于遗传图谱构建。同时由于 BES-SSR 来自于 BAC 克隆,而 BAC 克隆可以定位于物理图谱上,因此 BES-SSR 在物理图谱与遗传图谱整合方面得到广泛的应用<sup>[20-22]</sup>。然而在栉孔扇贝中, BES-SSR 的开发研究及其在图谱构建和整合研究中的应用还未见报道。本课题组 2008 年成功构建了栉孔扇贝 BAC 文库<sup>[23-24]</sup>,最近又获得了大量 BESs,为 BES-SSR 的开发工作奠定了坚实的基础。由这些 BAC 克隆构建的物理图谱已经完成<sup>[25]</sup>,因此可以通过 BES-SSR 可以实现物理图谱和遗传图谱的整合,对栉孔扇贝的 QTL 分析、基因克隆和分子标记辅助育种以及全基因组测序都具有重要的现实意义。

本研究利用栉孔扇贝 BES 进行了微卫星标记的开发,并对大连群体和青岛群体这两个不同地理群体进行遗传多样性研究,分析其遗传结构和分化水平。这些工作是对栉孔扇贝 BAC 文库以及 BES-SSR 应用的一种尝试,同时也可以为栉孔扇贝种质资源保护和扇贝养殖业的健康发展提供理论指导。

## 1 材料与方法

### 1.1 序列来源和试验材料

从栉孔扇贝 scallop-CBE 和 scallop-CME 这两个 BAC 文库中随机挑取 10 237 个 BAC 克隆,在 ABI 3130 xl 测序仪上使用 Big Dye Terminator

V3.1 Cycle Sequencing Kit(Applied Biosystems, Foster City, CA, USA)进行双末端测序。利用 PHRED(Ewing and Green, 1998)<sup>[26-27]</sup>和 CROSS\_MATCH (<http://www.genome.washington.edu>)对测序结果进行数据处理,去除载体序列、重复序列、细菌污染以及小于 50 bp 的序列,得到的这些 BAC 末端序列(BES)即可用于栉孔扇贝 SSR 的开发。

试验所用材料为 Li 等<sup>[28]</sup>构建栉孔扇贝 AFLP 遗传图谱时所用的作图家系,包括亲本 Jd× Cg 和 93 个子代;以及采集于大连黑石礁海底的大连群体和来自于青岛崂山海域养殖笼的青岛群体。所有这些材料提取基因组 DNA,于 -80 °C 保存备用。

### 1.2 栉孔扇贝 BES 中的 SSR 分析

首先利用 tandem repeats finder (TRF, <http://tandem.bu.edu/trf/trf.html>)<sup>[29]</sup>程序对栉孔扇贝 BES 中 SSR 序列的分布进行初步分析,设置参数为重复序列  $\geq 12$  bp。其次,为了对高质量的 SSR 位点进行检索,设置更为严谨的检索参数如下:单核苷酸单元的重复数  $\geq 12$ ,二核苷酸单元的重复数  $\geq 6$ ,三核苷酸单元重复数  $\geq 4$ ,四核苷酸、五核苷酸和六核苷酸单元的重复数  $\geq 3$ 。最后,为了进一步了解 BES 中长 SSR 的情况,再设置参数为重复序列  $\geq 20$  bp,重新进行检索。

### 1.3 栉孔扇贝 BES-SSR 引物的设计和筛选

使用在线软件 Msatfinder(<http://www.genomics.ceh.ac.uk/msatfinder/>)<sup>[30]</sup>进行微卫星引物的设计和筛选。将经过 PHRED 和 CROSS\_MATCH 软件处理后的 BES 按照软件要求导入后,每个符合条件的 SSR 自动生成 5 对引物。为了更好地扩增目的 SSR,筛选引物的条件如下:SSR 长度  $\geq 18$  bp,引物长度为 18~23 bp,引物退火温度为 50~65 °C,上下引物之间退火温度相差不超过 5 °C,GC 含量为 40%~60%,扩增片段长度为 100~300 bp。引物由生工生物工程(上海)有限公司合成。

### 1.4 PCR 扩增检测引物的有效性

选择作图群体 2 个亲本 Jd、Cg 以及 3 个子代做成 DNA 池,以此 DNA 池为模板进行扩增来检测引物的有效性。PCR 扩增体系为 10  $\mu$ L,包括 5  $\mu$ L 的 PCR reaction mix<sup>HS</sup>(东盛),1  $\mu$ L 的 DNA 模板(10 ng/ $\mu$ L),上下游引物(10  $\mu$ mol/L)各 0.25  $\mu$ L, *Taq* DNA 聚合酶(2.5 U/ $\mu$ L)0.1  $\mu$ L,用水补足到 10  $\mu$ L。PCR

反应在 GeneAmp PCR System 9700 上进行, 扩增程序: 95 °C 预变性 5 min, 35 个循环(95 °C 变性 30 s,  $T_m$  °C 退火 30 s, 72 °C 延伸 45 s), 72 °C 充分延伸 10 min, 16 °C 保存。PCR 扩增产物在 1% 琼脂糖凝胶电泳上检测。

### 1.5 栉孔扇贝 BES-SSR 的多态性检测

选择有效扩增的引物对, 按照上述 PCR 体系和程序分别在作图群体父母本中进行扩增, PCR 产物利用 8% 非变性聚丙烯酰胺凝胶电泳分离带型, 检测多态性, 如果亲本之一存在无效等位基因(null 基因), 需选择几个子代进一步 PCR 扩增检测。

非变性聚丙烯酰胺凝胶电泳 80 V 恒压 3 h, EB 染色 10 min 后在紫外凝胶成像系统观察扩增片段谱带, 选用 50 bp DNA ladder 作为分子量标准。

### 1.6 栉孔扇贝遗传多样性研究

选择具有多态性的引物 14 对, 取大连群体(DL)34 只和青岛群体(QD)44 只, 进行 PCR 扩增。扩增条件和体系不变, PCR 产物经 1.0% 琼脂糖凝胶电泳, 有目的条带扩增的 PCR 产物再进行 8%(w/v) 非变性聚丙烯酰胺凝胶电泳, 利用 Quantity One\_v4.4.0.36 统计每个位点的等位基因, 通过 DataTrans 1.0<sup>[31]</sup> 将 bp 数据转换为基因型数据, 用 PopGen1.32 软件计算群体的等位基因数( $N_a$ )、有效等位基因数( $N_e$ )、观测杂合度( $H_o$ )和期望杂合度( $H_e$ )、Shannon 多样性指数( $I$ ), 并进行  $F$  检验, 计算基因分化系数  $F_{ST}$  和基因流  $N_m$ , 以及两群体

间的遗传相似指数和无偏遗传距离。利用 CerVus 3.0 软件<sup>[32-34]</sup> 计算位点的多态信息含量  $PIC$ 。通过 Genepop v 4.0.10<sup>[35]</sup> 进行 Hardy-Weinberg 平衡  $\chi^2$  检验概率值( $P$ ) 的计算(Fisher's method), 并根据基因杂合度计算 Hardy-Weinberg 平衡偏离指数 [ $D$ ,  $D=(H_o-H_e)/H_e$ ]。

## 2 结果

### 2.1 栉孔扇贝 BES-SSR 的种类和分布频率

测序所得序列用 PHRED 和 CROSS\_MATCH 软件进行处理后共得到 17 477 条 BAC 末端序列(BES), 长度为 7 773 272 bp。利用 TandemRepeats Finder(TRF)软件进行分析, 由表 1 可见, 不同的检索条件, 得到不同的 SSR 数量。当设置重复序列  $\geq 12$  bp 时, 从 7 314 BESs 中得到 8 550 个 SSR 序列, 六核苷酸 SSR 最多, 为 4 523 个, 占 51.0%; 五核苷酸 SSR 次之, 为 1 665 个占 18.8%; 三核苷酸 SSR 最少, 为 446 个, 占 5.0%。当重复单元为 1~4 bp 的 SSR 重复序列  $\geq 12$  bp、重复单元为 5 和 6 bp 的 SSR 重复数至少为 3 次时, 通过 TRF 分析获得 SSR 3 374 个, 其中单核苷酸 SSR 有 550 个, 占这 6 类 SSR 总数的 16.3%; 二核苷酸 SSR 有 480 个, 占 14.2%; 三核苷酸 SSR 有 446 个, 占 13.2%; 四核苷酸 SSR 有 896 个, 占 26.6%; 五核苷酸 SSR 有 598 个, 占 17.7%; 六核苷酸 SSR 有 404 个, 占 12.0%; 因此在栉孔扇贝的 6 种类型 BES-SSR 中, 以四核苷酸重复为主, 五核苷酸重

表 1 栉孔扇贝 BES-SSR 的分布情况  
Tab.1 SSRs distribution in *C. farreri* BESs

类型 type	第 1 组( $\geq 12$ bp) group1		第 2 组( $\geq 12$ bp) group2		第 3 组( $\geq 20$ bp) group3	
	数量 number	比例/% proportion	数量 number	比例/% proportion	数量 number	比例/% proportion
mono-nucleotide	550	6.2	550	16.3	138	11.8
di-nucleotide	480	5.4	480	14.2	131	11.2
tri-nucleotide	446	5.0	446	13.2	144	12.3
tetra-nucleotide	896	10.1	896	26.6	277	23.7
penta-nucleotide	1 665	18.8	598	17.7	203	17.4
hexa-nucleotide	4 513	51.0	404	12.0	274	23.5
total	8 550		3 374		1 167	

注: 第 1 组: 重复序列  $\geq 12$  bp; 第 2 组: 对于重复单元为 1~4 个核苷酸的 SSR, 重复序列  $\geq 12$  bp; 对于重复单元为 5 和 6 个核苷酸的 SSR, 重复数至少为 3 次; 第 3 组: 重复序列  $\geq 20$  bp。

Notes: group1: SSR length  $\geq 12$  bp; group2: for mono-, di-, tri- and tetra-nucleotides repeats, SSRs spanning more than 12 bp were counted (including 12); for penta- and hexa- nucleotides repeats, repeat times more than three were calculated (including three); group3: SSR length  $\geq 20$  bp.

复次之,六核苷酸重复最少。而当重复序列 $\geq 20$  bp时,得到1 167个SSR序列,六核苷酸SSR最多,占23.5%;五核苷酸SSR次之,占17.4%;二核苷酸SSR最少,占11.2%。

按照上述第2组的TRF检索参数,重复单元为1~6个核苷酸的SSR类型和数量分布见图1。其中,在单核苷酸重复SSR中,以A/T类型的单碱基重复为主,A/T为401个,占单核苷酸重复数的72.9%。在重复基元为二核苷酸的SSR中,有3种重复类型,分别是AT/TA、AC/GT和AG/CT。其中以AT/TA类型的重复为主,不存在GC/CG重复类型的SSR,AT/TA类型的有356个,占二核苷酸重复SSR总数的74.2%。在重复基元为三到六核苷酸的SSR中,出现频率最高的重复基元分别是AAT/ATT(37.2%)、AACC/GGTT(20.0%)、AGAAA/TTTCT(13.5%)和AGGTCA/TGACCT(6.4%)。

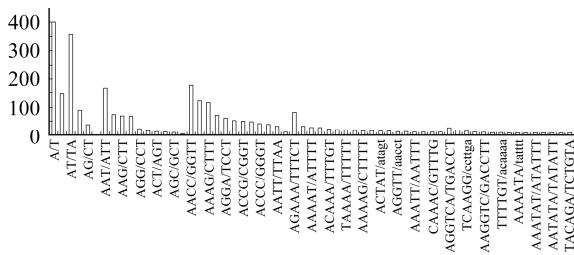


图1 重复单元为1~6核苷酸的SSR种类和数量分布  
重复单元为1~4个核苷酸的SSR,重复序列 $\geq 12$  bp; 重复单元为5和6个核苷酸的SSR,重复数至少为3次。

Fig. 1 The number and repeats unit in 1-6 bp tandem repeats

for mono-, di-, tri- and tetra-nucleotides repeats, SSRs spanning more than 12 bp were counted (including 12); for penta- and hexa-nucleotides repeats, repeat times more than three were calculated (including three).

2.2 栉孔扇贝 BES-SSR 引物的设计和筛选

将17 447条BES导入Msatfinder软件(http://www.genomics.ceh.ac.uk/msatfinder/),共搜索到1 229个SSR,其中可以进行引物设计的有974个,占总数的79.3%。根据引物的筛选条件,最终选择了128对引物用于位点的多态性检测和群体遗传分析。

2.3 引物的有效性检测

选择父母本和3个子代的混合DNA作为PCR模板,检测128对引物的有效性,产物经1%琼脂糖凝胶电泳检测,128对引物中有99对引物能扩增出目的条带,扩增效率为77.3%。部分电泳检测结果如图2。

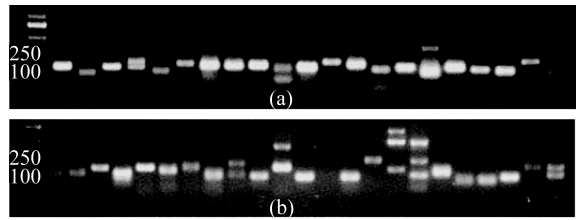


图2 部分引物在栉孔扇贝作图群体中的扩增结果电泳图

Fig. 2 Electrophoresis patterns of some loci in mapping population of *C. farreri*

2.4 栉孔扇贝 BES-SSR 多态性检测

将有效扩增的99对引物在亲本中进行PCR扩增,PCR产物经8%非变性聚丙烯酰胺凝胶电泳检测,43对引物具有多态性,可用于遗传图谱的构建。选择其中具有多态性的14对引物在大连群体和青岛群体中分别进行PCR扩增,1%琼脂糖凝胶电泳检测均能有效扩增,8%PAGE电泳进一步检测均具有多态性,这14对的引物信息见表2。

2.5 栉孔扇贝大连群体与青岛群体遗传多样性分析

栉孔扇贝两群体的遗传多样性信息见表3。14个基因座在两个群体中共产生489个等位基因,每个基因座检测到的等位基因数为12~38个,QD群体在9910基因座上检测到的基因数最多,DL群体在9852座位检测到最少的12个等位基因,平均每个位点获得的基因数为34.9286。本研究中所采用的14个位点在两个群体中的多态信息含量PIC分别为0.8940和0.9302(PIC>0.05),均为高度多态位点,可有效应用于栉孔扇贝遗传多样性分析。

两个群体的平均等位基因数 $N_a$ 分别为18.9286和26.2143,平均有效等位基因数 $N_e$ 为11.7505和17.0891,平均Shannon多样性指数 $I$ 为2.6060和2.9828,平均观察杂合度 $H_o$ 为0.5100和0.4204,期望杂合度 $H_e$ 为0.9156和0.9450,两群体的整体遗传多样性水平都较高,同时大连群体的遗传多样性略高于青岛群体。利用Hardy-Weinberg定律对两群体中每个位点的基因平衡状态进行检验,结果两群体中14个位点均极其显著偏离Hardy-Weinberg平衡( $P < 0.001$ ),遗传偏离指数 $D < 0$ 说明两群体都处于杂合子缺失状态,特别是青岛群体的9880标记偏离指数为-0.9266,说明该群体纯合子过量,杂合子严重缺失。

表 2 栉孔扇贝 14 对微卫星引物信息  
Tab. 2 Characteristics of 14 microsatellite primers in *C. farreri*

位点 locus	引物序列 primer sequences(5'-3')	退火温度/°C annealing temperature	产物/bp product size	核心重复序列 repeat motif
3942	F:GACCTTGACCAACTTCAAAGG R:CGGCAGAGCTACAAATACA	55	153~209	(TGT) <sub>11</sub>
3966	F:CGGTGTTTCAGAGGAAAAA R:TAACCGCTGGAATCTGGATA	55	191~291	(GAT) <sub>13</sub>
3978	F:TTGGGATGGATGAATCTTCG R:GGCACCGTTCAACAGAGTTT	55	110~230	(CTG) <sub>8</sub>
4030	F:CTTCCACAAGACGTTGGTC R:A GCAGCAAGAAGGCAGTCAT	55	96~155	(GCG) <sub>8</sub>
9852	F:TTGTCCTATCTGACCCAGACTC R:GGTCCTTTTCATCCTCACCA	55	99~124	(TC) <sub>9</sub>
9870	F:GCAGCATGGAAACATGCTTAT R:AAGCAGTGTTCTGCAACGAA	55	145~196	(TCT) <sub>6</sub>
9880	F:CGCAATCTTATTTGTGAAGC R:TATCGCCACGCTTTCGACTA	55	158~199	(GT) <sub>8</sub>
9886	F:TGCTTTTGGGCATAAGATGA R:TGCTCTCAAACCATCCTCA	55	197~232	(ATG) <sub>6</sub>
9892	F:GCATGGAAGTTAGCATGGAG R:ACACCCTCATAACCCACCT	55	230~272	(GT) <sub>8</sub>
9904	F:TGTTTCGGCGTGTAATAATCAA R:TGGACATCCGGTATTCTG	55	188~217	(AT) <sub>9</sub>
9910	F:GGTCAGGTCGGAGACTATTTTC R:GAATGGCACAATGCATCAAC	55	202~312	(AT) <sub>8</sub>
9930	F:TCCATGTGCGACAATGTTCT R:AGAAGGAGGGCGTGTTTAT	55	196~255	(GAA) <sub>8</sub>
9938	F:CGTCTGTACAAAAACGAA R:CTTTGGCGATAATCGGTGT	55	168~240	(ATC) <sub>10</sub>
9942	F:GCAATTTGGGGAGGTTAGGT R:CGAAATTTGGAGTCTCCTGGT	55	179~237	(GGTTA) <sub>8</sub>

## 2.6 栉孔扇贝大连群体和青岛群体间遗传结构和遗传分化分析

两个群体各位点的  $F$  统计量和基因流  $N_m$  见表 4。从中可见, 群体内平均  $F$  统计量分别为 0.493 4( $F_{IS}$ )、0.505 8( $F_{IT}$ )和 0.024 3, 14 个标记的  $F_{IS}$  都大于 0, 表示各位点处于杂合子缺失状态。两群体间的基因分化指数  $F_{ST}$  为 0.024 3, 小于 0.05, 表明两群体间的遗传分化较弱, 遗传变异主要来自于群体内个体之间。两群体间各位点基因流  $N_m$  为 6.496 8~30.976 1( $N_m \geq 4$ ), 说明遗传漂变未成为群体遗传结构变化的主要因素。

为了进一步研究栉孔扇贝大连群体和青岛群体的遗传分化程度, 我们又分析了两群体的无偏遗传相似性系数和遗传距离, 分别为 0.487 9 和 0.717 7, 可见大连群体与青岛群体的遗传距离较远, 这两群体间的遗传差异可能是由他们地理位置不同造成的。

## 3 讨论

### 3.1 栉孔扇贝中的 SSR

通过分析 BES 可以看出, 栉孔扇贝基因组中含有大量的 SSR。按照重复序列  $\geq 12$  bp 的标准来分析, 所有 17 447 条 BES, 7 314 BESs 中含有 8 550 个 SSR, 即 49.0%(8 550/17 447) 的 BES 中含有 SSR, 超过 40% 的 BES (7 314/17 447) 含有至少 1 条 SSR 序列, 每 909 bp 中含有 1 个微卫星。这远高于从 fosmid (317/3 780) 和 EST (66/3 467) 中的 SSR 分布<sup>[35-36]</sup>。即使从更为严谨的 SSR 标准, 第二组和第三组的 SSR 比例 (3 374/17 447, 1 167/17 447) 也是很高的。从微卫星的重复序列数目上看, 六碱基重复是最丰富的类型, 可以占到 SSR 的 51.0% (4 513/8 550) (重复序列  $\geq 12$  bp), 这与 fosmid 末端结果相符<sup>[34]</sup>, 然而六碱基 SSR 绝大多数大多 (91.0%) 重复 2 次左右, 因而当选取重复 3 次以上时, 六碱基 SSR 只有 404 条。从每种重复拷贝类别的重复序列数目上看, A/T 和 AT/TA 重复最多, 在其他三碱基至六碱基

表 3 栉孔扇贝大连群体和青岛群体 14 个位点的遗传多样性信息  
 Tab. 3 The polymorphic information at 14 microsatellite loci of DL and QD populations

群体 population	座位 locus	等位基因数 $N_a$	有效等位基因数 $N_e$	Shannon 多样性指数 $I$	观察杂合度 $H_o$
DL	3942	20	12.430 1	2.734 7	0.470 6
	3966	20	12.773 5	2.744 5	0.352 9
	3978	21	10.703 7	2.630 5	0.441 2
	4030	24	18.615 4	3.042 2	0.515 2
	9852	12	5.364 5	2.008 9	0.151 5
	9870	15	8.790 9	2.439 9	0.676 5
	9880	13	7.225 6	2.189 0	0.096 8
	9886	19	12.162 2	2.671 3	0.766 7
	9892	14	7.135 8	2.238 9	0.558 8
	9904	19	13.136 4	2.720 7	0.735 3
	9910	36	25.325 6	3.398 0	0.848 5
	9930	17	10.854 5	2.576 9	0.588 2
	9938	19	9.267 0	2.562 3	0.312 5
	9942	16	10.722 5	2.525 6	0.625 0
	平均 average	18.928 6±5.941 6	11.750 5±5.079 7	2.606 0±0.347 9	0.510 0±0.223 6
QD	3942	23	13.397 9	2.806 5	0.318 2
	3966	23	14.235 3	2.864 7	0.250 0
	3978	18	8.760 2	2.480 7	0.272 7
	4030	37	27.076 9	3.440 3	0.681 8
	9852	18	11.388 2	2.615 7	0.068 2
	9870	26	20.430 9	3.110 2	0.511 6
	9880	24	14.029 0	2.922 6	0.068 2
	9886	19	11.769 0	2.655 7	0.386 4
	9892	26	16.362 8	3.013 0	0.511 6
	9904	17	11.105 1	2.575 4	0.325 6
	9910	38	28.229 0	3.488 7	0.627 9
	9930	32	20.486 8	3.247 7	0.545 5
	9938	32	17.925 9	3.195 6	0.500 0
	9942	34	24.049 7	3.342 2	0.818 2
	average	26.2143±7.2343	17.0891±6.1604	2.9828±0.3323	0.4204±0.2198
群体 population	座位 locus	期望杂合度 $H_e$	多态信息含量 $PIC$	卡方检验 $P$	遗传偏离指数 $D$
DL	3942	0.933 3	0.914 0	0.000 0	-0.488 2
	3966	0.935 5	0.917 0	0.000 0	-0.617 1
	3978	0.920 1	0.899 0	0.000 0	-0.513 4
	4030	0.960 8	0.944 0	0.000 0	-0.455 6
	9852	0.826 1	0.795 0	0.000 0	-0.813 8
	9870	0.899 5	0.878 0	0.000 0	-0.236 7
	9880	0.875 7	0.847 0	0.000 0	-0.887 7
	9886	0.933 3	0.912 0	0.000 0	-0.164 6
	9892	0.872 7	0.846 0	0.000 0	-0.350 1
	9904	0.937 7	0.919 0	0.000 0	-0.204 1
	9910	0.975 3	0.962 0	0.000 0	-0.116 6
	9930	0.921 4	0.901 0	0.000 2	-0.352 1
	9938	0.906 2	0.884 0	0.000 0	-0.649 7
	9942	0.921 1	0.899 0	0.000 0	-0.310 7
	平均 average	0.915 6±0.038 2	0.894 0	0.000 0	-0.434 5
QD	3942	0.936 0	0.921 0	0.000 0	-0.656 2
	3966	0.940 4	0.926 0	0.000 0	-0.731 1
	3978	0.896 0	0.876 0	0.000 0	-0.692 1
	4030	0.974 1	0.962 0	0.000 0	-0.292 0
	9852	0.922 7	0.906 0	0.000 0	-0.925 3
	9870	0.962 2	0.949 0	0.000 0	0.462 0
	9880	0.939 4	0.925 0	0.000 0	-0.926 6
	9886	0.925 5	0.909 0	0.000 0	-0.577 8
	9892	0.949 9	0.936 0	0.000 0	-0.455 1
	9904	0.920 7	0.903 0	0.000 0	-0.642 2
	9910	0.975 9	0.964 0	0.000 0	-0.349 0
	9930	0.962 1	0.949 0	0.000 0	-0.426 6
	9938	0.955 1	0.942 0	0.000 0	-0.470 5
	9942	0.969 4	0.957 0	0.000 0	-0.146 3
	平均 average	0.945 0±0.023 4	0.930 2	0.000 0	-0.550 0

表 4 栉孔扇贝两群体的  $F$  统计量和基因流  
Tab. 4  $F$ -statistics ( $F_{ST}$ ) and gene flow ( $N_m$ ) for two populations in *C. farreri*

locus	$F_{IS}$	$F_{IT}$	$F_{ST}$	$N_m$
3942	0.572 5	0.580 1	0.018 0	13.661 7
3966	0.674 3	0.677 0	0.008 0	30.976 1
3978	0.601 7	0.613 8	0.030 5	7.955 1
4030	0.373 1	0.381 8	0.013 9	17.743 2
9852	0.872 7	0.877 4	0.036 9	6.519 9
9870	0.353 3	0.372 1	0.029 1	8.348 9
9880	0.907 9	0.910 5	0.028 7	8.449 7
9886	0.370 9	0.386 1	0.024 2	10.064 3
9892	0.404 9	0.421 7	0.028 2	8.609 1
9904	0.421 5	0.430 3	0.015 2	16.161 6
9910	0.233 1	0.241 4	0.010 8	22.841 2
9930	0.390 2	0.410 1	0.032 6	7.422 5
9938	0.557 5	0.573 9	0.037 1	6.496 8
9942	0.226 2	0.248 2	0.028 4	8.562 1
average	0.493 4	0.505 8	0.024 3	10.017 9

重复的类型中, 富含AT的SSR比例也是最高的, 这应该与栉孔扇贝基因组富含AT有关。

### 3.2 栉孔扇贝 SSR 的开发

SSR作为第二代分子标记, 在生物遗传研究和育种实践中已被广泛使用。目前已有多种SSR的开发技术应用于栉孔扇贝, 如: 直接文库筛选法、基于锚定PCR技术法、选择杂交富集法和生物信息学方法等等。战爱斌等<sup>[37]</sup>利用PCR方法筛选栉孔扇贝基因组DNA文库, 最终获得3个多态性SSR标记。Li等<sup>[38]</sup>通过构建并筛选海湾扇贝微卫星富集文库获得15对具有多态性的SSR标记。Zhan等<sup>[15]</sup>利用寡聚核苷酸(GA)<sub>20</sub>和(CA)<sub>20</sub>富集栉孔扇贝基因组DNA的AFLP预扩增产物中含有微卫星的片段, 构建微卫星富集文库, 筛选出70个具多态性的位点。

随着海洋生物基因组研究的不断开展, 通过生物信息学方法开发SSR成为一种重要途径。目前, 东方红鲂豚、文昌鱼、海胆和牡蛎等海产动物的全基因组测序已完成, 积累了大量的基因组序列和表达序列标签数据, 为其微卫星的开发提供了丰富的资源。李红蕾等<sup>[39]</sup>利用RepeatMasker软件从栉孔扇贝6 935条EST中搜索到微卫星序列, 并选用其中7条序列设计引物, 有6条可以有效扩增。Zhan等<sup>[17]</sup>利用自行设计的微卫星序列分析软件(Repeat Reporter v 1. 5), 从7 057个海湾扇贝EST序列中筛查了含有2~6个碱基重复的SSR序列

141个。选出其中10对引物在群体中检测, 结果多态位点平均扩增6.4个等位基因, 其中5个位点符合Hardy-Weinberg平衡。

本研究同样采用生物信息学的方法进行栉孔扇贝微卫星的开发, 不同的是所选用数据是栉孔扇贝BAC末端序列(BES), 属于基因组序列, 这些SSR多数处于非编码区。与EST-SSR相比后者来源于基因转录区, 具有较高的保守性, 而BES-SSR分子标记具有更丰富的遗传多态性, 广泛应用于遗传图谱的构建。而且由于BES-SSR来源于BAC末端序列, 直接与物理图谱相关联, 因此以BES-SSR分子标记构建的遗传图谱能够实现与物理图谱的整合, 从而可以极大推进QTL定位和全基因组序列组装。本研究中通过TRF软件从栉孔扇贝BAC末端序列中筛选到大量微卫星序列, 可设计引物的SSR有974个, 选择其中的128个微卫星序列在栉孔扇贝家系中进行PCR扩增, 99对(77.3%)有效扩增出SSR特征条带, 43对引物在亲本中检测出多态性, 多态性比例为33.6%。

BES-SSR高多态性的特点使其在群体遗传多样性研究和监测种群的遗传变异方面具有优势。在植物分子标记和遗传图谱构建中, Shultz等<sup>[40]</sup>利用从大豆BAC末端序列中开发的135个SSR标记实现了大豆物理图谱与遗传图谱的整合, BES-SSR在所研究的4个近交群体中的多态性比例分别为15%、18%、14%和15%。Bohra等<sup>[20]</sup>开发了842个多态性BES-SSR, 平均多态性比例为28.40%, 并构建了木豆(*Cajanus* spp.)第一张SSR标记的遗传图谱。本研究开发的BES-SSR可以促进高多态性遗传标记的收集, 对于扇贝遗传多样性及亲缘关系的研究、遗传图谱构建整合及遗传育种具有重要的实践和现实意义。

### 3.3 群体遗传多样性和遗传分化

遗传多样性的最直接表现形式就是遗传变异性的分布, 同时也包括遗传变异的分布, 即群体的遗传结构。本研究中两群体的平均多态信息含量分别为0.894 0和0.930 2, 根据Botstein<sup>[41]</sup>提出的衡量标准, 当 $PIC > 0.5$ 时, 意味着该位点为高度多态位点, 说明大连群体和青岛群体都具有非常丰富的遗传多样性。基因杂合度是衡量群体遗传变异水平的理想参数, 杂合度越高, 说明群体的遗传结构越复杂。14个位点在两群体中的平均

观察基因杂合度为 0.080 0~0.736 8, 只有位点 3966、位点 9852 和位点 9880 杂合度较低(分别是 0.294 9、0.103 9 和 0.080 0), 其他多数位点具有较高的基因杂合度, 因此可以认为大连群体和青岛群体的杂合度较高, 群体内遗传变异水平较大。两群体的平均观察杂合度分别为 0.510 0 和 0.420 4, 大连群体的杂合度略高于青岛群体, 但差别不大, 说明两群体间遗传结构的复杂程度相近。

本研究中各位点在两群体中均表现出杂合子缺失现象( $F_{IS}>0$ ), 而且 Hardy-Weinberg 平衡分析显示, 14 个位点在两群体中均表现出显著的偏离 H-W 平衡,  $H_o$ 、 $H_e$  量化了基因的遗传多样性, 平衡偏离指数  $D$  值则反映两者的平衡关系,  $D$  值越接近于零, 基因型分布越接近于平衡状态,  $D$  值越偏离零, 基因型分布越偏离平衡状态, 虽然目前还没有一个公认的标准来衡量  $D$  值达到多少被认为是偏离了 Hardy-Weinberg 平衡, 但是遗传偏离指数方法的正负直观地反映了种群内杂合子的过剩或缺失状态, 两种方法相结合更能正确、全面地反映种群的遗传平衡状况。结果表明  $D$  值偏离零, 与固定指数  $F_{IS}$  反应的结果一致, 说明两群体中有杂合子缺失现象。本研究在作图亲本中进行 SSR 多态性分析时发现存在无效等位基因(null allele), 这可能也是导致杂合子缺失的原因之一, 而普遍存在的杂合子缺失现象可能正是 Hardy-Weinberg 平衡偏离的主要原因。此外, 也可能是由人工养殖和过度捕捞等人为因素干扰造成的。结合遗传多态信息含量分析, 可见大连群体和青岛群体均具有较高的遗传多样性, 但也存在一定程度的信息丢失, 这对于栉孔扇贝种质资源的保护和苗种培育提供了有效的背景资料。

群体间遗传分化指数  $F_{ST}$  是衡量群体间遗传分化程度的重要指标。Wright<sup>[42]</sup>将  $F_{ST}$  划分为 4 个范围,  $F_{ST}$  为 0~0.05 时, 群体间遗传分化程度较弱; 0.05< $F_{ST}$ <0.15 时, 群体间遗传分化程度中等; 0.15< $F_{ST}$ <0.25 时, 群体间遗传分化程度较大;  $F_{ST}>0.25$  时, 群体间遗传分化程度极大。本研究中两群体的遗传分化指数  $F_{ST}$  为 0.008 0~0.036 9, 都小于 0.05, 平均遗传分化指数为 0.024 3, 因此可以说大连群体和青岛群体的遗传分化水平较低, 遗传变异主要来自于群体内个体之间。

基因流( $N_m$ )指生物个体从其发生地分散出去

而导致不同种群之间基因交流的过程, 其基本作用是消弱了种群间的遗传差异。Wright<sup>[43]</sup>认为, 当  $N_m>1$  时, 基因流就可以防止由遗传漂变引起的群体间的遗传分化。本研究中  $N_m>1$ , 说明遗传漂变未成为群体遗传结构变化的主要因素。两群体间无偏遗传距离为 0.717 7, 遗传相似度为 0.487 9, 两群体间的遗传距离较大, 主要可能是地理位置的原因导致两群体间的遗传分化。本研究结果也提示, 在进行栉孔扇贝育种工作中, 可选取遗传距离较远且性状优良的两群体进行杂交, 利用杂交优势获得优良新品种。

#### 参考文献:

- [1] Tautz D. Hypervariability of simple sequences as a general source for polymorphic DNA markers [J]. *Nucleic Acids Research*, 1989, 17(16): 6463-6471.
- [2] Toth G, Gaspari Z, Jurka J. Microsatellites in different eukaryotic genomes: survey and analysis [J]. *Genome Research*, 2000, 10(7): 967-981.
- [3] Gur-Arie R, Cohen C J, Eitan Y, et al. Simple sequence repeats in *Escherichia coli*: abundance, distribution, composition, and polymorphism [J]. *Genome Research*, 2000, 10(1): 62-71.
- [4] Innan H, Terauchi R, Miyashita N T. Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana* [J]. *Genetics*, 1997, 146(4): 1441-1452.
- [5] Somers D J, Isaac P, Edwards K. A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.) [J]. *Theoretical and Applied Genetics*, 2004, 109(6): 1105-1114.
- [6] McCouch S R, Teytelman L, Xu Y, et al. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.) [J]. *DNA Research*, 2002, 9(6): 199-207.
- [7] Morgante M, Hanafey M, Powell W. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes [J]. *Nature Genetics*, 2002, 30(2): 194-200.
- [8] 王运涛, 相建海. 栉孔扇贝大规模死亡的原因探讨[J]. *海洋与湖沼*, 1999, 30(6): 770-774.
- [9] 张福绥, 杨红生. 栉孔扇贝大规模死亡问题的对策与应急措施[J]. *海洋科学*, 1999(2): 1-5.
- [10] Ferguson M M, Drahuschak L R. Disease resistance and enzyme heterozygosity in rainbow-trout [J]. *Heredity*, 1990, 64: 413-417.
- [11] Garton D W, Koehn R K, Scott T M. Multiple-locus heterozygosity and the physiological energetics of growth in the coot clam, *Mulinia lateralis*, from a natural population [J]. *Genetics*, 1984, 108(2): 445-455.



- [12] Zhan A B, Bao Z M, Yao B, *et al.* Polymorphic microsatellite markers in the Zhikong scallop *Chlamys farreri* [J]. *Molecular Ecology Notes*, 2006, 6(1): 127–129.
- [13] Zhan A B, Bao Z M, Hu X L, *et al.* Isolation and characterization of 150 novel microsatellite markers for Zhikong scallop (*Chlamys farreri*) [J]. *Molecular Ecology Notes*, 2007, 7(6): 1015–1022.
- [14] Zhan A B, Hu X L, Bao L S, *et al.* Molecular identification of scallop planktonic larvae using species-specific microsatellites[J]. *Acta Oceanologica Sinica*, 2008, 27(5): 134–146.
- [15] Zhan A B, Bao Z M, Hu X L, *et al.* Characterization of 95 novel microsatellite markers for Zhikong scallop *Chlamys farreri* using FIASCO-colony hybridization and EST database mining [J]. *Fisheries Science*, 2008, 74(3): 516–526.
- [16] Roberts S, Romano C, Gerlach G. Characterization of EST derived SSRs from the bay scallop, *Argopecten irradians*[J]. *Molecular Ecology Notes*, 2005, 5(3): 567–568.
- [17] Zhan A, Hu J, Wang X, *et al.* A panel of polymorphic EST-derived microsatellite loci for the bay scallop (*Argopecten irradians*) [J]. *The Journal of Molluscan Studies*, 2006, 72(4): 435–438.
- [18] Zhan A B, Hui Z M, Wang M L, *et al.* Inheritance pattern of EST-SSRs in self-fertilized larvae of the bay scallop *Argopecten irradians* [J]. *Annales Zoologici Fennici*, 2007, 44: 259–268.
- [19] Bao Z, Zhan A, Hu J, *et al.* Construction of microsatellite-based linkage maps and identification of size-related quantitative trait loci for Zhikong scallop (*Chlamys farreri*)[J]. *Animal Genetics*, 2009, 40(6): 821–831.
- [20] Bohra A, Dubey A, Saxena R K, *et al.* Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus spp.*)[J]. *BMC Plant Biology*, 2011, 11(1): 56.
- [21] Xu J, Qian X, Wang X, *et al.* Construction of an integrated genetic linkage map for the A genome of *Brassica napus* using SSR markers derived from sequenced BACs in *B. rapa*[J]. *BMC Genomics*, 2010, 11: 594.
- [22] Cordoba J M, Chavarro C, Schlueter J A, *et al.* Integration of physical and genetic maps of common bean through BAC-derived microsatellite markers[J]. *BMC Genomics*, 2010, 11:436.
- [23] Zhang Y, Zhang X, Scheuring C F, *et al.* Construction and characterization of two bacterial artificial chromosome libraries of Zhikong scallop, *Chlamys farreri* Jones et Preston, and identification of BAC clones containing the genes involved in its innate immune system [J]. *Marine Biotechnology* (NY), 2008, 10(4): 358–365.
- [24] Zhang Y, Zhang X J, Scheuring C F, *et al.* Construction of bacterial artificial chromosome libraries for Zhikong Scallop *Chlamys farreri*[J]. *Chinese Journal of Oceanology and Limnology*, 2008, 26(2): 215–218.
- [25] Zhang X, Zhao C, Huang C, *et al.* A BAC-Based physical map of Zhikong scallop (*Chlamys farreri* Jones et Preston) [J]. *PLoS ONE*, 2011, 6(11): e27612.
- [26] Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities [J]. *Genome Research*, 1998, 8(3): 186–194.
- [27] Ewing B, Hillier L, Wendl M C, *et al.* Base-calling of automated sequencer traces using phred. I. Accuracy assessment [J]. *Genome Research*, 1998, 8(3): 175–185.
- [28] Li L, Xiang J, Liu X, *et al.* Construction of AFLP-based genetic linkage map for Zhikong scallop, *Chlamys farreri* Jones et Preston and mapping of sex-linked markers [J]. *Aquaculture*, 2005, 245(1-4): 63–73.
- [29] Benson G. Tandem repeats finder: a program to analyze DNA sequences [J]. *Nucleic Acids Research*, 1999, 27(2): 573–580.
- [30] Thurston M I, Field D. Msatfinder: detection and characterisation of microsatellites[OL]. 2005, <http://www.genomics.ceh.ac.uk/msatfinder/>.
- [31] 盖红梅, 任民. SSR 数据处理宏程序 DataTrans 1.0 [J]. *分子植物育种*, 2011, 9: 1359–1365.
- [32] Kalinowski S T, Taper M L, Marshall T C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment [J]. *Molecular Ecology*, 2007, 16(5): 1099–1106.
- [33] Marshall T C, Slate J, Kruuk L E B, *et al.* Statistical confidence for likelihood-based paternity inference in natural populations [J]. *Molecular Ecology*, 1998, 7(5): 639–655.
- [34] Slate J, Marshall T, Pemberton J. A retrospective assessment of the accuracy of the paternity inference program CERVUS [J]. *Molecular Ecology*, 2000, 9(6): 801–808.
- [35] Rousset F. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux [J]. *Molecular Ecology Research*, 2008, 8(1): 103–106.
- [36] 程洁, 张玲玲, 黄晓婷, 等. 栉孔扇贝 Fosmid 文库的构建及基因组结构特征分析[J]. *中国海洋大学学报: 自然科学版*, 2008, 38: 78–88.
- [37] 战爱斌. 栉孔扇贝(*Chlamys farreri*)微卫星标记的筛选及应用[D]. 青岛: 中国海洋大学, 2007.
- [38] Li H, Liu X, Hu J, *et al.* PRIMER NOTE: A set of polymorphic microsatellite loci for the bay scallop, *Argopecten irradians*[J]. *Molecular Ecology Notes*, 2006, 7(3): 422–424.
- [39] 李红蕾, 宋林生, 王玲玲, 等. 栉孔扇贝 EST 中微卫星标记的筛选[J]. *高技术通讯*, 2003, 12: 72–75.
- [40] Shultz J L, Kazi S, Bashir R, *et al.* The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean [J]. *Theoretical and Applied Genetics*, 2007, 114(6):

- 1081–1090.
- [41] Botstein D. Construction of a genetic linkage map in man using restriction fragment length polymorphisms [J]. *American Journal of Human Genetics*, 1980, 32: 314–331.
- [42] Wright S. Evolution in Mendelian population [J]. *Genetics*, 1931, 16: 97–159.
- [43] Wright S. The genetical structure of populations [J]. *Annals of Eugenics*, 1951, 15: 323–354.

## The development of BAC-end sequence-based microsatellite markers and analysis on population genetic diversity in Zhikong scallop (*Chlamys farreri*)

ZHANG Xiu-ying<sup>1,2</sup>, ZHANG Xiao-jun<sup>1</sup>, ZHAO Cui<sup>1,2</sup>, CHANG Ya-qing<sup>3</sup>, HUAN Pin<sup>1</sup>,  
LI Fu-hua<sup>1</sup>, XIANG Jian-hai<sup>1\*</sup>

(1. *The Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China;*

2. *Graduate University of the Chinese Academy of Sciences, Beijing 100039, China;*

3. *College of Fisheries and Life Science, Dalian Ocean University, Dalian 116023, China)*

**Abstract:** In this study, microsatellite markers were developed from the BAC-end sequences and used to analyze the genetic structure and genetic differentiation in two populations in Zhikong scallop (*Chlamys farreri*). A total of 3 374 microsatellites were identified from the 17 477 BAC-end sequences (BESs), of which tetra-nucleotide motifs were most abundant (26.6%), followed by penta-nucleotide motifs (17.7%) and hexa-nucleotide motifs were the least (12.0%). 77.3% SSRs were successfully amplified (99/128), and 43 SSRs (33.6%) were polymorphic in the parentage of mapping population. In order to apply these BES-SSR markers, 14 polymorphic SSRs were chosen to amplify and then analyze the genetic diversity in Dalian population and Qingdao population. A total of 395 alleles were obtained at the fourteen microsatellite markers and the number of alleles in each locus ranged from 8 to 38 in the two populations. The average number of alleles ( $N_a$ ) was 18.928 6 and 26.214 3 respectively. The average effective number of alleles ( $N_e$ ) was 11.750 5 and 17.089 1 respectively. The mean observed heterozygosity ( $H_o$ ) was 0.510 0 and 0.420 4. The mean expected heterozygosity ( $H_e$ ) was 0.915 6 and 0.945 0. The data suggested both populations have high genetic diversities. The mean polymorphic information content ( $PIC$ ) was 0.894 0 and 0.930 2, which were both greater than 0.5, indicating the fourteen loci were highly polymorphic. The unbiased genetic identity index was 0.487 9, and the genetic distance was 0.717 7. The coefficient of gene differentiation ( $F_{ST}$ ) and gene flow ( $N_m$ ) between two populations were 0.024 3 and 10.017 9 respectively. Low genetic differentiation was observed between the two populations, and the variance mainly came from individual difference. Significant deviation was detected by Hardy-Weinberg equilibrium test. There was heterozygote deficiency at all loci. The results showed BAC-end sequences were an effective resource for development of SSR markers for genetic and genomic researches.

**Key words:** *Chlamys farreri*; BAC-end sequences (BES); microsatellite; genetic diversity

**Corresponding author:** XIANG Jian-hai. E-mail: jhxiang@ms.qdio.ac.cn