

长牡蛎 17 个 fosmid-SSR 标记的开发与分析

李慧娟¹, 亓海刚², 李莉², 闫喜武^{1*}, 张国范²

(1. 大连海洋大学生命科学与技术学院, 辽宁 大连 116023;

2. 中国科学院海洋研究所, 山东 青岛 266071)

摘要: 研究所用序列由长牡蛎 fosmid 文库末端测序获得。首先用 TRF 程序扫描序列获得一批重复单元为 2 碱基的候选 SSR 位点, 然后在部分 SSR 序列两侧的保守区设计 50 对引物, 对取自山东青岛的一个长牡蛎野生群体进行基因型分析。结果显示, 共有 17 个 SSR 位点显示多态性, 等位基因数 (N_a) 平均为 4, 有效等位基因数 (N_e) 平均为 2.32, 平均杂合度观测值 (H_o) 和期望值 (H_e) 分别为 0.395 9 和 0.628 8。其中, 11 个位点的多态信息含量 (PIC) 值均大于 0.5, 共有 49 个等位基因, 适合对长牡蛎群体遗传结构的分析; 6 个位点 $0.25 < PIC < 0.5$, 为中度多态位点。 χ^2 检验估计 Hardy-Weinberg 平衡, 经 Sequential Bonferroni 校正后, 除 DEC_1 以外, 其他位点仍偏离平衡。将 SSR 附近 10 000 bp 内基因组预测的编码蛋白通过与 NCBI 的 nr 库 blastp 分析, 共有 13 条微卫星序列获得连锁的相关基因功能注释。

关键词: 长牡蛎; 简单重复序列; 基因组; 遗传多样性

中图分类号: Q 786; S 917.4

文献标志码: A

长牡蛎 (*Crassostrea gigas*) 又称太平洋牡蛎, 属软体动物门 (Mollusca)、瓣鳃纲 (Lamellibranchia)、翼形亚纲 (Pteriomorphia)、珍珠贝目 (Pterioidea)、牡蛎科 (Ostreidae), 是一种广盐、广温的种类, 广泛分布于我国北方海区, 其肉味鲜美、营养丰富, 具有很高的经济价值。据 FAO 统计, 2008 年世界牡蛎养殖产量为 4 164 010 t, 我国牡蛎产量 335 万 t, 长牡蛎至少占 15.6%。目前, 牡蛎产业也同样面临着育种方法、病害防治等问题, 而结合现代生物技术, 系统深入地开展长牡蛎遗传多样性研究, 培育出生长快、品质好、抗逆性强、成活率高和肉壳比率高的新品种^[1], 将是实现经济贝类资源合理开发与持续利用的有效途径。

微卫星标记又称简单重复序列 (simple sequence repeat, SSR)^[2], 具有较高稳定性、多态性、检测方便^[3]、引物通用性、位点特异性和分析

操作简单等^[4]特点, 已广泛应用于水产动物的遗传多样性分析^[5]、寻找与优良的性状相连锁的基因进行基因定位^[6]、建立遗传连锁图^[7-8]、鉴定亲缘关系^[9]等。于红^[10-11]对 GenBank 的 EST-SSR 进行筛选, 聚类拼接去除冗余后得到 130 条微卫星表达序列, 用 20 个 EST-SSRs 标记对 3 个长牡蛎家系进行遗传分离模式分析, 有 5 组显著偏离孟德尔分离定律, 其中 4 组采用无效的等位基因调整后符合孟德尔遗传, 并用 7 个多态性较高的 SSR 标记分析了 10 个长牡蛎群体的遗传多样性和群体间遗传分化, Hardy-Weinberg 平衡 (HWE) 多位点检测中有 28 组偏离平衡。YU 等^[12]用 AFLP 和 5 个微卫星标记对 4 个养殖群体和一个野生群体进行了遗传学分析, 发现在两个抗病群体中含有较少的等位基因数目, 可进一步分析这些位点是否和某些抗病基因相关。REECE 等^[13]报道了美洲牡蛎 26 个微卫星分离中有 8 个分离

收稿日期: 2011-02-12 修回日期: 2011-05-06

资助项目: 国家“九七三”重点基础研究发展计划 (2010CB126402); 国家自然科学基金重点项目 (40730845); 国家公益性行业 (农业) 科研专项 (3-53)

通讯作者: 闫喜武, E-mail: yanxiwu2002@163.com

显著偏离孟德尔规律。HUBERT 等^[8]利用 102 个 SSR 标记对 3 个家系构建出微卫星连锁图,雄性图谱包括 88 个标记,覆盖 616.1 cM,雌性图谱包括 86 个标记,覆盖 770.5 cM。经过校正,可达到长牡蛎的单倍体的 10 个连锁群的分布。2009 年,HUBERT 等^[14]又选取 56 个卫星标记和淀粉酶基因对 7 个三倍体牡蛎家系进行着丝粒作图,测定了基因和标记与着丝粒间的距离。

2010 年 7 月 31 日,中国科学院海洋研究所在青岛宣布牡蛎基因组序列图谱绘制完成。伴随长牡蛎基因组水平的深入研究和重要功能基因的批量发掘,为重要经济性状的解析、新品种的分子选育提供理论基础。本实验初步探讨了 fosmid 文库测序序列的 SSR 开发,为下一步大量 SSR 筛查奠定基础。

1 材料与方 法

1.1 SSR 引物开发

序列来源和候选 SSR 位点的获得 实验所用数据为长牡蛎基因组 fosmid 末端测序所得部分数据。用 TRF(Tandem Repeat Finder)软件分析 fosmid 序列,二碱基 SSR 的筛查标准为重复单元基至少重复 6 次,并挑选 SSR 两侧翼大于 50 bp 的序列。

长牡蛎 SSR 引物设计 利用软件 Primer 5.0 和 Oligo 6.0 在 SSR 两端序列保守区内设计和评价引物,引物长度 18~25 bp,GC 含量 40%~60%,由上海生工生物工程有限公司合成。

1.2 DNA 样品的采集与处理

取青岛 48 只野生长牡蛎闭壳肌,采用常规“酚—氯仿”抽提法提取其基因组 DNA。用 1.0% 琼脂糖凝胶电泳检测所提取的 DNA 质量,并用紫外分光光度计测定 DNA 浓度,-20℃保存备用。把 DNA 工作液调整浓度至 30 ng/ μ L,4℃保存。

1.3 长牡蛎候选 SSR 位点检测

先用 6~8 个野生长牡蛎 DNA 对所有引物进行初筛,然后用于其群体分析。15 μ L 反应体系为 20~30 ng DNA、0.75 U Taq DNA 聚合酶、10 mmol/L dNTP、10 \times PCR Buffer、100 μ mol/L 引物。PCR 循环参数设置为 94℃ 4 min,94℃ 40 s,最适退火温度下 30 s,72℃ 30 s,共 35 个循环,循环结束后 72℃ 延伸 10 min,10℃保存。扩增产物经

12% 的非变性聚丙烯酰胺凝胶电泳, GelRed 染色。

2 结果

2.1 长牡蛎 SSR 位点的多态性推测

实验选自组序列的 50 个 SSR 位点基序长度主要集中在 12~16 bp,基序最长的为 40 bp,为二核苷酸基元的 20 次重复。与地高辛标记的二碱基单元(CA)15 作探针筛选澳洲鳗鲡(*Anguilla australis*)^[15]的部分基因组文库所得的 SSR 标记基序长度的结果一致(绝大部分重复次数在 10 次以下)。本研究所得的二碱基重复基元属于低级基元,依照当 SSR 基序长度长于 20 bp 时多态性较高,长度在 12~20 bp 时多态性中等,而长度在 12 bp 以下时多态性极低^[16]的标准,可推测本实验的大多数 SSR 位点具有中等多态性,DEC06_47 具有较高的多态性。

2.2 长牡蛎 SSR 引物筛选结果

在初筛结果中,共有 22 对引物出现扩增失败,4 对引物的产物大小与目的片段不符,相差 100 bp 以上。把其余 24 对引物用于自然群体分析,其中 2 对引物出现的产物单一条带,5 对引物只在少数个体中扩增出目的条带,经反复试验结果不变。共有 17 对 SSR 引物在野生群体中扩增出了稳定、清晰的 DNA 目的片段并在个体间表现出不同程度的多态性,能准确对长牡蛎野生群体进行基因分型(表 1)。

2.3 野生长牡蛎群体 17 个 SSR 位点分析

用 PopGene 软件统计每个微卫星位点的等位基因数(observed number of alleles, N_a),有效等位基因数(effective number of alleles, N_e),观测杂合度(observed heterozygosity, H_o),期望杂合度(expected heterozygosity, H_e),多态信息含量(polymorphism information content, PIC)(表 2)。

多态信息含量(PIC)是衡量基因变异程度的重要指标^[17]。在本研究中,17 个 SSR 位点中的 11 个的 PIC 值均大于 0.5,共有 49 个等位基因,适合对长牡蛎野生群体遗传结构的分析;6 个位点 $0.25 < PIC < 0.5$,为中度多态位点。 N_a 、 N_e 、 H_o 和 H_e 等能反映群体遗传多样性,数值越大基因丰富度越高。 N_e 是反映群体遗传变异大小的一个指标,等位基因在群体中分布越均匀,其数值越接近检测到的 N_a 。本研究中, N_a 平均为 4, N_e 平均为 2.82, H_o 和 H_e 分别为 0.395 9 和 0.628 8。 χ^2

检验估计 Hardy-Weinberg 平衡,对结果采用 Sequential Bonferroni 校正,结果显示在长牡蛎野生群体中除 DEC_1 以外,其余位点仍偏离平衡。

2.4 17 个 SSR 位点的连锁标记

将这些微卫星序列在 NCBI 网站进行 blastx

分析后发现无阳性匹配,将 SSR 附近 10 000 bp 内基因组预测的编码蛋白通过与 NCBI 的 nr 库 blastp 分析,共有 13 条微卫星序列获得连锁的相关基因的功能注释(表 3)。

表 1 17 对微卫星引物序列情况
Tab.1 17 pairs of SSR primer sequence

| 位点 locus | 重复序列 repetitive sequence | 引物序列 primer sequence | |
|-------------|-----------------------------|---------------------------|----------------------------|
| DEC06_1 | (TA)8 | CACCAGATTGATGGTTTGAACAGA | CCAAGGTAATGGCATGTAGCAAT |
| DEC06_4 | (AT)6 | CGACCACATACAGAGAAATGTGA | ATCCAAGACGGAACATAGGTAA |
| DEC06_7 | (TA)6 | AGCTAGATTGATAAGCGGATGGA | AAGTGACTAATACACGCAAAGGG |
| DEC06_8 | (CA)6 | CCAGTGACTTCAAATTGGTTTCC | TTGCTAGTGAGACCACTGATTGA |
| DEC06_11 | (AG)11 | GTGCCACATTTCTGTGTTAT | ATTCAGGGACTATCTGATTGGA |
| DEC06_12 | (AT)7 | CAGATCACAGCGGATATGTAGGA | GAATTTGTTCTTGCCCATGTC |
| DEC06_13 | (AG)6 | CAATGTATAGCTCAGGAAAGAGG | GCCATTACATAAGTTTGCCCATC |
| DEC06_19 | (AC)6 | GCTGTCTAAACCTTACTACACAAGT | TTTAGTTGCTCAGGTGAACGATG |
| DEC06_29 | (AT)7 | CTGCCTGTTCTTATTACCCAAG | GTAGGAAACGATGGAAAGCCAAA |
| DEC06_31 | (GA)8 | TACTCGCTTAGGAAGTGGTGT | CGAGTCACAGCTCGGTTTAATTC |
| DEC06_32 | (AT)7 | CAGGAATACCGATGTAGAAAGT | ATATACCGACGAATACGGAAGA |
| DEC06_34 | (AT)8 | GACACACCGCAAATCTGTCATC | CATACACGACTCTTCCGTTTGG |
| DEC06_38 | (CT)7 | TGTTGCGTACTAGCATATCTCCA | CCTTGTGTTTCATTAGGAGTCTACAG |
| DEC06_40 | (AC)7 | CTCTGACAACAAGGGTGAAGGG | GGGCACTTTGTGAAAGAAATCATAG |
| DEC06_43 | (TC)7 | ACCTTATCTTCTCAAGTCCT | GTTCAAGTCAGGTGTAACGAAATAG |
| DEC06_46 | (GA)8 | AAGGGAAAGCCCTACATAAAGGT | CCAAGATGGCCGTAAGTAACAAA |
| DEC06_47 | (TC)20 | AATCCAAACAGGCATCTCAAGTG | GCAAGAGAGAGACAGAAAGAGAG |

表 2 长牡蛎野生群体的遗传多态性及 Hardy-Weinberg 平衡检验
Tab.2 Genetic diversity and P-values of *C. gigas* wild population

| 位点 locus | N_a | N_e | H_o | H_e | P | PIC |
|----------|-------|---------|---------|---------|-----------|-----------|
| DEC_1 | 2 | 1.967 7 | 0.461 5 | 0.498 2 | 0.641 644 | 0.370 857 |
| DEC_4 | 2 | 1.777 1 | 0.229 2 | 0.441 9 | 0.000 722 | 0.341 668 |
| DEC_7 | 4 | 3.474 5 | 0.521 7 | 0.720 0 | 0.017 639 | 0.659 704 |
| DEC_8 | 4 | 3.100 4 | 0.553 2 | 0.684 7 | 0.000 259 | 0.615 627 |
| DEC_11 | 6 | 4.322 9 | 0.574 5 | 0.776 9 | 0.006 625 | 0.615 054 |
| DEC_12 | 5 | 3.035 5 | 0.378 4 | 0.679 7 | 0.000 000 | 0.488 129 |
| DEC_13 | 4 | 2.227 8 | 0.172 4 | 0.560 8 | 0.000 000 | 0.470 838 |
| DEC_19 | 4 | 2.045 7 | 0.179 5 | 0.517 8 | 0.000 000 | 0.621 227 |
| DEC_29 | 4 | 3.112 5 | 0.522 7 | 0.686 5 | 0.001 961 | 0.731 051 |
| DEC_31 | 4 | 3.038 5 | 0.446 8 | 0.678 1 | 0.000 262 | 0.622 869 |
| DEC_32 | 4 | 2.965 3 | 0.916 7 | 0.669 7 | 0.000 000 | 0.598 673 |
| DEC_34 | 4 | 2.122 5 | 0.479 2 | 0.534 4 | 0.000 000 | 0.482 623 |
| DEC_38 | 5 | 3.785 1 | 0.186 0 | 0.744 5 | 0.000 000 | 0.688 691 |
| DEC_40 | 5 | 3.049 0 | 0.468 1 | 0.679 2 | 0.000 000 | 0.612 819 |
| DEC_43 | 3 | 2.101 2 | 0.020 8 | 0.529 6 | 0.000 001 | 0.431 087 |
| DEC_46 | 4 | 3.589 0 | 0.425 5 | 0.729 1 | 0.000 000 | 0.669 195 |
| DEC_47 | 4 | 2.224 9 | 0.194 4 | 0.558 3 | 0.000 000 | 0.503 181 |
| 平均值 mean | 4 | 2.820 0 | 0.395 9 | 0.628 8 | — | — |

表 3 13 条微卫星序列功能注释
Tab.3 Functional annotations of 13 SSR

| 位点 locus | 预测功能 functional annotations | E 值 E-value | 登录号 accession no. |
|-------------|--|----------------|----------------------|
| DEC06_1 | 生发中心激酶 III (germinal centre kinase III) | 9.00E-143 | NP_650596.1 |
| DEC06_7 | 预测:与 GA16498-PA 相似 | 2.00E-50 | XP_974849.2 |
| DEC06_8 | Hox 蛋白:LIM/homeobox protein Lhx2 | 6.00E-74 | NP_990220.1 |
| DEC06_13 | 囊泡蛋白分拣相关蛋白(vacuolar protein sorting-associated protein) | 0.00E+00 | XP_002427738.1 |
| DEC06_19 | 葡萄糖抑制的分裂蛋白 A(glucose inhibited division protein A) | 5.00E-98 | XP_002423494.1 |
| DEC06_29 | 有机物-负离子转运子(organic anion transporter) | 0.00E+00 | XP_002433182.1 |
| DEC06_31 | 类似于脱脂叶绿素 A 结合蛋白的前体(chlorophyllide A binding protein precursor) | 2.00E-07 | XP_789199.2 |
| DEC06_32 | 线粒体基质转运蛋白(mitochondrial solute carrier protein) | 1.00E-57 | ABF18426.1 |
| DEC06_34 | 预测:一种假设蛋白 | 5.00E-104 | XP_001198541.1 |
| DEC06_38 | ρ /rac 鸟苷酸交换因子 18(rho/rac guanine nucleotide exchange factor(GEF)18) | 7.00E-36 | XP_002192527.1 |
| DEC06_40 | 钠离子渗出通道的非选择性蛋白(Sodium leak channel non-selective protein) | 0.00E+00 | EFN65909.1 |
| DEC06_46 | PREDICTED:类犬尿氨酸酶(kynureninase-like) | 1.00E-151 | XP_002937401.1 |
| DEC06_47 | 一种假设蛋白:BRAFLDRAFT_120093 | 0.00E+00 | XP_002606640.1 |

3 讨论

本研究在长牡蛎野生群体中发现,5 对引物只在自然群体的极少数个体中扩增出产物,大部分个体无 PCR 产物,在此作为无效引物;本研究在长牡蛎野生群体中发现 DEC_12、DEC_13、DEC_38、DEC_47 存在无效等位基因,部分位点重新设计引物后能够消除无效等位基因。这说明与微卫星多变的核重复序列相比较,微卫星侧翼序列只是具有相对保守性,在长牡蛎漫长进化过程中侧翼序列可能会发生较大变化,引物 3' 区域如果设计在侧翼突变区,可能会出现部分个体无 PCR 产物的现象。当两等位基因长度差距大时,会出现长等位基因扩增弱于短等位基因的现象,通过优化长等位基因的扩增条件^[18]能有效避免长等位基因扩增劣势导致的基因分型错误。

13 条微卫星序列附近 10 000 bp 内基因组预测的编码蛋白通过与 NCBI 的 nr 库 blastp 分析获得功能注释,结果表明,SSR 位点与长牡蛎生长发育和新陈代谢相关基因有关联,可以很好地作为分子标记。例如:Hox 蛋白在决定昆虫的体节规定性过程中起着重要作用^[19],并能够影响昆虫的形态进化。犬尿氨酸酶在人^[20]和小鼠^[21]中主要是色氨酸代谢通路中涉及血压调节的关键成分,对于生活在潮间带的长牡蛎受潮汐变化的影响,必然有一套完善的渗透压调节系统,这可能与犬尿氨酸酶有关。

本研究筛选的 17 个 SSR 位点中 11 个位点

PIC 值 > 0.5 具有高度多态信息含量,数量多于长牡蛎 SSR 位点的多态性推测,猜测可能与长牡蛎本身的高变异性有关。卡方检验 Hardy-Weinberg 平衡分析,经 Sequential Bonferroni 校正后除 DEC_1 没有显著偏离外,其它位点的基因型分布均不同程度偏离平衡。这可能由于长牡蛎附在海岸礁石上生长,对采样地点的过渡采挖、产卵场遭到破坏、水体污染、周围生态环境人为改变或者海上养殖群体缺乏种质保护与管理,致使长牡蛎群体规模和基因型频率发生了较大的变化,导致遗传漂变现象明显。

微卫星标记已成为现代遗传学研究的有利手段,为长牡蛎遗传多样性分析、定向遗传育种、构建遗传图谱及杂交鉴定等提供了可能。长牡蛎基因组的完成不但有助于牡蛎养殖产业的健康可持续发展,而且也有利于牡蛎种质资源的保护。

感谢王家丰、韩斐斐和余智彩同学在本实验过程中给予的大力支持和帮助。

参考文献:

- [1] 刘小林,相建海. 重要经济贝类选择育种及遗传力研究进展[J]. 海洋科学,2003,27(6):15-20.
- [2] ZIETKIEWICZ E, RAFALSKI A, LABUDA L, et al. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain-reaction amplification [J]. Genomics, 1994, 20(2): 176-183.

- [3] 刘瑾,舒妙安. 微卫星标记及其在贝类中的应用 [J]. 水利渔业,2007,27(2):17-19.
- [4] 杜长斌,孙效文,楼允东,等. 微卫星在水产动物种质资源研究方面的应用[J]. 水产学杂志,2000,13(1):68-73.
- [5] SOBOLEWSKA A, BEAUMONT A, HAKILTON A R. Dinucleotide microsatellites isolated from the European flat oyster, *Ostrea edulis* [J]. Molecular Ecology Notes,2001,1(1):79-80.
- [6] 付春鹏,傅洪拓,蒋速飞,等. 水产动物重要经济性状相关联的微卫星标记研究进展[J]. 中国农学通报,2010,26(4):314-317.
- [7] LI G, HUBERT S, BUCKLIN K, et al. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas* [J]. Molecular Ecology Notes,2003,3(2):228-232.
- [8] HUBERT S, HEDGECOCK D. Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas* [J]. Genetics, 2004, 168: 351-362.
- [9] 汪桂玲,袁一鸣,李家乐. 中国五大湖三角帆蚌群体遗传多样性及亲缘关系的 SSR 分析[J]. 水产学报,2007,31(2):152-158.
- [10] YU H, LI Q. Exploiting EST databases for the development and characterization of EST-SSRs in the Pacific oyster (*Crassostrea gigas*) [J]. Journal of Heredity,2008,99(2):208.
- [11] 于红. 牡蛎良种选育的遗传学基础研究[D]. 青岛:中国海洋大学,2010.
- [12] YU Z N, GUO X M. Genetic analysis of selected strains of eastern oyster (*Crassostrea virginica* Gmelin) using AFLP and microsatellite markers[J]. Marine Biotechnology,2004,6(6):575-586.
- [13] REECE K S, RIBERO W L, GAFFNEY P M, et al. Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): Confirmation of null alleles and non-Mendelian segregation ratios [J]. Journal of Heredity,2004,95(4):346-352.
- [14] HUBERT S, COGNARD E, D. HEDGECOCK. Centromere mapping in triploid families of the Pacific oyster *Crassostrea gigas* (Thunberg) [J]. Aquaculture,2009,288(3-4):172-183.
- [15] 龚小玲,李思发,蔡完其,等. 澳洲鳎微卫星分子标记的筛选与检测[J]. 中国水产科学,2009,16(1):133-138.
- [16] RAYMOND M, ROUSSET F. GENEPOP (Version 1.2)-Population genetics software for exact tests and ecumenicism [J]. Journal of Heredity,1995,86(3):248-249.
- [17] BOTSTEIN D, WHITE R L, SKOLNICK M, et al. Construction of a genetic-linkage map in man using restriction fragment length polymorphisms [J]. American Journal of Human Genetics,1980,32(3):314-331.
- [18] 李琪. 海洋贝类微卫星 DNA 标记的开发及其在遗传学研究中的应用 [J]. 中国水产科学,2006,13(3):502-509.
- [19] 姚鹏程,叶恭银. Hox 基因与昆虫体躯决定 [J]. 生命的化学,2003,23(2):105-107.
- [20] 张怡,张奎星,何鑫,等. 高血压定位区域犬尿氨酸酶基因多态性与高血压病相关 [J]. 中华心血管病杂志,2005,33(7):588-591.
- [21] 肖冰,李宁宁,吴永杰,等. 脑组织过表达犬尿氨酸酶的转基因小鼠对血压的影响 [J]. 中国分子心脏病学杂志,2009,9(3):144-149.

Development and analysis of 17 SSR from *Crassostrea gigas* fosmid database

LI Hui-juan¹, QI Hai-gang², LI Li², YAN Xi-wu^{1*}, ZHANG Guo-fan²

(1. College of Life Science & Technology, Dalian Ocean University, Dalian 116023, China;

2. Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China)

Abstract: Pacific oyster (*Crassostrea gigas*) is a widespread, delicious and nutritious species and is of high economic value. The codominant molecular microsatellite marker (also called simple sequence repeat, SSR) complies with Mendelian inheritance principles and is a powerful tool for analyzing population genetic structure, breeding and constructing genetic map because of its highly polymorphic and accurate genotyping. Sequences used in the study were obtained from end sequencing of fosmid library of the *C. gigas*. TRF software was used to exploit SSR loci with dinucleotide repeat, and 50 pairs of primers were acquired based on conserved SSR flanking sequence. We did preliminary screening in eight individuals and 17 of the above 50 primers were proved polymorphic. These primers were then screened and verified in a wild population in Qingdao, Shandong Province. Of these loci, 17 showed polymorphism. The average number of allele (N_a), number of effective alleles (N_e), observed heterozygosity (H_o), and expected heterozygosity (H_e) was 4, 2.82, 0.395 9, and 0.628 8, respectively. Among these loci, polymorphism information content (PIC) of 11 loci was more than 0.5, and 49 alleles in total were suitable for *C. gigas* population genetic structure analysis. The other 6 loci was moderately polymorphic and $0.25 < PIC < 0.5$. Through estimating the agreement of Hardy-Weinberg principle (χ^2 test) and Sequential Bonferroni calibration, we found that all loci except DEC_1 deviated equilibrium, suggesting that the growth and frequencies of genotypes of *C. gigas* changed. Blastp alignment analysis of 10 000 bp flanking sequences of SSR loci in nr database of NCBI showed that 13 linked gene sequences were development or metabolism related and can be used as molecular markers. The present results indicated that the development of SSR molecular marker based on the fosmid library can effectively remedy the predicament of lacking SSR markers of Pacific oyster to provide basis to protect and utilize the genetic diversity of this species.

Key words: *Crassostrea gigas*; simple sequence repeat (SSR); genome; genetic diversity

Corresponding author: YAN Xi-wu. E-mail: yanxiwu2002@163.com