

24 个泥蚶 EST-SSR 标记的开发与分析

史松富^{1,2}, 姚韩韩¹, 林志华¹, 周小龙^{1,2}, 齐晓艳³, 董迎辉¹

(1. 浙江万里学院 浙江省水产种质资源高效利用技术研究重点实验室, 浙江 宁波 315100; 2. 上海海洋大学 水产与生命学院, 上海 201306; 3. 宁波大学 海洋学院, 宁波 315211)

摘要: 利用 3 000 条泥蚶(*Tegillarca granosa*)单拷贝序列, 筛选到 132 条含有 SSR 候选位点序列, 成功设计引物 50 条。群体多态性检测结果显示, 24 个 SSR 位点为多态性, 各位点的等位基因介于 2~5 个, 平均等位基因数 3.5 个; 观测杂合度 H_o 、期望杂合度 H_e 、多态信息含量 PIC 范围分别为 0.000~0.591、0.045~0.727、0.043~0.696; 多态位点中有 11 个高度多态、4 个中度多态、9 个低度多态; 进行哈迪-温伯格平衡检测后, 用 Bonferroni correction 法进行校正, 结果有 17 个位点显著偏离平衡。

关键词: 泥蚶(*Tegillarca granosa*); EST-SSR; 单拷贝序列; 遗传多样性

中图分类号: S 917.4

文献标识码: A

文章编号: 1000-3096(2013)08-0042-05

微卫星标记又称简单重复序列(simple sequence repeat, SSR), 其核苷酸序列是以 2~6 个核苷酸为一个单位重复排列。微卫星因其在基因组中分布广泛、多态性丰富、共显性遗传、重复性好以及实验操作简单等特点, 已被广泛应用于遗传多样性分析、分子标记辅助育种、遗传连锁图构建、QTL 定位以及物种亲缘关系鉴定^[1-3]。SSR 标记的开发与分析在太平洋牡蛎(*Crassostrea gigas*)^[4-5]、泥蚶(*Tegillarca granosa*)^[6-7]、魁蚶(*Scapharca broughtonii*)^[8]、栉孔扇贝(*Chlamys farreri*)^[9-10]、文蛤(*Meretrix meretrix*)^[11]等海洋双壳贝类中已有报道。

泥蚶(*Tegillarca granosa*), 俗称血蚶、花蚶、粒蚶, 为我国四大海水养殖贝类之一。泥蚶肉质鲜美、营养丰富, 市场前景广阔。关于泥蚶分子标记的研究已有很多报道, 如李太武等^[12]利用 RAPD 标记技术分析了 5 个泥蚶群体的遗传多样性, 姚韩韩等^[13]利用 AFLP 标记技术对泥蚶 4 个快速生长家系的遗传变异进行了分析, 顾晓英等^[14]通过磁珠富集法筛选了 6 对多态性 SSR 引物; 董迎辉等^[6]利用泥蚶转录组文库开发了 34 个 EST-SSR 标记, Liu 等^[7]开发了 39 个泥蚶 EST-SSR 位点。本研究在泥蚶转录组文库的基础之上, 探讨利用单拷贝序列筛选 SSR 标记的可行性及优缺点, 为泥蚶的种质资源保护、遗传多样性分析、遗传连锁图谱的构建以及种群鉴定提供有力工具。

1 材料与方法

1.1 实验材料与 DNA 提取

实验样品采自浙江宁波奉化海区, 随机取样 30

粒, 活体解剖后取其闭壳肌, 保存于-20℃冰箱中备用。采用酚-氯仿法提取基因组 DNA, 用 1%的琼脂糖凝胶电泳检测 DNA 的质量, 紫外分光光度计检测 DNA 的纯度与浓度, 调节 DNA 浓度最终为 100 ng/μL。

1.2 候选 SSR 位点的获得与引物设计

用 SSR Hunter1.3 软件在单拷贝序列中筛选微卫星位点, 筛选标准为二碱基、三碱基、四碱基、五碱基重复单元重复至少依次为 6、4、3、3 次, 侧翼序列大于 150 bp。利用软件 Primer5.0 设计引物, 引物长度 18~24 bp, GC 含量 40%~60%, 引物由上海生工合成(表 1)。

1.3 泥蚶候选 SSR 位点检测与 PCR 扩增

随机取 6 个泥蚶 DNA 混成基因池, 对所有引物进行初步筛选。PCR 反应体系为 20 μL, 包括 100 ng DNA, 0.2 mmol/L dNTP, 0.5U Taq 酶, 10×PCR buffer, 上游引物与下游引物各 1 μmol/L, 1.5 mmol/L MgCl₂。PCR 反应条件: 94℃变性 5 min; 94℃变性 45 s, 退火 45 s, 72℃ 45 s, 共 35 循环; 72℃延伸 8 min; 4℃保存。每对引物所需的退火温度见表 1。用 8%的非变性聚丙烯酰胺凝胶电泳检测 PCR 产物,

收稿日期: 2013-06-06; 修回日期: 2013-06-16

基金项目: 国家现代贝类产业技术体系项目(CARS-48); 浙江省农业新品种选育重大科技专项(2012C12907-4); 宁波市科技创新团队项目(2011B82017); 宁波市科技局农业择优委托项目(2010C10011)

作者简介: 史松富(1988-), 男, 硕士研究生, 主要研究方向海水贝类养殖研究, 电话: 13486686591, E-mail: shisongfu123@163.com; 董迎辉, 通信作者, E-mail: dongyinghui118@126.com

表 1 泥蚶 24 个 EST-SSRs 标记的特征
Tab. 1 Characteristics of 24 EST-SSRs for *Tegillarca granosa*

位点	引物序列(5'-3')	重复模式	退火温度 (°C)	产物大小 (bp)	等位基因数 Na	观测杂合度 Ho	期望杂合度 He	多态信息含量 P/I	P 值	GeneBank 登录号
Tg01	F:TAAACAGTGGCATCTATCTTGGCT R:ACTCTGTC TTGTTGGCTTCT	(CAA) ₄ (ACG) ₄ (ACA) ₅	64.5	251~280	2	0.045	0.045	0.043	0.0000	KF171286
Tg02	F:02CAAAGTTGACCGCTGCTT R:GCTGTGACTATGATAGTTGTTCTC	(ATA) ₄	64.5	140~158	5	0.136	0.566	0.517	0.0000	KF171287
Tg03	F:TGACCTTAAATAGAGTTGTCC R:CACTGATGAGGAATGATTTACCG	(TA) ₆	64.5	118~129	5	0.591	0.709	0.648	0.0000	KF171288
Tg04	F:GAGTAGCACATTCATTTCTG R:GTGTGTTTCAATTTAAGTCACTG	(ATA) ₄	64.5	152~155	3	0.000	0.664	0.574	0.0000	KF171289
Tg05	F:GACGACACGGCTTTACACAT R:ATAACAAGGACAAAGCACAT	(TAT) ₄	68.7	197~203	4	0.273	0.665	0.595	0.0000	KF171290
Tg06	F:TCTTCAAACTACAAAGTTTACG R:CTCTTCCACATACAGAACCAT	(TAA) ₆	61.1	190~200	3	0.000	0.260	0.237	0.0000	KF171291
Tg07	F:TGAATGATTTGTAAACTGGG R:TCGGTTGAACAGAGGGAAT	(GATT) ₃	61.1	278~280	2	0.000	0.089	0.083	0.0000	KF171292
Tg08	F:CTATCAGCCACAGCCTTT R:CTACACAAAATCCAAAGATC	(GAT) ₅	61.1	208~215	5	0.045	0.652	0.586	0.0000	KF171293
Tg09	F:ATTGTTAAGAAGTTGTGTAG R:TAACATACCTGTCAAAAC	(ATT) ₆	68.7	300~306	4	0.091	0.452	0.411	0.0000	KF171294
Tg10	F:CTAAATCTGTTTCTGGTTGA R:AGGAGCGTTACAAAAGATG	(ATAA) ₃	68.7	155~160	2	0.045	0.130	0.119	0.0725	KF171295
Tg11	F:AGTCATTTGTTATAGGTGTC R:GACATAGTTAGAAGGATG	(ATA) ₄	53.4	147~151	2	0.000	0.089	0.083	0.0222	KF171296
Tg12	F:TCTGTGTGTGTCGTGGT R:GTAGCAATACAGCATATCAGTAC	(TTG) ₄	53.4	159~190	3	0.143	0.261	0.240	0.0161	KF171297
Tg13	F:TGTCCCTTTTGAACGC R:CATCACACAATACACTGGTTAA	(TTA) ₄	61.1	99~105	4	0.136	0.449	0.405	0.0004	KF171298
Tg14	F:CCACATTTGTGATTTCTTG R:AGCGTGACATTTGCGTA	(AAC) ₄	48.5	190~210	4	0.000	0.490	0.427	0.0000	KF171299
Tg15	F:CAATAAGTTTCTCAATGC R:CAAAACAGTATCAGCAAGAC	(GA) ₆	53.4	235~255	4	0.409	0.621	0.558	0.0015	KF171300
Tg16	F:AGCAGCGGCCACATAACAT R:TGACCAACGACACACACAGAATA	(TTG) ₄	50.8	175~200	5	0.318	0.727	0.666	0.0000	KF171301
Tg17	F:TACAATGAACATCGGCAC R:GGCTAGTGTTCGAAGAGG	(GCT) ₄	61.1	140~146	2	0.000	0.089	0.083	0.0234	KF171302
Tg18	F:TAAGTGAATGGAACAAC R:TGTGTGTTATAGATTTGTGG	(ACC) ₄	58.7	230~256	2	0.000	0.089	0.083	0.0224	KF171303
Tg19	F:GGTCAGACTTGACTTTAATACA R:TTCACTCCAAAACCTCGG	(AT) ₇	58.7	220~226	5	0.000	0.757	0.696	0.0000	KF171304
Tg20	F:CGAAGGACCTACAAATGAA R:GTTGCAATTTTACCAATGT	(TGA) ₆	58.7	105~110	4	0.136	0.479	0.432	0.0000	KF171305
Tg21	F:GCCATAATCTGCCTGTAA R:CTGTGAACCTTCACCTTGTA	(ATT) ₄	58.7	180~185	2	0.000	0.089	0.083	0.0236	KF171306
Tg22	F:CTGGGCCACAACTTAA R:CTGCTGGTTTCCAAATGCT	(TTC) ₄	46.6	200~215	4	0.091	0.678	0.599	0.0000	KF171307
Tg23	F:ACAAGATTACACAAGAGCCTG R:CACGCAAGGATTCGTCAAT	(AC) ₆	61.1	194~200	5	0.333	0.694	0.633	0.0003	KF171308
Tg24	F:GTACAAACACAAACCACTG R:ACTCAGTTGAATCACCTT	(CCA) ₅	56.1	119~139	3	0.682	0.678	0.588	0.4544	KF171309
平均值	—	—	—	—	3.5	0.145	0.434	0.391	—	—

电泳缓冲液为 $1\times$ TBE, 电压为 180 V, 电泳时间为 3~4 h, EB 染色, 用 Biorad 凝胶成像系统检测结果。

1.4 泥蚶 SSR 位点分析

利用筛选的引物对泥蚶 30 个个体进行微卫星标记的多态性检测, 用 8% 的非变性聚丙烯酰胺凝胶电泳检测 PCR 产物。用 CERVUS 3.0 软件处理结果, 获得重要的遗传学参数: 等位基因数(N_a)、期望杂合度(H_e)、观测杂合度(H_o)和多态信息含量(PIC)。用在线软件 GENEPOP 进行哈迪-温伯格平衡(Hardy-Weinberg equilibrium, HWE)检测, 利用 Bonferroni correction 法对结果进行校正。

2 结果

2.1 泥蚶 SSR 位点的筛选

利用 SSR Hunter1.3 软件搜索 3000 条单拷贝序列 (singleton), 获得 132 条含有 SSR 位点的序列, 成功设计引物有 50 条, 成功率为 37.88%; 利用 Primer5.0 在各 SSR 位点两侧保守区设计引物, 41 对引物获得稳定清晰的条带; 在泥蚶群体中对 41 对引物进行多态性检测, 24 对引物的扩增产物表现为多态性。

在 24 个 EST-SSR 位点中, 三碱基重复类型的最多为 18 个, 占总 SSR 的 75%; 其次二碱基重复类型的 SSR 有 4 个, 占总数的 16.67%; 四碱基与五碱基重复类型的 SSR 最少均为 1 个, 均占总多态位点的 4.17%(图 1)。

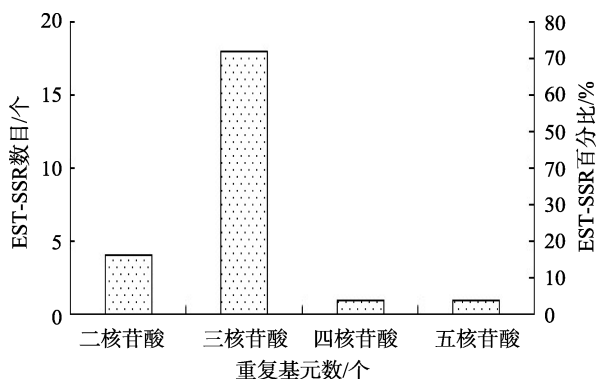


图 1 泥蚶 24 个 EST-SSR 位点的核苷酸重复碱基的分布特征

Fig. 1 Distribution of 24 EST-SSR motifs types in *Tegillarca granosa*

2.2 24 个 SSR 位点对泥蚶群体的多态性评价

24 个多态性位点在泥蚶 30 个个体中共获得 84 个等位基因, 多态位点的等位基因数为 2~5 个, 平均

等位基因数为 3.5 个, 其中 Tg02、Tg03、Tg08、Tg16、Tg19、Tg23 六个位点的等位基因数为 5 个。观测杂合度、期望杂合度分别为 0.000~0.591、0.045~0.727(表 1)。 PIC 值是衡量群体变异程度的重要参数^[15], 24 个多态位点的 PIC 值介于 0.043~0.696, 平均值为 0.391, 其中有 11 个高度多态位点($PIC > 0.5$)、4 个中度多态位点($0.25 < PIC < 0.5$)、9 个低度多态位点($PIC < 0.25$)。各多态位点经哈迪-温伯格平衡检测, 并用 Bonferroni correction 法进行校正, 除 Tg10、Tg11、Tg12、Tg17、Tg18、Tg21、Tg24 之外, 其余位点均显著偏离 Hardy-Weinberg 平衡。

3 讨论

目前 SSR 开发主要是基于构建的基因组文库和 EST 文库, EST-SSR 的开发具有操作简单、成本低、高效率等特点, 越来越受到研究者的青睐。在海洋贝类中, 利用 EST 文库筛选 SSR 的研究已较深入, 如 Wang 等^[16]利用从文蛤 EST 文库筛选出 2970 个候选 SSR 位点; Hou 等^[17]通过虾夷扇贝转录组文库得到 2700 个 SSR 候选位点。本研究在泥蚶转录组的单拷贝序列 (singleton) 中筛选 SSR 位点, 并将相关信息与从重叠群 (contig) 中筛选 SSR 位点的研究相比较, 分析 singleton 与 contig 筛选 SSR 位点的差异。

本研究利用泥蚶 singleton 筛选的 SSR 引物设计成功率为 37.88%, 董迎辉等^[6]利用泥蚶 contig 筛选 SSR 位点的引物设计成功率 73.97%, 引物设计成功率较低的原因有: (1) singleton 的序列长度较短, 泥蚶 singleton 的平均长度为 285.46 bp^[18], 用于设计引物的区域相对较小, 而 contig 的平均长度为 964.2 bp。(2) singleton 中的一小部分被认为是在测序过程中被其他生物的序列污染或者是人工序列的产生^[17]。

singleton 来源的 SSR 同 contig 来源的 SSR 一样能反映生物群体的重要遗传学信息, 单拷贝序列的充分利用有利于群体的遗传多样性分析、种质资源的保护等, 如董迎辉等^[6]利用 contig 筛选 34 个 EST-SSR 的研究中平均等位基因数为 3.59、平均多态信息含量为 0.396、平均期望杂合度为 0.447, Liu 等^[7]关于 39 个 EST-SSR 的研究中平均等位基因数、平均多态信息含量、平均期望杂合度分别为 3.9、0.402、0.449, 本实验利用 singleton 筛选的 24 个 SSR 多态位点的平均等位基因数为 3.5、平均多态信息含量 0.391、平均期望杂合度 0.434。目前 EST-SSR 筛

选主要来自 contig, 而关于 singleton 的开发应用以及
与 contig 的比较研究较少。本研究证实利用 singleton
开发 SSR 标记是可行的, singleton 在转录组文库中所
占的比例较大且包含丰富的遗传信息, 特别在表达
丰度较低基因的开发方面较 contig 具有更重要的应
用价值^[17]。

参考文献:

- [1] 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.
- [2] Morgante M, Hanafey M, Powell W. Microsatellites are
preferentially associated with nonrepetitive DNA in
plant genomes [J]. Nature Genetics, 2002, 30(2):
194-200.
- [3] 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.
- [4] Qi L, Wang Q Z, Qi M J, et al. Development,
characterization, and inheritance of 113 novel EST-SSR
markers in the Pacific oyster (*Crassostrea gigas*)[J].
Genes & Genomics, 2011, 33:313-316.
- [5] 李慧娟, 亓海刚, 李莉, 等. 长牡蛎 17 个 fosmid-
SSR 标记的开发与分析[J]. 水产学报, 2011,
35(10):1464-1467.
- [6] 董迎辉, 吴国星, 姚韩韩, 等. 泥蚶 34 个 EST-SSR 标
记的开发及在格粗饰蚶中的通用性检测[J]. 水产学
报, 2013, 37(1):70-77.
- [7] Liu B, Teng S S, Shao Y Q, et al. Isolation and
characterization of 39 novel polymorphic EST-SSR loci
for the blood clam, *Tegillarca granosa*[J]. Conservation
Genet Resour, 2012, 4:375-378.
- [8] 吴彪, 梁超, 杨爱国, 等. 基于 SSR 标记的魁蚶
(*Scapharca broughtonii*)不同群体遗传结构的分析[J].
海洋与湖沼, 2012, 43(4):863-869.
- [9] 李红蕾, 宋林生, 王玲玲. 栉孔扇贝 EST 中微卫星标
记的筛选[J]. 高技术通讯, 2003: 72-75.
- [10] Zhan A B, Bao Z M, Hu X L, et al. Characterization of
95 novel microsatellite marker for Zhikong scallop
Chlamys farreri using FIASCO-colony hybridization
and EST database mining[J]. Fisheries Science, 2008,
74(3): 516-526.
- [11] 朱东丽, 董迎辉, 林志华, 等. 利用微卫星标记对文
蛤 9 个壳色花纹品系的遗传分析[J]. 水产学报, 2011,
26(2): 202-209.
- [12] 李太武, 李成华, 宋林生, 等. 5 个泥蚶群体遗传多样
性的 RAPD 分析[J]. 生物多样性, 2003, 11(2):
118-124.
- [13] 姚韩韩, 董迎辉, 林志华, 等. 泥蚶 4 个快速生长家
系的遗传变异分析[J]. 水产学报, 2011, 35(3):
350-357.
- [14] 顾晓英, 曾庆国, 尤仲杰, 等. 泥蚶(*Tegillarca
granosa*)6 个微卫星引物的分离和鉴定[J]. 海洋与湖
沼, 2008, 39(6):661-664.
- [15] 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.
- [16] Wang H X, Huan P, Lu X, et al. Mining of EST-SSR
marker in clam *Meretrix meretrix* larvae from 454
shotgun transcriptome[J]. Conservation Genetics
Resource, 2011, 3(4):655-658.
- [17] Hou R, Bao Z M, Wang S, et al. Transcriptome
sequencing and de nove analysis for Yesso scallop
(*Patinopecten yessoensis*) using 454 GS FLX[J]. Plos
One, 2011, 6(6):e21560.
- [18] 董迎辉. 泥蚶高通量转录组分析及生长相关基因的
克隆与表达研究[D]. 青岛: 中国海洋大学, 2012:
1-144.

Characterization and analysis of 24 polymorphic EST-SSR loci in *Tegillarca granosa*

SHI Song-fu^{1,2}, YAO Han-han¹, LIN Zhi-hua¹, ZHOU Xiao-long^{1,2}, QI Xiao-yan³, DONG Ying-hui¹

(1. Zhejiang Key Laboratory of Aquatic Germplasm Resources, Zhejiang Wanli University, Ningbo 315100, China; 2. College of Fisheries and Life Sciences, Shanghai Ocean University, Shanghai 201306, China; 3. College of Marine Sciences, Ningbo University, Ningbo 315211, China)

Received: Jun., 6, 2013

Key words: *Tegillarca granosa*; EST-SSR loci; single copy sequence; polymorphism

Abstract: 50 pairs of primers were designed successfully by using 132 candidates of SSR loci sequences which were identified from 3000 single copy sequences of *Tegillarca granosa*. It was revealed from the investigation of polymorphism that 24 SSR loci were polymorphic. The number of alleles of the 24 EST-SSR loci varied from 2 to 6, and the mean value is about 3.5. The range of the observed heterozygosity, the expected heterozygosity and the polymorphic information content was 0.000 to 0.591, 0.045 to 0.727 and 0.043 to 0.696, respectively. There are 9 high polymorphic SSR locus, 4 moderate polymorphic SSR locus, and 11 high polymorphism SSR locus. 17 SSR locus are significant deviations from Hardy-Weinberg equilibrium after Bonferroni correction.

(本文编辑: 康亦兼)