

泥蚶 34 个 EST-SSR 标记的开发及 在格粗饰蚶中的通用性检测

董迎辉¹, 吴国星¹, 姚韩韩¹, 边平江², 赵德旻¹, 林志华^{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. 浙江省海洋水产养殖研究所, 浙江 温州 325005)

摘要: 利用泥蚶转录组高通量测序、拼接获得的大量 EST 序列开发 SSR 标记, 在 1 123 条 EST 序列里筛选到 73 条含有 SSR 位点的 EST 序列, 其中 54 个位点适合设计引物, 在位点两侧设计引物并进行 PCR 扩增。结果显示, 46 对引物获得稳定扩增的位点, 引物在泥蚶奉化群体的多态性检测中发现, 有 34 对引物表现出多态性, 共扩增出 122 个等位基因, 各位点的等位基因数 N_a 为 2~7 个, 平均每个位点产生 3.59 个等位基因, 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 0.000~0.600、0.078~0.771、0.106~0.718; Hardy-Weinberg 平衡检测显示, 13 个位点偏离了平衡状态; 用 Nr 和 Swiss-Prot 蛋白质数据库对含有多态性 SSR 的 EST 进行了基因注释, 25 个 SSR 位点来自注释基因序列。将 34 对泥蚶多态性 SSR 引物在格粗饰蚶中进行了通用性检测, 结果有 11 对成功扩增, 8 对表现为多态, 通用率为 23.53%, 这些通用引物可用于两种蚶的遗传多样性评价、系统进化分析、比较作图和基因发掘等研究。

关键词: 泥蚶; EST-SSR; 多态性; 通用性; 格粗饰蚶

中图分类号: S 917.4

文献标志码: A

微卫星标记 (simple sequence repeat, SSR), 具有多态性丰富、共显性遗传、重复性好、操作分析简单等优点, 已被广泛用于种群遗传多样性评价、遗传连锁图谱构建、分子标记辅助育种等领域^[1]。近年来, 随着高通量测序产生的海量 EST (expressed sequence tag), 基于 EST 序列大量开发 SSR 标记的技术正在得到广泛应用。EST-SSR 来源于基因表达序列, 与基因组来源的 SSR 相比, 具有可能与基因功能相关、种间通用性好、开发成本低且周期短等优点^[2]。目前, 在海洋双壳贝类中已见海湾扇贝 (*Argopecten irradians*)^[3]、栉孔扇贝 (*Chlamys farreri*)^[4-5]、太平洋牡蛎 (*Crassostrea gigas*)^[6-7]、文蛤 (*Meretrix meretrix*)^[8-10]、马氏珠母贝 (*Pinctada martensii*)^[11]、缢蛏 (*Sinonovacula constricta*)^[12] 等开展过 EST-SSR 标记的开发和分析。

泥蚶 (*Tegillarca granosa*) 是江、浙、闽沿海居

民喜食的海鲜佳品, 也是我国四大养殖贝类之一。然而, 目前有关泥蚶分子标记方面的研究报道不多, 主要采用 RAPD、AFLP 技术对泥蚶不同地理群体^[13-14]、快速生长家系^[15] 的遗传结构和遗传变异进行了分析, 而泥蚶 SSR 标记的研究仅见通过磁珠富集法从基因组中开发了 6 对多态性 SSR 引物^[16-17]。本研究在 454 高通量转录组测序的基础上, 利用测序并拼接组装获得的大量 EST 序列规模化开发微卫星标记, 并在具有较大保护价值和开发潜力的格粗饰蚶 (*Anadara craticulata*) 中进行通用性检测, 可望为泥蚶和格粗饰蚶种质资源保护、遗传多样性评价、系统进化分析以及比较基因组学研究等提供有力工具。

1 材料与方法

1.1 样品采集及 DNA 提取

实验所用泥蚶和格粗饰蚶均为野生群体, 泥

收稿日期: 2012-08-23 修回日期: 2012-10-11

资助项目: 国家现代贝类产业技术体系项目 (CARS-48); 国家“八六三”高技术研究发展计划 (2012AA10A410); 浙江省海洋生物技术产业科技创新团队项目 (2012R10029); 宁波市农业择优委托项目 (2010C10011)

通信作者: 林志华, E-mail: zhihua9988@126.com

蚶采自浙江宁波奉化海区(29°51'N,121°46'E),格粗饰蚶取自广东湛江东海岛海区(21°53'N,110°45'E),活体运回实验室,各取 30 粒进行解剖,剪取闭壳肌,保存于 -80 °C 冰箱中备用。采用常规酚/氯仿/异戊醇法提取基因组 DNA,用 1% 琼脂糖凝胶电泳检测 DNA 的完整性,超微量紫外分光光度计(NanoVue)检测 DNA 浓度和纯度。调整 DNA 浓度至 100 ng/ μ L,用于后续的 PCR 扩增。

1.2 SSR 序列的获得及引物设计

根据本实验室泥蚶 454 转录组测序(登录号 SRA028346)、拼接后的 EST 序列,利用 SSR Hunter1.3 软件进行微卫星序列搜索,搜索标准是:二碱基至少重复 6 次,三碱基至少重复 4 次,四碱基至少重复 3 次,五碱基至少重复 3 次,六碱基至少重复 3 次。将含有微卫星序列的 EST 片段进行序列比对,去除重复序列,用 Primer 5.0 软件设计引物,设计的引物由上海生工合成(表 1)。

表 1 泥蚶 34 个多态性 EST-SSR 标记的特征
Tab.1 Characteristics of 34 polymorphic EST-SSRs for the blood clam *T. granosa*

位点 locus	引物序列(5'-3') primer sequence	重复 模式 repetitive mode	退火 温度/ $^{\circ}$ C annealing temperature	产物 大小/bp product size	等位 基因数 N_a	观测杂合度/ 期望杂合度 H_o/H_e	多态信息 含量 PIC	P 值 P value	功能注释* / E 值 functional annotations/ E value
Tg01	F:GTTTTGGTGTAGTTGGTCGC R:TACCCAACGACTCCAATGAC	(GTT) ₅	64.5	175 ~ 195	7	0.440/0.449	0.426	0.591 4	葡萄糖脱氢酶 [受体]/8.00E-81
Tg02	F:GTGCCAGTGAGTATTCTGATT R:ACTGGAATCCAAATCCACAA	(TTG) ₅	64.5	175 ~ 177	2	0.080/0.078	0.074	1.000 0	葡萄糖脱氢酶 [受体]/8.00E-81
Tg03	F:CAGGACCAGGGTTCATTTTTG R:GGGGCAAACCTCAAATCCAC	(TTG) ₄	64.5	185 ~ 200	5	0.600/0.728	0.664	0.029 6	葡萄糖脱氢酶 [受体]/8.00E-81
Tg04	F:GCAGGACCAGGGTTCATTT R:GCACAAAGTAACTGGGGCAAAC	(TTG) ₅	64.5	196 ~ 205	4	0.560/0.675	0.595	0.457 2	葡萄糖脱氢酶 [受体]/8.00E-81
Tg05	F:CGACAACCTAACTTCACTTTCAC R:CCGTTGACGATGGTGCTG	(CAA) ₅	64.5	274 ~ 278	3	0.200/0.254	0.232	0.016 2	无
Tg06	F:CTCCAGCCGAAAATCAAACACAA R:TCCCACTCAACCCATCATTCCAAA	(AAC) ₆	61.1	185 ~ 192	3	0.000/0.222	0.205	0.000 0	无
Tg07	F:ACCGAAGTCTGTTGATTGC R:TGGTCTGAAAATGTTTAGCC	(AAAT) ₃	61.1	143 ~ 155	3	0.200/0.318	0.289	0.000 0	无
Tg08	F:GGCAAATCCATCAAACCTTCTC R:GCCCTGTTCTGTTTGTGGTC	(AATAA) ₃	63.1	212 ~ 218	3	0.174/0.305	0.273	0.063 1	室管膜蛋白/ 1.00E-07
Tg09	F:GAAATCCGCCGACAAAGT R:TTGTATTTTCGACGGCATCA	(GAC) ₄	63.1	236 ~ 240	3	0.040/0.254	0.254	0.000 1	铜胺氧化酶 N-端结 构域蛋白/2.00E-07
Tg10	F:GAAGGGGATGATGTATTGGAAA R:TTGTTTTACCTGGTTTGCTGC	(GAT) ₅	58.7	190 ~ 195	3	0.320/0.280	0.280	1.000 0	铜胺氧化酶 N-端结 构域蛋白/2.00E-07
Tg11	F:GAAGGGGATGATGTATTGGAAAAC R:TTTACCTGGTTTGCTGCTTGAA	(ATG) ₄	58.7	186 ~ 192	3	0.360/0.308	0.308	1.000 0	铜胺氧化酶 N-端结 构域蛋白/2.00E-07
Tg12	F:GACTTTTATCGCAGTTTC R:CAITGATGGTGTCATTTTG	(TTTA) ₅	56.1	176 ~ 178	2	0.240/0.216	0.189	1.000 0	类赖氨酸组氨酸转 运蛋白 2/2.00E-18
Tg13	F:GAGCAAGCAGTTCAAGAAGATA R:TCTCAGCTATCTCCATGTTTA	(ATA) ₅	61.1	276 ~ 280	3	0.280/0.469	0.410	0.002 5	无
Tg14	F:CTTGACCCTATGACAGA R:GTTACCGTAGAGATAACACA	(TTTCT) ₃	58.7	132 ~ 138	4	0.240/0.404	0.365	0.005 9	无
Tg15	F:CAGACTTCAAGGAGCAAACAG R:TTGCTTTGGTGGAACCTGC	(CAA) ₇	58.7	132 ~ 142	2	0.640/0.509	0.375	0.248 0	无
Tg16	F:GACTTTGAGTATTCTCTTAGC R:CGGGGAGTGCTAAAAA	(TTG) ₆	58.7	197 ~ 205	4	0.520/0.456	0.411	1.000 0	A 型病毒蛋白重复 包含蛋白/2.00E-13
Tg17	F:AGAATCATAATCACTCTCCTC R:AAAACAGACAAGCCAAGA	(TCA) ₄	53.4	170 ~ 180	3	0.240/0.222	0.205	1.000 0	转录调控因子 ATRX/3.00E-49

续表 1

位点 locus	引物序列(5'-3') primer sequence	重复 模式 repetitive mode	退火 温度/°C annealing temperature	产物 大小/bp product size	等位 基因数 N_a	观测杂合度/ 期望杂合度 H_o/H_e	多态信息 含量 PIC	P 值 P value	功能注释* / E 值 functional annotations/ E value
Tg18	F:CTGTGATTGGAGACACCATAAACT R:CTGAGGACGAAGTAGAAAGAAGAC	(GGT) ₅	46.6	225 ~ 238	5	0.200/0.542	0.502	0.000 0	剪接因子 3A 亚基 1/e-119
Tg19	F:AGGAGCCCAGAGAAATGA R:TTTTGGACTCTTTAGCAGGT	(AGA) ₄	45.3	220 ~ 238	5	0.240/0.519	0.482	0.000 2	无
Tg20	F:CATCAGAATCTTCCTGCTTT R:GGAGGTTACCAACTACAA	(TCA) ₅	45.3	156 ~ 165	4	0.240/0.291	0.273	0.089 9	DNA 拓扑异构酶 1/4.00E-31
Tg21	F:TGAAGGGGAACCTGTCAAAA R:CTACCATTTCCTTGATTTGTGC	(GAA) ₅	50.8	100 ~ 110	3	0.440/0.373	0.334	1.000 0	过渡型内质网 ATP 酶/0.00E-0
Tg22	F:TGGCAGTATTTTTCTGTGTGTGG R:AAGTAGGTCTGTGGGTGATTTGA	(TCC) ₆	64.5	194 ~ 200	4	0.280/0.409	0.376	0.000 0	Rho GTP 酶激活蛋白 6/7.00E-08
Tg23	F:GGGAAAGTTATGGAAGTG R:ACAGGCATTGTTGATATG	(ATTTT) ₃	50.8	220 ~ 224	2	0.040/0.115	0.106	0.062 0	放射辅头样蛋白 4 同源 A/e-155
Tg24	F:GCCTGAACCACCAAATGT R:TTTTCTGGAGAGAAGTCA	(ACAA) ₃	58.7	145 ~ 158	5	0.400/0.706	0.643	0.0016	预测:似 HrPalE/ 5.00E-25
Tg25	F:TTGGCTTGATTTCCATAGG R:TGAGGCACTGAAAAACGAT	(ATG) ₅	56.1	160 ~ 170	4	0.200/0.555	0.463	0.000 0	微管蛋白 α -1 链/ 0.00E-0
Tg26	F:TATCTGAGACGGCATTG R:GCTGCCAATACAGATGAA	(ATC) ₅	56.1	148 ~ 170	6	0.600/0.771	0.718	0.187 6	Claspin 蛋白/ 6.00E-12
Tg27	F:TTCAGGGGGACAACCTCT R:TCTACCAGCGAACATTTG	(CT) ₇	53.4	100 ~ 105	3	0.391/0.405	0.333	0.303 0	CD151 抗原/ 1.00E-14
Tg28	F:CATTATCCTCATCTCAAGTGTC R:CTTTCTGTTTACCTCTCCTA	(TCA) ₇	61.1	180 ~ 186	3	0.240/0.549	0.430	0.000 3	AF-9 蛋白/ 4.00E-43
Tg29	F:CAAATAACCTTGCCAGTGA R:TCTCTCATCACATTCATAACAC	(TA) ₇	61.1	194 ~ 198	3	0.200/0.581	0.475	0.000 0	同源框蛋白 SIX2/1.00E-79
Tg30	F:GCCAGTGAATCCCAACAA R:CCAAACACAAACATCAATATAGC	(TG) ₆	61.1	225 ~ 230	3	0.280/0.603	0.508	0.001 0	同源框蛋白 SIX2/1.00E-79
Tg31	F:TACCATTTCAGTACCTCA R:GTCAAACCTGTAGCCAACA	(TA) ₇	56.1	142 ~ 152	4	0.560/0.663	0.602	0.106 2	无
Tg32	F:AATGTCAGTCGGTGTAGAG R:CACTTTTGGACAGTGTAT	(AAT) ₆	58.7	178 ~ 190	4	0.440/0.690	0.613	0.001 3	凝溶胶蛋白样蛋白 1/1.00E-70
Tg33	F:CCAAACATTGAACTGCTCCTGCT R:CAAGAAAAAGTGACACCGAAACCT	(TCA) ₄	46.6	280 ~ 286	3	0.440/0.613	0.533	0.000 0	无
Tg34	F:TCTTTCCACAGCGCAAT R:ATCGGCAAAGTGATGGTATTAG	(AT) ₈	64.5	225 ~ 233	4	0.240/0.673	0.603	0.000 0	蛋白磷酸酶 1 调节 亚基 16A/7.00E-08
平均值 mean		-	-	-	3.59	0.311/0.447	0.396		

注: * 利用 Nr 和 Swiss-Prot 蛋白质数据库对含有多态性 SSR 的 EST 进行了基因注释(E 值为 $1e-5$)。

Notes: * : The SSR-containing ESTS were annotated using BlastX against Nr or Swiss-Prot database with an E -value cutoff of $1e-5$.

1.3 微卫星引物的筛选与优化

先将 8 个泥蚶的 DNA 混合成基因池,进行引物的初步筛选与条件优化。PCR 反应体系为 20 μ L, 包含 0.5 U *Taq* 酶 (TaKaRa), 1 \times PCR buffer, 1.5 mmol/L $MgCl_2$, 0.2 mmol/L dNTP, 1 μ mol/L 上、下游引物和 100 ng DNA 模板。PCR

反应程序为:94 $^{\circ}C$ 预变性 5 min;94 $^{\circ}C$ 变性 45 s, T_m 退火 45 s(根据各引物情况而定),72 $^{\circ}C$ 延伸 45 s,进行 35 个循环,最后 72 $^{\circ}C$ 延伸 7 min,4 $^{\circ}C$ 保存。PCR 产物用 8% 非变性聚丙烯酰胺胶进行电泳检测,电压 180 V,电泳 3~4 h,EB 染色,Bio-rad 凝胶成像系统拍照记录,根据电泳图谱结果确

定引物的最佳退火温度。

1.4 泥蚶奉化群体 SSR 标记的多态性分析

根据引物筛选结果,在奉化群体 30 个个体中进行 SSR 标记的多态性验证,扩增产物用 8% 非变性聚丙烯酰胺电泳检测。数据处理采用 CERVUS 3.0 软件计算等位基因数(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 GENEPOP 在线软件(<http://wbiomed.curtin.edu.au/genepop/>)计算 Hardy-Weinberg 平衡,用 Bonferroni correction 对结果进行校正。用 Nr 和 Swiss-Prot 蛋白质数据库对含有多态性 SSR 的 EST 进行基因注释。

1.5 泥蚶 SSR 位点在格粗饰蚶中的通用性检测及多态性评价

将在泥蚶中扩增获得的多态性 SSR 引物,在格粗饰蚶群体中进行通用性检测。先进行退火温度的筛选,挑选目标条带清晰的温度作为群体扩增的最佳温度,然后在格粗饰蚶 30 个个体中进行扩增,统计这些标记在格粗饰蚶中的通用率,并分析这些位点在泥蚶和格粗饰蚶中的多态性差异。

2 结果

2.1 泥蚶微卫星位点的筛选及对奉化群体的多态性评价

利用 SSR Hunter1.3 软件在泥蚶 1 123 条拼接后的大片段 EST 序列里搜索 SSR 位点,结果检测到 73 条含有二至六核苷酸重复的 EST 序列,其中 54 个位点适合设计引物,在各位点两侧保守区域设计引物,并进行 PCR 扩增和最佳退火温度筛选。结果显示,有 46 对引物获得了稳定扩增的清晰位点,扩增成功率为 85.19%。46 个位点在泥蚶奉化群体中进行了多态性检测,其中 34 个位点表现出多态性($N_a \geq 2$),多态引物比例达到 73.91%。多态位点的最佳扩增条件见表 1。

在检测出的 34 个 EST-SSR 位点中,包括二至五核苷酸重复类型,其中三核苷酸重复类型的 SSR 数量最多(23 个),占总 EST-SSR 的 67.65%,TTG 重复类型有 3 个,CAA、TCA 重复类型各 2 个;其次二核苷酸重复类型的 SSR 数有 5 个,占总 EST-SSR 的 14.71%;而三、四核苷酸重复类型的 SSR 数都是 3 个,占总 EST-SSR 的 8.82% (图 1)。

34 个多态位点在泥蚶奉化群体 30 个个体中进行了扩增,共获得 122 个等位基因,各位点的等位基因数 N_a 为 2~7 个,平均每个位点产生 3.59

个等位基因,其中 Tg01 位点的等位基因数最多(7 个)。泥蚶奉化群体的各遗传参数见表 1,观测杂合度 H_o 、期望杂合度 H_e 、多态信息含量 PIC 分别为 0.000~0.600、0.078~0.771、0.106~0.718;Tg03、Tg04、Tg15、Tg26、Tg31 等位点的杂合度较高,适合进行群体遗传多样性分析;从 PIC 值来看,在 34 个多态性位点中有 10 个高度多态位点($PIC > 0.5$)、18 个中度多态位点($0.25 < PIC < 0.5$),可作为遗传变异研究的工具。利用在线 Genepop 软件对各位点进行哈温平衡检验,经过 Bonferroni correction 校正,仍有 13 个位点显著偏离 Hardy-Weinberg 平衡($P < 0.05$)。

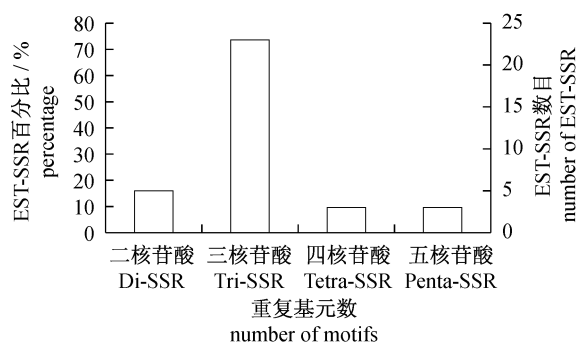


图 1 泥蚶 34 个 EST-SSR 位点核苷酸碱基的分布特征

Fig. 1 Distribution of 34 EST-SSR motifs types in *T. granosa*

2.2 泥蚶多态性 EST-SSR 标记的功能注释

利用 NCBI 上的 Nr 和 Swiss-Prot 蛋白质数据库,对含有 34 个多态性 SSR 的 28 条 EST 序列进行了基因功能注释(E 值为 $1e-5$),结果有 25 个 SSR 位点来自注释基因序列,与 19 个已知功能基因相关(表 1),其中葡萄糖脱氢酶[受体]基因中有 4 个 SSR 位点,铜胺氧化酶 N-端结构域蛋白基因中有 3 个位点,同源框蛋白 SIX2 基因中有 2 个位点,其余基因有 1 个位点。这些 SSR 位点可能与基因的表达调控相关,将对相应基因的准确定位和功能分析有很大帮助。

2.3 泥蚶多态性 EST-SSR 在格粗饰蚶中的通用性检测及群体多态性评价

用 34 对泥蚶多态性 SSR 引物在格粗饰蚶中进行了通用性验证,结果发现 11 对引物能成功扩增,其中 8 对表现为多态,通用率为 23.53%,两个物种的遗传参数比较见表 2。结果表明,在通用的 8 对多态性 EST-SSR 引物中,三碱基重复和二碱基重复类型的通用性较高,5 对为三碱基重复类型,2 对为二碱基重复类型。

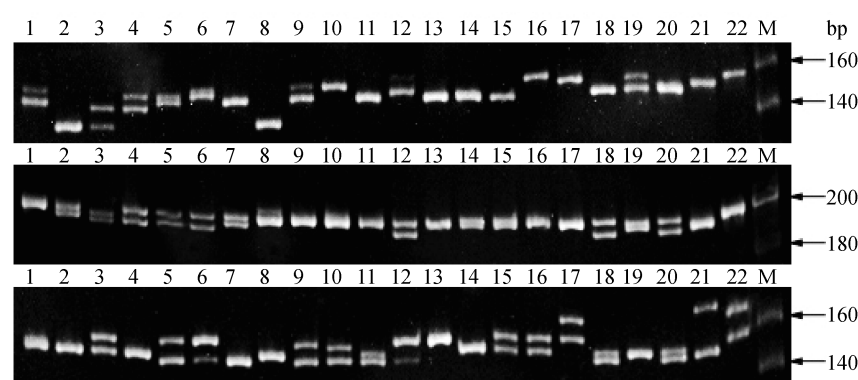


图 2 Tg24、Tg16、Tg26 位点在泥蚶奉化群体部分个体中扩增的电泳图谱

Fig. 2 Amplified allele of loci Tg24, Tg16 and Tg26 in partial Fenghua population of *T. granosa*

表 2 8 对泥蚶多态性 SSR 标记引物在格粗饰蚶中的通用性分析

Tab. 2 Cross-species amplification of 8 polymorphic EST-SSRs from *T. granosa* in *A. craticulata*

位点 locus	重复模式 repetitive mode	泥蚶 <i>T. granosa</i> (n = 30)					格粗饰蚶 <i>A. craticulata</i> (n = 30)				
		等位 基因数 N_a	观测 杂合度 H_o	期望 杂合度 H_e	多态 信息含量 PIC	P 值 P	等位 基因数 N_a	观测 杂合度 H_o	期望 杂合度 H_e	多态 信息含量 PIC	P 值 P
Tg03	(TTG) ₄	5	0.600	0.728	0.664	0.029 6	2	0.125	0.120	0.110	1.000 0
Tg04	(TTG) ₅	4	0.560	0.675	0.595	0.457 2	3	0.292	0.297	0.270	0.500 3
Tg12	(TTTA) ₅	2	0.240	0.216	0.189	1.000 0	3	0.208	0.194	0.178	1.000 0
Tg16	(TTG) ₆	4	0.520	0.456	0.411	1.000 0	3	0.167	0.160	0.150	1.000 0
Tg19	(AGA) ₄	5	0.240	0.519	0.482	0.000 2	3	0.417	0.649	0.561	0.000 0
Tg25	(ATG) ₅	4	0.200	0.555	0.463	0.000 0	3	0.375	0.387	0.348	0.521 0
Tg29	(TA) ₇	3	0.200	0.581	0.475	0.000 0	5	0.458	0.652	0.597	0.044 9
Tg30	(TG) ₆	3	0.280	0.603	0.508	0.001 0	3	0.375	0.531	0.435	0.065 7
平均值 mean	-	3.75	0.355	0.542	0.473	-	3.13	0.302	0.374	0.331	-

利用 8 对多态性 SSR 引物在格粗饰蚶 30 个个体中进行了群体检测,共扩增出 25 个等位基因,每个位点产生的等位基因数在 2~5 之间,平均每对引物产生 3.13 个等位基因,观测杂合度为 0.125~0.458,期望杂合度为 0.120~0.652,多态信息含量为 0.110~0.597,仅有 1 个位点 (Tg19) 显著偏离 Hardy-Weinberg 平衡。相同 SSR 位点在不同物种中扩增出的等位基因数有一定差别,如 Tg03 位点在泥蚶中扩增出 5 个等位基因,而在格粗饰蚶中只扩增出 2 个等位基因;Tg29 位点在泥蚶中扩增出 3 个等位基因,而在格粗饰蚶中扩增出 5 个等位基因(表 2)。然而,大多数位点在两物种中扩增出的等位基因数基本上一致,泥蚶和格粗饰蚶的平均等位基因数分别为 3.75 和 3.13。总体来看,泥蚶的平均观测杂合度、期望杂合度和多态信息含量均比格粗饰蚶要高,这反映泥蚶较高的杂合度和多态性。

3 讨论

3.1 泥蚶 EST-SSR 位点的筛选及多态性分析

随着高通量测序技术的迅猛发展和海量 EST 序列的产生,EST-SSR 技术已成为微卫星开发的主要方法之一。与传统富集文库法相比,EST-SSR 标记的开发过程不需构建文库、克隆测序等复杂步骤,具有开发成本低、效率高、方法简便等优点^[18]。目前在海洋贝类中,已从高通量测序的转录组中鉴别出了大量候选 EST-SSR 标记,如 Clark 等^[19]在南极双壳类 *Laternula elliptica* 外套膜的转录组文库中鉴别出 46 235 个 SSR 位点; Hou 等^[20]从 454 高通量测序获得的虾夷扇贝 (*Patinopecten yessoensis*) 转录组序列中筛查到二至六碱基的候选 SSR 标记 2 700 个; Wang 等^[8]从 75 万条文蛤 454 转录组序列中鉴定出 2 970 个 EST-SSR,并对 40 个位点进行了扩增和分析。本

研究利用泥蚶转录组 454 测序、拼接得到的大量 EST 序列,大规模开发 EST-SSR 标记,为泥蚶 SSR 引物的高效开发开辟途径。

EST 序列大多源自基因编码区,由于其更易受到选择压力的作用从而表现出较高的保守性,因此 EST-SSR 的多态性(H_o 、 H_e 、 PIC 、 N_a)要比基因组 SSR(G-SSR)低^[21-22]。Corderio 等^[23]在甘蔗(*Saccharum* spp.)中筛选出了 17 个 EST-SSR 位点,每个位点平均扩增出 3.3 个等位基因, PIC 值(0.23)远远低于 G-SSR 引物的 PIC 值(0.72)。在小麦^[24]和向日葵^[25]中也发现 EST-SSR 引物的多态性要低于基因组 SSR 的多态性。在贝类中也有相似报道,如 Lu 等^[9]在开发文蛤 G-SSR 和 EST-SSR 位点中发现,基因组来源的 SSR 位点的多态性(H_o 、 H_e 、 PIC 、 N_a)要高于 EST-SSR 位点的多态性;Zhan 等^[26]在栉孔扇贝中也发现 EST-SSR 位点的多态性低于基因组来源的 SSR 位点的多态性。目前,关于泥蚶 EST-SSR 标记的开发国内外还未见报道,仅顾晓英等^[17]利用富集文库法开发了 6 个泥蚶基因组来源的 SSR 位点,每个位点平均扩增出 5.4 个等位基因,平均 PIC 值 0.546,而本研究筛选到的 34 个多态性 EST-SSR 位点,平均每个位点扩增出 3.59 个等位基因,平均 PIC 值为 0.4,多态性明显低于基因组来源的 SSR 位点。

EST-SSR 来自于基因的转录区,很可能与物种的功能基因连锁,在一定程度上更容易标记出功能基因,反映物种真实的遗传特性^[2]。在本研究中,筛选到得到的 34 个 EST-SSR 位点中,有 25 个位点来自功能基因注释,这将为鉴定泥蚶功能基因、重要功能基因定位和分子标记辅助育种奠定重要基础。

3.2 泥蚶 EST-SSR 的通用性研究

EST-SSR 源于基因的转录区,侧翼序列在不同物种间具有高度保守性,故在属内具有很高的通用性,甚至在同科不同属间也具较高的通用性^[27],但 SSR 引物的通用性会随着种间遗传距离的增加而下降^[28]。EST-SSR 的通用性研究,有利于在不同物种间进行比较作图、系统进化分析、基因发掘和 QTLs 定位等研究。目前,在海洋双壳贝类中已开展了一些近缘物种间 EST-SSR 的通用性研究,如 Huvet 等^[29]分析了 4 个太平洋牡蛎微卫星在巨蛎属(*Crassostrea*)及小蛎属(*Saccostrea*)共 12 种

牡蛎中的保守性,结果仅在近缘种葡萄牙牡蛎(*C. angulata*)、熊本牡蛎(*C. sikamea*)和明巨牡蛎(*C. ariakensis*)中有扩增;Li 等^[30]将从太平洋牡蛎中新开发出的 15 对 EST-SSR 引物在巨蛎属的褶牡蛎(*C. plicatula*)、香港巨牡蛎(*C. Hongkongensis*)、明巨牡蛎、日本巨牡蛎(*C. nippona*)、熊本牡蛎 5 个近缘种间进行了扩增,结果发现 15 对引物均获得了成功,至少在一个近缘种中扩增出来,认为这些位点对群体遗传分析和基因定位有较大价值。Wang 等^[31]用 32 对太平洋牡蛎多态性微卫星引物在三角帆蚌(*Hyriopsis cumingii*)中进行了种间适用性扩增,13 对成功获得扩增条带,10 个位点表现多态性,通用率为 31.25%。目前,关于 EST-SSR 在蚶科贝类中的通用性研究尚未见报道。本实验在获得 34 对泥蚶多态性 EST-SSR 引物的基础上,对其在蚶科不同属的格粗饰蚶中进行了通用性检测,结果 11 对引物能成功扩增,其中 8 对表现为多态,通用率相对较低(25.81%)。微卫星的通用性及多态位点比例随着物种间遗传距离的增加而呈下降的趋势^[32]。泥蚶和格粗饰蚶 EST-SSR 位点的通用率低,可能与二者隶属于不同的属,种间遗传距离较大有关。

有研究表明,SSR 核心单元碱基重复类型对 SSR 位点在物种中的通用性具有很大的影响^[33]。在植物中,一般核心单元三核苷酸碱基重复类型的 SSR 通用性比例最高,其次是二核苷酸碱基重复类型^[33-36]。文明富等^[37]研究了木薯 EST-SSR 在麻疯树和橡胶树中的通用性,发现木薯三核苷酸重复类型的 EST-SSR 引物的通用性比例在麻疯树和橡胶中均最高。在本研究中,8 对多态性引物中有 5 对为三碱基重复类型,2 对为二碱基重复类型,表明三碱基和二碱基重复类型的 SSR 位点有较高的通用性,这与植物中的研究结果一致。本实验为蚶科贝类借助泥蚶转录组大量开发 EST-SSR 引物开辟了新途径,通用性 SSR 标记的开发必将在蚶科贝类种质资源保护、遗传多样性评价、系统进化分析以及比较基因组学等研究中发挥重要作用。

宁波市海洋与渔业研究院的尤仲杰研究员对本实验中的格粗饰蚶进行了种类鉴定,谨表谢忱。

参考文献:

- [1] Wright J M, Bentzen P. Microsatellites: genetic markers for future[J]. Reviews in Fish Biology and

- Fisheries, 1994, 4(3): 384 - 388.
- [2] Subramanian S, Mishra R K, Singh L. Genome-wide analysis of microsatellite repeats in humans; their abundance and density in specific genomic regions [J]. Genome Biology, 2003, 4(2): 1301 - 1310.
- [3] 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.
- [4] 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.
- [5] 张秀英, 张晓军, 赵翠, 等. 栉孔扇贝 BES-SSR 的开发及遗传多样性分析 [J]. 水产学报, 2012, 36(6): 815 - 824.
- [6] Li Q, 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(3): 313 - 316.
- [7] Bai J, Li Q, Cong R, et al. Development and characterization of 68 expressed sequence tag derived simple sequence repeat markers in the Pacific oyster, *Crassostrea gigas* [J]. Journal of the World Aquaculture Society, 2011, 42(3): 444 - 455.
- [8] Wang H X, Huan P, Lu X, et al. Mining of EST-SSR markers in clam *Meretrix meretrix* larvae from 454 shotgun transcriptome [J]. Genes and Genetic Systems, 2011, 86(3): 197 - 205.
- [9] Lu X, Wang H X, Dai P, et al. Characterization of EST-SSR and genomic-SSR markers in the clam, *Meretrix meretrix* [J]. Conservation Genetics Resource, 2011, 3(4): 655 - 658.
- [10] 朱东丽, 董迎辉, 林志华, 等. 利用微卫星标记对文蛤 4 个壳色花纹品系的遗传分析 [J]. 水产学报, 2011, 26(2): 202 - 209.
- [11] 石耀华, 洪葵, 郭希明, 等. 马氏珠母贝 EST 微卫星的筛选 [J]. 水产学报, 2008, 32(2): 174 - 181.
- [12] 刘博, 邵艳卿, 滕爽爽, 等. 缢蛭 (*Sinonovacula constricta*) EST-SSR 分布特征及引物开发利用 [J]. 海洋与湖沼, 2012, 43(1): 132 - 137.
- [13] 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.
- [14] 李太武, 李成华, 宋林生, 等. 福建南北泥蚶种内分化的 RAPD 分析 [J]. 动物学研究, 2003, 24(5): 362 - 366.
- [15] 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.
- [16] 曾庆国, 林志华, 尤仲杰. 泥蚶 (*Tegillarca granosa*) GT 微卫星位点的筛选和性质鉴定 [J]. 海洋与湖沼, 2008, 39(2): 174 - 177.
- [17] 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.
- [18] Varshney R K, Graner A, Sorrells M E. Genomic microsatellite markers in plants: features and applications [J]. Trends in Biotechnology, 2005, 23(1): 48 - 55.
- [19] Clark M, Thorne M A, Vieira F A, et al. Insight into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing [J]. BMC Genomics, 2010, 11: 362.
- [20] Hou R, Bao Z M, Wang S, et al. Transcriptome sequencing and *de novo* analysis for Yesso scallop (*Patinopecten yessoensis*) using 454 GS FLX [J]. PloS One, 2011, 6(6): e21560.
- [21] Chabane K, Ablett G A, Cordeiro G M, et al. EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley [J]. Genetic Resources and Crop Evolution, 2005, 52: 903 - 909.
- [22] Eujayl I, Sorrells M E, Baum M, et al. Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat [J]. Theoretical and Applied Genetics, 2002, 104: 399 - 407.
- [23] Cordeiro G M, Casu R, McIntyre C L, et al. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum [J]. Plant Science, 2001, 160: 1115 - 1123.
- [24] Gupta P K, Rustgi S, Sharma S, et al. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat [J]. Molecular Genetics and Genomics, 2003, 270: 315 - 323.
- [25] Pashley C H, Ellis J R, McCauley D E, et al. EST databases as a source for molecular markers: Lessons from *Helianthus* [J]. Journal of Heredity, 2006, 97: 381 - 388.
- [26] 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.
- [27] Semagn K, Bjrnsstad A, Ndjioudjop M N. An

- overview of molecular marker methods for plants [J]. African Journal of Biotechnology, 2006, 5: 2540 – 2568.
- [28] Barbar T, Palma-Silva C, Paggi G M, *et al.* Cross-species transfer of nuclear microsatellite markers; potential and limitations [J]. Molecular Ecology, 2007, 16: 3759-3767.
- [29] Huvet A, Boudry P, Ohresser M, *et al.* Variable microsatellites in the Pacific oyster *Crassostrea gigas* and other cupped oyster species [J]. Animal Genetics, 2000, 31: 6 – 79.
- [30] Li Q, Liu S K, Kong L F. Microsatellite within genes and ESTs of the Pacific oyster *Crassostrea gigas* and their transferability in five other *Crassostrea* species [J]. Electronic Journal of Biotechnology, 2009, 12 (3): 1 – 6.
- [31] Wang G L, Wang J J, Li J L. Preliminary study on applicability of microsatellite primers developed from *Crassostrea gigas* to genomic analysis of *Hyriopsis cumingii* [J]. Journal of Fisheries of China, 2006, 30 (1): 15 – 20.
- [32] Primmer C R, Moiler A P, Ellegren H. A wide range survey of cross-species microsatellite amplification in birds [J]. Molecular Ecology, 1996, 5: 365 – 378.
- [33] Thiel T, Michalek W, Varshney R K, *et al.* Exploiting EST database for the development and characterization of gene derived SSR markers in barley (*Hordeum vulgare* L.) [J]. Theoretical and Applied Genetics, 2003, 106: 411 – 422.
- [34] Gao L F, Tang J F, Li H W, *et al.* Analysis of microsatellites in major crops assessed by computational and experimental approaches [J]. Molecular Breeding, 2003, 12: 245 – 261.
- [35] Chen H M, Li L Z, Wei X Y, *et al.* Development, chromosome location and genetic mapping of EST-SSR markers in wheat [J]. Chinese Science Bulletin, 2005, 50: 2328 – 2336.
- [36] Li L Z, Wang J J, Guo Y, *et al.* Development of SSR markers from ESTs of gramineous species and their chromosome location on wheat [J]. Progress in Natural Science, 2008, 18(12): 1485 – 1490.
- [37] 文明富, 陈新, 王海燕, 等. 木薯基因组 SSR 和 EST-SSR 在麻疯树和橡胶树中的通用性分析 [J]. 作物学报, 2011, 37(1): 74 – 78.

Characterization of 34 polymorphic EST-SSR markers in *Tegillarca granosa* and their transferability in *Anadara craticulata*

DONG Yinghui¹, WU Guoxing¹, YAO Hanhan¹, BIAN Pingjiang², ZHAO Demin¹, LIN Zhihua^{1*}

(1. Zhejiang Key Laboratory of Aquatic Germplasm Resources, Zhejiang Wanli University, Ningbo 315100, China;

2. Zhejiang Mariculture Research Institute, Wenzhou 325005, China)

Abstract: EST-SSR is one of important methods to isolate microsatellite loci, which is low cost, high transferability among phylogenetically related species and a probability of being associated with functional regions of the genome. 34 polymorphic EST-SSR markers were isolated and characterized in the blood clam, *Tegillarca granosa*, from the ESTs generated by 454 sequencing. The number of alleles for the 34 SSR markers in 30 individuals of Zhejiang Fenghua population varied from 2 to 7, with an average of 3.59 alleles per locus. The observed heterozygosity ranged from 0.000 to 0.600, while the expected heterozygosity varied from 0.078 to 0.771. Polymorphic information content ranged from 0.106 to 0.718, and 13 SSR loci deviated from Hardy-Weinberg equilibrium. 25 EST-SSR identified from annotated genes were expected to be useful for mapping these genes in linkage maps. Interspecific transferability of the 34 markers revealed that 8 were polymorphic in *Anadara craticulata*, resulting in a transferability rate of 23.53%. To our knowledge, this is the first report on the survey of SSR transferability in blood clams. These novel polymorphic EST-SSR markers should be particularly useful for comparative mapping, gene tagging and QTL mapping among blood clams.

Key words: *Tegillarca granosa*; *Anadara craticulata*; EST-SSR; transferability

Corresponding author: LIN Zhihua. E-mail: zhihua9988@126.com