

基于转录组数据的马氏珠母贝 EST-SSR 位点的信息分析及其多态性检测*

王忠良^{1,2} 丁 燊^{1,4} 许尤厚³ 简纪常^{1,4} 鲁义善^{1,4}
王 蓓^{1,4} 陈 刚^{1,2} 吴灶和^{4,5}

(1. 广东海洋大学水产学院 湛江 524088; 2. 广东海洋大学南海水产经济动物增殖重点实验室 湛江 524088; 3. 广西北部湾海洋生物多样性养护重点实验室(钦州学院) 钦州 535000; 4. 广东省水产经济动物病原生物学及流行病学重点实验室 湛江 524088; 5. 仲恺农业工程学院 广州 510225)

摘要 为获得稳定可靠的马氏珠母贝 SSR 分子标记, 本文利用 MISA 软件对转录组测序数据进行了大规模的 EST-SSR 标记发掘, 并分析了位点信息及其多态性。结果表明, 马氏珠母贝转录组测序所获得的 74007 条 Unigenes 中检测出 9872 个 EST-SSR 位点, 位点出现频率为 13.34%, 平均每 5102 bp 含有 1 个 SSR 位点。在转录组 SSR 中共有 132 种重复基元类型, 其中单核苷酸重复基元为主要类型, 占 SSR 总数的 81.46%; 单核苷酸重复以 A/T 基序为主, 占 SSR 总数的 71.27%。基于筛选的 SSR 序列, 应用 Primer3 软件进行引物的批量设计, 共为 5922 条 EST-SSR 序列成功设计出 17766 对引物。随机选择 80 对引物对 EST-SSR 多态性进行检测, 共有 62 对引物成功扩增出稳定条带, 占引物总数的 77.5%; 其中, 17 对 EST-SSR 引物表现出个体多态性, 多态性比率为 27.42%。以上研究为马氏珠母贝遗传多样性、遗传图谱构建及分子辅助育种研究提供了有效工具, 对于马氏珠母贝种质资源保护、优良品种培育和珍珠养殖业的健康发展均具有重要意义。

关键词 马氏珠母贝; 简单重复序列; 转录组; 分子标记
中图分类号 Q789 **doi:** 10.11693/hyhz20141200361

马氏珠母贝(*Pinctada martensii*)又称合浦珠母贝(*Pinctada fucata*), 是人工培育海水珍珠的主要母贝。自 1965 年我国成功开展马氏珠母贝人工育苗以来, 海水珍珠养殖业发展迅速, 尤其是进入 20 世纪 90 年代后, 海水珍珠养殖已成为广东、广西和海南沿海地区的支柱产业(史兼华等, 2006; Wang *et al.*, 2009)。然而, 由于长期的人工养殖和近亲繁殖, 马氏珠母贝种质已呈现明显的退化现象, 如种苗活力下降、死亡率增加、育珠贝变小、珍珠质量差等(Miyazaki *et al.*, 1999; Tomaru *et al.*, 2001; He *et al.*, 2008)。为改善这一状况, 马氏珠母贝的遗传改良和优良品种培育是当前海水

珍珠养殖业的首要任务, 而筛选稳定可靠的分子标记评价其遗传多样性对马氏珠母贝良种选育和种质资源保护都具重要的指导意义。

简单重复序列(Simple Sequence Repeats, SSR), 又称微卫星 DNA, 是指由 1-6 个核苷酸组成的简单串联重复 DNA 序列(Powell *et al.*, 1996; Liu *et al.*, 2013)。与诸多分子标记相比, 具有数量丰富、多态性高、多等位性、共显性和可重复性高等优点, 已广泛应用于动植物遗传图谱构建、种质鉴定、基因定位、遗传多样性、分类与进化及分子标记辅助育种等研究领域(赵莹莹等, 2006; 徐美佳等, 2009; 李云峰等, 2010;

*国家自然科学基金项目, 31202023 号; 广东海洋大学优秀青年骨干教师特别资助计划, 2014001 号; 广西北部湾海洋生物多样性养护重点实验室(钦州学院)开放课题, 2015KB05 号。王忠良, 博士, 讲师, E-mail: leong2006@126.com

通讯作者: 吴灶和, 教授, 博士生导师, E-mail: wuzaohe@163.com

收稿日期: 2014-12-17, 收修改稿日期: 2015-01-26

孙国华等, 2010; 刘博等, 2012; 王东等, 2014)。根据来源, SSR 标记又可分为基因组 SSR(Genomic SSR, gSSR)和表达序列标签 SSR(Expressed sequence tag SSR, EST-SSR)两种。与 gSSR 标记相比, EST-SSR 标记无需构建基因组文库, 节省了大量人力物力, 开发成本较低, 且来源于基因的转录区, 其多态性与基因功能可能直接相关, 同时在相近物种间具有良好的通用性(Varshney *et al.*, 2005; 李琪, 2006; 张琼等, 2010)。

近年来, 随着 EST 和 cDNA 大规模测序技术的快速发展, 多种海产经济贝类 SSR 标记的开发利用研究已相继展开(李红蕾等, 2003; Sato *et al.*, 2005; 李琪, 2006; 李云峰等, 2010), 但现有的 SSR 标记还非常有限, 仍不能满足进一步研究的需要。为此, 本文基于前期 Illumina 高通量测序技术获得的马氏珠母贝转录组数据(未发表), 利用计算机软件辅助进行大规模 EST-SSR 标记的发掘, 并对其组成、分布及特征等信息进行了分析, 同时对 SSR 标记的多态性进行了评价, 以期期为马氏珠母贝遗传多样性、遗传图谱构建及分子辅助育种研究提供有效工具, 促进马氏珠母贝种质资源保护、优良品种培育和珍珠养殖业的健康发展。

1 材料与方法

1.1 马氏珠母贝的暂养及血细胞采集

马氏珠母贝(平均壳长 70mm)购自广东省湛江市徐闻迈陈珍珠贝养殖场, 暂养于室内玻璃钢水槽中(200L 水体), 每槽饲养 20 只; 水温 25°C, 盐度 28, 饲养期间连续充气; 暂养过程中以螺旋藻粉为饵料投喂, 每天换水(100%)一次; 室内暂养一周后采集血淋巴。

利用 1mL 无菌注射器于马氏珠母贝闭壳肌处采集血淋巴, 每只 0.5mL; 采集的血淋巴分装至 1.5mL 离心管中, 每管 1mL; 4°C, 800g 离心 10min 收集血细胞, 并立即进行 DNA 提取。

1.2 马氏珠母贝转录组数据来源

马氏珠母贝转录组数据来源于课题组前期利用 Illumin/HiSeq-2000 高通量测序平台对马氏珠母贝血细胞进行的全转录组测序。转录组测序共获得 70407878 条原始读数(Raw reads), 经去除含有接头、重复及测序质量较低的原始读数后, 获得 56345139 条纯净读数(Clean reads)。使用转录组 *de novo* 组装软件 Trinity(Grabherr *et al.*, 2011)对 Clean reads 进行组装, 并进行去冗余处理和进一步拼接, 共得到 74007 条非冗余独立基因(Unigene)。(以上转录组数据尚未发表,

转录组 Raw reads 已提交至 NCBI 的 SRA 数据库, 登录号为 SRP041567。)

1.3 马氏珠母贝转录组 EST-SSR 位点筛选

为检测马氏珠母贝转录组中 EST-SSR 位点, 使用软件 MicroSatellite (MISA) (Lu *et al.*, 2013)对组装得到的 Unigenes 序列进行 SSR 搜索和定位。所检测 SSR 位点共 6 类, 分别为单核苷酸重复、二核苷酸重复、三核苷酸重复、四核苷酸重复、五核苷酸重复和六核苷酸重复, 其筛选标准为: 单核苷酸重复 ≥ 10 次; 二核苷酸重复 ≥ 6 次; 三核苷酸重复、四核苷酸重复、五核苷酸重复和六核苷酸重复 ≥ 5 次。

1.4 马氏珠母贝转录组 EST-SSR 引物设计

使用 Primer3 软件进行 SSR 引物设计, 每条 SSR 产生 3 对引物; 引物设计的主要参数设置如下: 引物长 18—23bp, 最适长度 23 bp; PCR 扩增产物长度 100—300bp; GC 含量 40%—65%, 最适含量 50%; 退火温度 55—65°C, 最佳退火温度 55°C。

1.5 马氏珠母贝转录组 EST-SSR 多态性检测

1.5.1 模板 DNA 的制备 采用 Universal Genomic DNA Mini-Isolation Kit 试剂盒(上海生工生物工程有限公司), 并按照说明书方法进行血细胞 DNA 的提取。经紫外分光光度计定量后, -20°C 保存备用。

1.5.2 PCR 扩增及产物检测 为了筛选和验证所设计的 SSR 引物, 随机挑选 80 对引物(引物由上海生工生物工程有限公司合成), 以上述血细胞 DNA 为模板进行 PCR 扩增。PCR 反应体系为 20 μ L, 包括 2 \times ES Taq MIX 10 μ L, 上下游引物(10 μ mol/L)各 1 μ L, 模板 DNA 1 μ L, ddH₂O 补足至 20 μ L。PCR 扩增程序为: 94°C 预变性 5 min, 30 个循环(94°C 变性 30s, 最佳 T_m 退火 30s, 72°C 延伸 30s), 最后 72°C 延伸 5 min。PCR 扩增产物用 8%非变性聚丙烯酰胺凝胶电泳结合银染法显色进行检测。

2 结果与分析

2.1 马氏珠母贝转录组 EST-SSR 的分布及频率

利用 MISA 软件对马氏珠母贝转录组中的 74007 条 Unigenes 序列进行搜索, 共检测到 9872 个 SSR 位点, 分布于 8135 条 Unigenes 序列中。SSR 发生频率(含 SSR 的 Unigene 数/Unigene 总数)为 10.99%, SSR 出现频率(检出 SSR 位点数/Unigene 总数)为 13.34%; 平均每 5102bp 含有 1 个 SSR 位点(检测序列总碱基数 50368155 bp)。其中, 1350 条 Unigene 含有 1 个以上 SSR 位点, 复合型 SSR 数目为 518 个。

马氏珠母贝 EST-SSR 重复类型丰富, 除六核苷酸重复外, 1—5 核苷酸重复均有发现。其中, 单核苷酸重复有 4 种类型, 二核苷酸重复有 11 种类型, 三核苷酸重复有 52 种类型, 四核苷酸重复有 58 种类型, 五核苷酸重复有 7 种类型(表 1)。从各类型 SSR 位点数量看, 出现最多的为 1—3 核苷酸重复, 占总 SSR 位点数的 98.32%。其中, 单核苷酸重复比例最高, 占 81.46%; 其次为二核苷酸重复和三核苷酸重复, 分别占 10.48%和 6.38%; 4—5 核苷酸重复 SSR 数量较少, 共占 1.68%(表 1)。

表 1 马氏珠母贝 EST-SSRs 不同重复基元分布情况
Tab.1 Distribution of different repeat motifs of EST-SSRs in *P. martensii* transcriptome

重复类型	重复基元类型数量	SSR 数量	占总 SSR 比例(%)	SSR 分布频率(%)
单核苷酸	4	8042	81.46	10.87
二核苷酸	11	1035	10.48	1.40
三核苷酸	52	630	6.38	0.85
四核苷酸	58	158	1.60	0.21
五核苷酸	7	7	0.08	0.01
总计	132	9872	100	13.34

马氏珠母贝转录组中 EST-SSR 的长度在 12—20bp 的有 3677 条, 其中单核苷酸重复 SSR 共计 2226 条; 长度超过 20bp 的 SSR 有 336 条, 并以单核苷酸重复 SSR 最多, 达 129 条(表 2)。SSR 位点重复次数方面, 以重复 10 次(3883)最多, 占总 SSR 位点的 39.33%; 其次为重复 11 次和 12 次, 分别占 17.12%和 8.88%; 重复 5—9 次、13—17 次及 18—24 次的 SSR 位点个数分别为 1671、1279 和 472(图 1)。

2.2 马氏珠母贝转录组 EST-SSR 的特征

在马氏珠母贝转录组 9872 个 SSR 位点中, 共检测到 132 种重复基元(表 1)。其中, 以单核苷酸重复基元 A/T 和 C/G 最多, 分别占总 SSR 的 71.27%和

表 2 马氏珠母贝 EST-SSRs 长度分布情况
Tab.2 Lengthdistribution of EST-SSRs in *P. martensii* transcriptome

重复类型	12—20 bp	>20bp
单核苷酸	2226	129
二核苷酸	855	59
三核苷酸	477	91
四核苷酸	119	51
五核苷酸	0	6
总计	3677	336

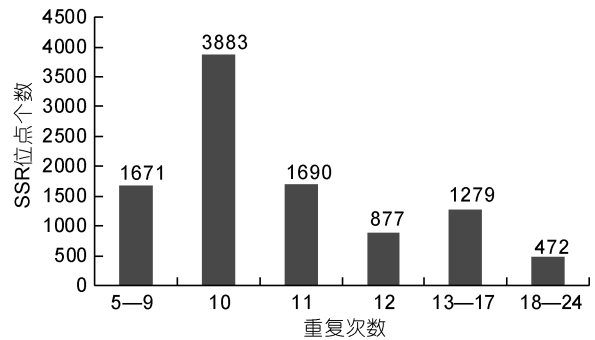


图 1 马氏珠母贝 EST-SSRs 重复次数分布图
Fig.1 Distribution of EST-SSR repeats frequency in *P. martensii* transcriptome

10.19%, 发生频率分别为 9.507%和 1.359%; 其次为二核苷酸重复基元 AT/AT 和 AG/CT, 分别占 4.66%和 4.24%; 三核苷酸重复单元中, 以 ATC/ATG、AAG/CTT 和 AAC/GTT 出现次数最多, 分别占总 SSR 的 2.07%、1.29%和 1.16%; 四核苷酸和五核苷酸重复基元种类较多, 但数量较少, 频率较低, 共占 SSR 总数的 1.67%(表 3)。

2.3 马氏珠母贝转录组 EST-SSR 引物设计及多态性检测

为更好地利用马氏珠母贝 SSR 位点, 本文应用 Primer3 软件对上下游序列均不小于 150 bp 的 EST-SSR 设计引物, 每条序列产生 3 对引物。经 SSRfinder 校验及除去不符合条件的引物后, 共为 5922 条 EST-SSR 序列成功设计出 17766 对引物, 占马氏珠母贝 SSR 总数的 59.99%。

从上述设计引物中随机挑选 80 对引物, 并以随机挑选的 20 只健康马氏珠母贝血细胞基因组 DNA 为模板进行 PCR 扩增以验证设计的 EST-SSR 引物。结果表明, 共有 62 对引物成功扩增出稳定条带, 占引物总数的 77.5%, 但其中有 23 对引物扩增产物片段与预期产物片段大小不符; 在 62 个 SSR 位点中, 以二碱基、三碱基重复类型的 SSR 位点扩增成功率较高; 此外, 17 对 EST-SSR 引物(占扩增引物总数的 21.25%)表现出个体多态性(表 4), 多态性比率为 27.42%; 17 对多态性引物所在 SSR 位点中, 有 7 个为二碱基重复, 6 个为三碱基重复, 2 个为单碱基重复, 四碱基重复和五碱基重复各 1 个。

3 讨论

目前, 有关贝类微卫星标记的研究主要是通过构建 cDNA 文库来获得含有微卫星的序列, 并且设计

引物对微卫星标记进行筛选(Sato *et al.*, 2005; 李春艳等, 2009)。如石耀华等(2008)从马氏珠母贝 9 个组织的 SMART cDNA 文库中共获得 6979 条 EST 序列, 并筛选到 268 个 SSR 位点, SSR 出现频率为 3.48%; 同样, 李云峰等(2010)从虾夷扇贝 (*Patinopecten yessoensis*) cDNA 文库中发现了 74 个 SSR 重复序列(2007 条 EST 序列), SSR 出现频率为 3.69%。与上述传统 EST-SSR 标记开发技术相比, 基于高通量转录组测序的 SSR 标记开发技术不仅可以借助更全面的数据信息, 提高遗传多样性研究的准确性, 而且开发的 SSR 标记同样具有 EST-SSR 标记的诸多优点。王学颖(2012)从马氏珠母贝珍珠囊转录组的 99127 条无冗余 EST 序列中共发现了 6595 个 SSR 位点, SSR 出现频率为 6.65%; 同样, 本文利用 MISA 软件从马氏

珠母贝转录组 74007 条 Unigenes 中共检索出 9872 个 SSR 位点, 分布于 8135 条 Unigenes 中, SSR 出现频率高达 13.34%。因此, 结合其它动植物 SSR 开发相关文献和本研究结果表明, 基于转录组数据筛选 SSR 标记是发掘非模式生物 EST-SSR 的一种快速、高效的方法, 应用前景广阔。

根据现有报道, 虾夷扇贝和马氏珠母贝的 EST-SSR 均以二核苷酸重复基元为主, 其所占比例分别为 40.54%和 48.5% (石耀华等, 2008; 李云峰等, 2010); 缢蛏(*Sinonovacula constricta*)SSR 以三核苷酸重复 SSR 比列最高, 为 37.13%(刘博等, 2012); 此外, 基于马氏珠母贝珍珠囊转录组的 SSR 分析中, 四碱基重复基元占总 EST-SSR 的 40%(石耀华等, 2008)。而本研究中以单核苷酸重复基元为主, 比列达到

表 3 马氏珠母贝 EST-SSRs 重复基元的类型及频率
Tab.3 EST-SSR repeats motifs and the frequency in *P. martensii* transcriptome

重复类型	重复基元	数量	各重复类型总数	所占比例(%)	发生频率(%)		
单核苷酸	A/T	7036	8042	71.27	9.507		
	C/G	1006		10.19	1.359		
二核苷酸	AC/GT	155	1035	1.57	0.209		
	AG/CT	419		4.24	0.566		
	AT/AT	460		4.66	0.622		
	CG/CG	1		0.01	0.001		
三核苷酸	AAC/GTT	115	630	1.16	0.155		
	AAG/CTT	127		1.29	0.172		
	AAT/ATT	66		0.67	0.089		
	ACC/GGT	30		0.30	0.041		
	ACG/CGT	12		0.12	0.016		
	ACT/AGT	10		0.10	0.014		
	AGC/CTG	8		0.08	0.011		
	AGG/CCT	57		0.58	0.077		
	ATC/ATG	204		2.07	0.276		
	CCG/CGG	1		0.01	0.001		
	四核苷酸	ACAG/CTGT		48	158	0.49	0.065
		AAAT/ATTT		33		0.33	0.045
AATG/ATTC		17	0.17	0.023			
AAAC/GTTT		17	0.17	0.023			
其它四核苷酸		43	0.44	0.058			
五核苷酸	AAAAG/CTTTT	1	7	0.01	0.001		
	AACAG/CTGTT	1		0.01	0.001		
	AACTG/AGTTC	1		0.01	0.001		
	AAGAT/ATCTT	1		0.01	0.001		
	AATAC/ATTGT	1		0.01	0.001		
	AATCC/ATTGG	1		0.01	0.001		
	AATTC/AATTG	1		0.01	0.001		

表 4 17 对多态性 EST-SSRs 引物序列及扩增参数
Tab.4 Sequences and amplification parameters of the 17 EST-SSRs primers

序号	引物编号	SSR 类型	引物序列(5'—3')	产物大小(bp)	退火温度(°C)
1	PF-13	(TA)6	F: AGACTTTGATCTTTTGCCAGATGC R: ACGTAAATCTGCTGTACTGAAAAA	118	58
2	PF-42	(CGTA)5	F: TTGTTTCGTGCAATGTGGCAA R: AGCACAAGAGAGACACTGCG	279	59
3	PF-35	(AC)8	F: GGTCACCTCGTTTGATTGCA R: GCAGACGTAGACCAGCTGTT	271	57
4	PF-61	(ACC)6	F: AGAGACCAATGGGCAATCGG R: ATGAGCTCGCTTCTGTGAC	252	60
5	PF-07	(CT)6	F: TCCTCTCAATTCGATTGACCCT R: GGCTGTGGAACCTTAGAACTCA	130	58
6	PF-44	(AT)6	F: CCTTTGTTGTTGCCACCACC R: GCACCCCTGCATTGCTTAAC	278	60
7	PF-28	(AATAC)5	F: TGCCTAGGGGACAATTCGC R: GACATGTATGCCTCTCGGGG	199	59
8	PF-54	(T)10	F: GGGAGTGAATGTCAAGTCA R: GGTCCTGAAAGTTCACATGGC	138	57
9	PF-53	(GTT)5	F: GCTTCGCTCACAACACCAT R: ACATGAACATGGGACGGGAC	171	59
10	PF-67	(TC)10	F: ATCACTAGGGCAGCTACAGC R: GCCTTCCTCGGCTTATCTGA	177	58
11	PF-24	(AG)6	F: ATCACAAAACGGAGAGCGT R: CCAATGAAGAGCTTGGTGCAC	125	57
12	PF-73	(ATG)7	F: CGACTTGCCCTGATGATCGA R: ACCTGAAGTCACTTGAAGTGT	233	58
13	PF-40	(TTC)5	F: ATGACCCCTATCTGCCTGGT R: GCAGCTCAGACAAGACGGAT	277	60
14	PF-03	(GGA)5	F: GGACATGGGCTGGACAAAGA R: TCCATCTCCACCTTACCT	207	56
15	PF-22	(AG)8	F: TGATACATGCACATTCCTCTGA R: AGAAAGGGCAAGTGAGTAGGT	130	57
16	PF-13	(T)10	F: TGTGAGACCGATTCTGCACC R: CTCCCACAACCTCACCCAG	148	60
17	PF-49	(TGA)6	F: AACAGCCCGCGAATCTAAA R: TTTCAGCTGCCTTGTTGAGT	231	58

81.46%。造成以上差异的原因, 除 EST 序列获取方法不同外(王学颖, 2012), 在一定程度上与搜索 SSR 时设置的参数有很大的关系(Wei *et al*, 2011)。Aggarwal 等(2007)发现, 如果改变 SSR 搜索标准(提高二核苷

酸重复基元的重复次数, 或降低三核苷酸重复基元的重复次数), 原本二核苷酸重复 SSR 占优势的研究中, 便会出现三核苷酸重复 SSR 占优势的情况; 同样, 若本研究中提高单核苷酸 SSR 的重复次数, 则 SSR

将以二核苷酸重复基元为主。因此,亟需建立一个统一的 SSR 检测参数标准以便有效比较不同研究中同一或相近物种 SSR 分布的类型与特点。

研究表明,SSR 标记可用性的重要依据——多态性的高低,主要取决于的其长度大小(李珊等,2010)。当 SSR 长度 $\geq 20\text{bp}$ 时其多态性较高,长度在 12—20bp 之间的 SSR 则呈现中等多态性,而长度 $< 12\text{bp}$ 时多态性较低(杨华等,2011)。本研究中,长度在 12—20bp 之间的 SSR 有 3677 条(占 SSR 总数的 37.25%),这些 SSR 具中等多态性;而长度 $\geq 20\text{bp}$ 的 SSR 有 336 条(占 SSR 总数的 3.4%),此类 SSR 呈较高的多态性。此外,本研究中 20bp 以上的低级重复基元(一、二、三核苷酸重复基元)较多,共计 279 条(占 20bp 以上 SSR 的 83.04%)。由此可预见马氏珠母贝转录组来源的 SSR 具有较高的多态性潜能,在分子标记研究方面将具有较高的利用价值(Dreisigacker *et al*, 2004)。

为检测转录组中 EST-SSR 的多态性,随机挑选了 80 对引物进行 PCR 扩增,其中 62 对引物成功扩增出稳定条带,但仍有 18 条引物未能扩增出产物。这可能是由于所设计的引物序列位于两个外显子上,或者基因组对应序列含有内含子而不具备 SSR 序列特征等原因造成(Varshney *et al*, 2005)。此外,本研究出现了 23 对引物扩增产物片段与预期产物片段大小不符的现象,而该现象同样出现在栉孔扇贝和缢蛭的 EST-SSR 扩增产物中(李红蕾等,2003;刘博等,2012)。究其原因,有学者认为其与引物的非特异性结合(错配)或基因组 DNA 序列含有内含子等相关(李红蕾等,2003; Saha *et al*, 2004)。17 对多态性引物所在 SSR 位点中,多为二、三核苷酸重复,这从某种程度上也验证了低级重复基元多态性高于高级重复基元多态性的推断(Dreisigacker *et al*, 2004)。

4 结论

本研究利用马氏珠母贝的转录组数据,筛选获得了 9872 个 EST-SSR 位点,位点出现频率达 13.34%;同时,共为 5922 条 EST-SSR 序列成功设计出 17766 对引物;随机选择的 80 对扩增引物中,共有 17 对 EST-SSR 引物表现出个体多态性,多态性比率为 27.42%。以上研究结果对于丰富马氏珠母贝分子标记类型、加速功能基因资源的利用、探究种群遗传结构、分析遗传多样性、保护种质资源和培育优良品种均具有重要意义。

参 考 文 献

- 王 东,曹玲亚,高建平,2014. 党参转录组中 SSR 位点信息分析. 中草药,45(16): 2390—2394
- 王学颖,2012. 马氏珠母贝微卫星引物的批量开发及选育系和珍珠囊的标记分析. 湛江: 广东海洋大学硕士学位论文
- 石耀华,洪 葵,郭希明等,2008. 马氏珠母贝 EST 微卫星的筛选. 水产学报,32(2): 174—181
- 史兼华,何毛贤,黄良民等,2006. 马氏珠母贝选育系遗传变异的 AFLP 分析. 高技术通讯,16(5): 534—538
- 刘 博,邵艳卿,滕爽爽等,2012. 缢蛭(*Sinonovacula constricta*)EST-SSR 分布特征及引物开发利用. 海洋与湖沼,43(1): 132—137
- 孙国华,杨建敏,宋志乐等,2010. 刺参(*Apostichopus japonicus*)EST 序列中微卫星分布分析及其标记的筛选. 海洋与湖沼,41(1): 133—139
- 李 珊,周天华,赵桂仿等,2010. 马蹄香表达序列标签资源的 SSR 信息分析. 中草药,41(3): 464—468
- 李 琪,2006. 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. 中国水产科学,13(3): 502—509
- 李云峰,刘卫东,高祥刚等,2010. 虾夷扇贝外套膜和肾脏组织 cDNA 文库构建以及 EST 的初步分析. 中国水产科学,17(3): 578—585
- 李红蕾,宋林生,王玲玲等,2003. 栉孔扇贝 EST 中微卫星标记的筛选. 高技术通讯,13(12): 72—75
- 李春艳,丁 君,常亚青等,2009. 虾夷扇贝微卫星标记的分离及其养殖群体的遗传结构分析. 中国水产科学,16(1): 39—46
- 杨 华,陈 琪,韦朝领等,2011. 茶树转录组中 SSR 位点的信息分析. 安徽农业大学学报,38(6): 882—886
- 张 琼,刘小林,李喜莲等,2010. EST-SSR 分子标记在水生动物遗传研究中的应用. 水产科学,29(5): 302—306
- 赵莹莹,朱晓琛,孙效文等,2006. 虾夷扇贝的多态性微卫星座位. 动物学报,52(1): 229—233
- 徐美佳,张 妍,鲁翠云等,2009. 利用 EST-SSR 座位对鲤鱼 4 种生长性状的单标记回归分析. 水产学杂志,22(1): 15—18
- Aggarwal R K, Hendre P S, Varshney R K *et al*, 2007. Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theor Appl Genet, 114(2): 359—372
- Dreisigacker S, Zhang P, Warburton M L *et al*, 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. Crop Sci, 44(2): 381—388
- Grabherr M G, Haas B J, Yassour M *et al*, 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol, 29(7): 644—652
- He M X, Guan Y Y, Yuan T *et al*, 2008. Realized heritability and response to selection for shell height in the pearl oyster *Pinctada fucata* (Gould). Aquac Res, 39(8): 801—805
- Liu T M, Zhu S Y, Fu L L *et al*, 2013. Development and characterization of 1, 827 expressed Sequence tag-derived

- simple sequence repeat markers for ramie (*Boehmeria nivea* L. Gaud). PLoS One, 8(4): e60346
- Lu X, Wang H X, Liu B Z *et al*, 2013. Three EST-SSR markers associated with QTL for the growth of the clam *Meretrix meretrix* revealed by selective genotyping. Mar Biotechnol, 15(1): 16—25
- Miyazaki T, Goto K, Kobayashi T *et al*, 1999. Mass mortalities associated with a virus disease in Japanese pearl oysters *Pinctada fucata martensii*. Dis Aquat Org, 37(1): 1—12
- Powell W, Machray G C, Provan J, 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci, 1(7): 215—222
- Saha M C, Rouf Mian M A, Eujayl I *et al*, 2004. Tall fescue EST-SSR markers with transferability across several grass species. Theor Appl Genet, 109(4): 783—791
- Sato M, Kawamata K, Zaslavskaya N *et al*, 2005. Development of microsatellite markers for Japanese scallop (*Mizuhopecten yessoensis*) and their application to a population genetic study. Mar Biotechnol, 7(6): 713—728
- Tomaru Y, Kawabata Z, Nakano S, 2001. Mass mortality of the Japanese pearl oyster *Pinctada fucata martensii* in relation to water temperature, chlorophyll *a* and phytoplankton composition. Dis Aquat Org, 44(1): 61—68
- Varshney R K, Graner A, Sorrells M E *et al*, 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol, 23(1): 48—55
- Wang Z L, Wu Z H, Jian J C *et al*, 2009. Cloning and expression of heat shock protein 70 gene in the haemocytes of pearl oyster (*Pinctada fucata*, Gould 1850) responding to bacterial challenge. Fish Shellfish Immunol, 26(4): 639—645
- Wei W L, Qi X Q, Wang L H *et al*, 2011. Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. BMC Genomics, 12: 451

POLYMORPHISM OF EST-SSRs FROM *PINCTADA MARTENSII* BASED ON TRANSCRIPTOME DATASETS

WANG Zhong-Liang^{1,2}, DING Yu^{1,4}, XU You-Hou³, JIAN Ji-Chang^{1,4}, LU Yi-Shan^{1,4},
WANG Bei^{1,4}, CHEN Gang^{1,2}, WU Zao-He^{4,5}

(1. Fisheries College, Guangdong Ocean University, Zhanjiang 524088, China; 2. Key Laboratory of Aquaculture in South China Sea for Aquatic Economic Animal, Regular High Education Institute of Guangdong Province, Zhanjiang 524088, China; 3. Guangxi Key Laboratory of Beibu Gulf Marine Biodiversity Conservation, Qinzhou University, Qinzhou 535000, China; 4. Guangdong Provincial Key Laboratory of Pathogenic Biology and Epidemiology for Aquatic Economic Animals, Zhanjiang 524088, China; 5. Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China)

Abstract To enrich the SSR markers from the pearl oyster *Pinctada martensii*, the EST-SSRs (expressed sequence tag-derived-simple sequence repeats) were developed on a large scale based on the transcriptome sequencing data using program MISA, and polymorphism of SSRs were analyzed and characterized. The results indicate that 9872 EST-SSRs were recognized from 74007 unigenes generated from the transcriptome datasets, accounting for 13.34% of the total unigenes (one EST-SSR per 5102 bp sequence on average). In addition, 132 types of repeat motifs were classified in all EST-SSRs, of which the type of mono-nucleotide repeat was dominant (81.46%), and A/T was the most abundant mono-nucleotide motif (71.27%). Based on the identified SSRs, 17766 pairs of primer sets were designed from 5922 unigenes using program Primer 3. Eighty pairs of primer were randomly selected for PCR amplification to analyze the polymorphism of EST-SSRs, and 62 primers were amplified successfully, among which 17 were polymorphic and the polymorphic rate was 27.42%. These results will provide a useful tool for studies in genetic diversity, genetic map construction and molecular assisted breeding of *P. martensii*, and be of great importance in genetic resource conservation, stock breeding of *P. martensii* and sound development of the pearls culture.

Key words *Pinctada martensii*; simple sequence repeat; transcriptome; molecular marker