

凡纳滨对虾 EST 微卫星标记初步筛选

王艳红, 胡超群*, 张吕平, 夏建军, 任春华

(中国科学院南海海洋研究所, 中国科学院海洋生物资源可持续利用重点实验室,
广东应用海洋生物学重点实验室, 广东 广州 510301)

摘要: 对 161 075 条凡纳滨对虾 dbEST (database of “Expressed Sequence Tags”) 的数据进行 SSR 序列筛选, 搜索参数为二~六核苷酸重复 6 次以上 (包括 6 次), 共获得含 SSR 的 EST 序列 12 600 条, 占整个凡纳滨对虾 dbEST 数据库的 7.8%, 其中二核苷酸重复 10 104 条, 占总 EST-SSR 序列的 80.2%; 三核苷酸重复 2 036 条 (16.1%), 四核苷酸重复 336 条 (2.7%), 五核苷酸重复 35 条 (0.3%), 六核苷酸重复 89 条 (0.7%)。大部分发现的微卫星序列为完美型 (perfect) 的重复序列。在二核苷酸所获得的 SSR 重复单元中, AG/CT 是优势重复单元, 共分布 4 820 条, 占二核苷酸各重复单元的 47.8%; 其次是 AC/GT 重复, 共分布 3 584 条, 占二核苷酸各重复单元的 35.4%; 最后是 AT/TA 和 CG/GC, 分别占二核苷酸各重复单元的 16.6% 和 0.2%。通过对凡纳滨对虾 EST 序列中 SSR 位点信息的发掘分析, 共设计 77 对微卫星引物, 在 10 个个体中成功扩增的引物有 54 对, 成功扩增率为 70.1%。其中多态性丰富的引物 28 对, 多态率为 51.9%。等位基因数为 2~5, PCR 产物大小为 140~600 bp。本研究筛选的 EST-SSR 标记将为凡纳滨对虾的分子遗传学研究、种质鉴定等提供可靠的工具。

关键词: 凡纳滨对虾; 表达序列标签; 微卫星; 分子标记

中图分类号: Q 958.1; S 917

文献标志码: A

凡纳滨对虾 (*Litopenaeus vannamei*), 俗称南美白对虾、太平洋白对虾, 自然分布于秘鲁至墨西哥太平洋沿岸。它具有生长快、营养要求低、出肉率高、环境适应力强等优点, 是世界上三大养殖对虾中单产量最高的虾种。目前, 凡纳滨对虾是我国最主要的养殖对虾品种^[1], 在对虾渔业及养殖业中占有重要地位。

微卫星分子标记 (microsatellite) 是指以 2~6 个核苷酸为基本重复单元的多次串联重复的 DNA 序列, 亦称简单序列重复 (simple sequence repeats, SSR)^[2]。微卫星标记与其他分子标记技术相比有以下优点: 多态性高、符合孟德尔遗传模式、共显性遗传等优势, 因此, 在种群遗传结构分析、连锁图谱构建以及系谱认证等工作中, 微卫星标记技术得到了广泛的应用。

表达序列标签 (expressed sequence tag, ESTs), 是指通过对不同组织 cDNA 文库测序得到的 DNA 序列^[3], EST 基本能反映 mRNA 的信息, 在功能基因的筛选和遗传图谱构建等工作中具有重要意义。大量的 ESTs 数据为 SSR 标记的开发提供了一个十分便捷的来源, 既节省了传统开发 SSR 所需要的人力物力, 又为功能基因提供信息。从 ESTs 数据库中筛选微卫星标记的工作, 在多种水产物种中已见报道, 如太平洋牡蛎 (*Crassostrea gigas*)^[4] 等, 在对虾属的研究包括斑节对虾 (*Penaeus mortodort*)^[5-6] 和中国明对虾 (*Fenneropenaeus chinensis*)^[7] 等。凡纳滨对虾 SSR 研究领域, 已开发出大量来自基因组的微卫星标记^[8-10], PEREZ 等^[11] 报道了关于凡纳滨对虾 EST 微卫星标记筛选和应用, 而这些数据来自

收稿日期: 2011-01-13 修回日期: 2011-03-28

资助项目: 国家“八六三”高技术研究发展计划 (2010AA10A401); 广东省海洋渔业科技推广专项 (A200899A02 & A200901B03); 广东省/中科院省院合作项目 (2009B091300088); 中国科学院知识创新工程重要方向性项目 (KSC XZ-YW-N-4701)

通讯作者: 胡超群, E-mail: cqhu@scsio.ac.cn

于(<http://www.marinegenomics.org>),并没有从NCBI上的大量数据中筛选EST-SSR标记的报道。

国内多家相关机构正在开展该对虾新品种的遗传育种与改良,并取得阶段性成果^[12],因此对于可靠的EST-SSR标记的需求十分迫切。为了更进一步研究凡纳滨对虾种群结构特征、结合经济性状定位以及构建更加精细的遗传连锁图谱,迫切需要开发大量的微卫星标记。本研究从凡纳滨对虾EST数据库中进行了微卫星序列筛选,并进一步分析EST微卫星的特征,开发出候选EST-SSR标记77个,为开发新型的凡纳滨对虾EST-SSR标记奠定理论基础。

1 材料与方法

1.1 凡纳滨对虾EST来源

截至2009年12月,GenBank上已经公布了16万多条凡纳滨对虾的EST序列,从NCBI dbEST数据库(www.ncbi.nlm.nih.gov/dbEST/index.html)批量下载所有凡纳滨对虾EST序列。保存成FASTA格式备用。

1.2 EST-SSR的发掘

本研究利用The Simple Sequence Repeat Identification Tool(SSRIT)^[13]软件进行微卫星序列分析(<http://www.gramene.org/db/markers.ssrtool>),搜索微卫星设定参数是:二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸重复次数分别为6次以上(包括6次),包括复合型微卫星。微卫星的重复次数越多,相应的检验出等位基因数目就越多。

1.3 EST-SSR微卫星引物设计

从获得的微卫星序列中,优先选取重复次数在16次以上的二~六核苷酸碱基重复序列,使用引物设计软件Primer Premier 5.0(<http://www.premierbiosoft.com/>)进行微卫星引物的设计。引物设计的主要参数为:引物长度为18~24 bp;产物长度为100~300 bp;引物复性温度为45~65℃,GC含量为40%~70%,最适为50%。选出最优的引物对,由上海生工生物技术公司合成。

1.4 样品来源及DNA提取

凡纳滨对虾样品购自广州市黄沙水产市场,共10个个体,取肌肉组织0.1 g保存在95%的乙醇中。蒸馏水漂洗两遍,剪碎后用海洋动物组织

基因组DNA提取试剂盒[天根生化科技(北京)有限公司]提取基因组DNA,于室温干燥,0.1×TE溶解,电泳检测后备用。

1.5 PCR扩增与电泳检测和筛选

PCR扩增总体积为25 μL,包括模板DNA 50~100 ng,10×PCR混合缓冲液2.5 μL,10 mmol/L dNTP混合液0.4 μL,10 μmol/L上下游引物各0.4 μL,5.0 U的Taq DNA聚合酶[天根生化科技(北京)有限公司]0.2 μL,加ddH₂O补足体系。扩增反应均在PTC-2000型PCR仪(Bio-Rad Laboratories,USA)上完成。反应程序为94℃预变性2 min;94℃变性30 s,退火温度退火30 s(退火温度依引物而定),72℃延伸30 s,35个循环;最后72℃延伸10 min。扩增产物采用2%的琼脂糖凝胶电泳检测。

2 结果与分析

2.1 EST微卫星序列筛选

凡纳滨对虾EST数据库包含161 075条序列,用SSRIT搜索SSR,结果表明,共获得含SSR的EST序列12 600条(7.8%),其中包括二核苷酸重复10 104条,三核苷酸重复2 036条,四核苷酸重复336条,五核苷酸重复35条,六核苷酸重复89条,二核苷酸重复序列是最丰富的重复单元,占80.2%,其次为三核苷酸重复和四核苷酸重复分别占16.1%和2.7%,六核苷酸重复占0.7%五核苷酸最少只有0.3%(表1)。

表1 凡纳滨对虾SSR在EST中出现的频率

Tab.1 The frequency of *L. vannamei* SSR in the EST

类型 type	数目 number	各类型比例/% rate
二核苷酸 dinucleotide	10 104	80.2
三核苷酸 trinucleotide	2 036	16.1
四核苷酸 tetranucleotide	336	2.7
五核苷酸 pentanucleotide	35	0.3
六核苷酸 hexanucleotide	89	0.7
合计 total	12 600	100.00

在二核苷酸所获得的SSR重复单元中,AG/CT重复单元是优势重复序列,共分布4 820条,占二核苷酸各重复单元的47.8%,其次是AC/GT重复,共分布3 584条,占二核苷酸各重复单元的35.4%,AT/TA和CG/GC,分别占二核苷酸

各重复单元的 16.6% 和 0.2% (表 2)。

在三核苷酸中,最为丰富的 3 种重复单元是 AAT/TTA/ATT/ATA/TAT/TAA, 共分布 576 条,占三核苷酸重复中各重复单元的 28.3% (表 3); AAG/AGA/GAA/CTT/TTC 共分布有 378 条,占总重复单元的 18.6%; ACT/CTA/TCA/ATG/TGA/GAT 筛选出 276 条,占三核苷酸总重复单元的 13.6%; 还包括有七种重复基元,分别是 AGG/GAG/GGA/CCT/CTC/TCC 分布有 197 条,占三核苷酸总重复单元的 9.7%; AGT/GTA/TAG/ATC/CAT/TCA 筛选出 182 条,占 8.9%; CGA/GAC/CTG 筛选出 167 条,占 8.2%; ACG/CAG/GCA/AGC/TCG/CGT 有 129 条,占总重复单元的 6.3%; 剩余 3 种核苷酸重复单元总共有 131 条序列,其中 GCC/GGC 只有 5 条,占 0.2%。

表 2 二核苷酸重复中各重复单元的数量及比例
Tab.2 Numbers and proportions of all the repeat motifs in the dinucleotide repeats

类型 type	数目 number	各类型比例/% rate
AG/CT	4 820	47.8
AC/GT	3 584	35.4
AT/TA	1 673	16.6
CG/GC	27	0.2
合计 total	10 104	100

表 3 三核苷酸重复中各重复单元的数量及比例
Tab.3 Numbers and proportions of all the repeat motifs in the trinucleotide repeats

类型 type	数目 number	各类型比例/% rate
AAT/TTA/ATT/ATA/TAT/TAA	576	28.3
AAG/AGA/GAA/CTT/TTC	378	18.6
ACT/CTA/TAC/ATG/TGA/GAT	276	13.6
AGG/GAG/GGA/CCT/CTC/TCC	197	9.7
AGT/GTA/TAG/ATC/CAT/TCA	182	8.9
CGA/GAC/CTG	167	8.2
ACG/GAC/CGA/ACG/TGC/GCT	129	6.3
ACC/CAC/CCA/GGT/GTG/TGG	65	3.2
ACA/AAC/CAA/GTT/TTG/TGT	61	3.0
GCC/GGC	5	0.2
合计 total	2 036	100

2.2 EST-SSR 引物设计及多态性引物的筛选

经过筛选,有 1 187 条 EST 序列所包含的微卫星核心重复序列在 16 bp 以上(包含 16 bp),通过 Primer Premier 5.0 进行引物设计,共设计 77

对候选 EST-SSR 标记,其余 EST 序列中没有找到合适的引物结合位点。77 对引物在 10 个个体中的扩增结果为 23 对引物没有扩增产物,其余 54 对引物获得预期的目标条带,成功扩增率为 70.1%,28 对在所检测的 10 个个体中具有不同程度的多态性,多态比例为 51.9%。等位基因数为 2~5,PCR 产物大小为 140~600 bp。

3 讨论

EST 中 SSR 的含量虽然在各个物种之间有所不同,但基本分布在 1%~5%^[2]。5 种主要农作物[大麦(*Hordeum vulgare* Linn)、玉米(*Zea mays* L)、水稻(*Oryza Sativa* Linnaeus)、高粱(*Sorghum vulgare*)和小麦(*Triticum aestivum* L)]ESTs 中微卫星的含量在 1.5%~4.7%^[14]。而在水产动物太平洋牡蛎 ESTs 中的微卫星的比例是 4.5%^[4],栉孔扇贝(*Chlamys farreri*) 0.6%^[14],而沟鲂(*Ictalurus punctatus*)则达到 4.66%^[16],在对虾研究中,中国明对虾 EST 中 SSR 的比例为 2.19%^[7],斑节对虾 EST 中 SSR 的比例高达 9.8%^[6],前人研究凡纳滨对虾 EST 序列中 SSR 所占比例是 4.9%^[11]。本研究从凡纳滨对虾 161 075 条 EST 序列中查找获得 12 600 条微卫星序列,占整个 ESTs 数据库的 7.8%,这个结果和斑节对虾相似。这和微卫星筛选时的标准有关,本研究标准是:二~六核苷酸重复次数分别为 6 次以上(这样的筛选标准对于二核苷酸重复的筛选稍低,而对于五~六核苷酸重复本筛选标准稍高。由于考虑到凡纳滨对虾 dbEST 数据库信息量很大,因此统一了筛选标准)。EST 序列中 SSR 的比例也有可能和数据库丰富程度也有关系^[17],本研究中凡纳滨对虾数据库有 16 万多条 EST,因此凡纳滨对虾 EST 中 SSR 比例较高。

凡纳滨对虾 EST 微卫星中序列的比例差异比较大,其中二核苷酸重复所占比例最高,其次是三核苷酸重复,它们分别占总 SSR 的 80.2% 和 16.1%,和中国明对虾、斑节对虾中的研究结果相似,凡纳滨对虾也是以二核苷酸重复为主导,凡纳滨对虾中的六核苷酸重复所占比例比是 0.7%,比五核苷酸重复所占比例高出 0.4%,这有可能和三个碱基翻译有关系,具体的原因有待进一步的研究。这个结果和马氏珠母贝的研究类似^[18]。

在重复形式上,凡纳滨对虾双碱基重复微卫

星最多的类型是 (AG/CT)_n, 约占总位点数的 47.8%, 这与人类和其他脊椎动物及秀丽隐杆线虫的结果相同^[14], 条斑紫菜、水稻和玉米等生物中双碱基的 AG/CT 形式的微卫星序列在各自类重复型中占了绝大多数^[19]。但是中国明对虾基因组^[20-21]、斑节对虾^[5-6]、马氏珠母贝^[17] EST 中的研究表明, 双碱基重复微卫星最多是 AT/TA, 分别占到 42.44% 和 29.1%。在凡纳滨对虾双碱基重复微卫星最少的类型是 (GC/CG)_n 有 27 条占总位点数的 0.2%, 而在模式植物拟南芥、水稻中均未见 GC 重复的报道, 同时在水产动物中国明对虾和马氏珠母贝中也没有 GC 重复的报道, 凡纳滨对虾中 0.2% 的 GC 重复单元的出现, 在水产动物中 GC 虽然出现的频率很低, 但也不缺乏这种重复序列的存在, 推测和凡纳滨对虾 EST 数据库庞大有关。凡纳滨对虾的三核苷酸中最为丰富的 3 种重复单元是 AAT/TTA, ACT/CTA 和 AAG/AGA 这 3 种类型共占总数 60.50% 的这与中国对虾和马氏珠母贝中最为丰富的类型

相似。

引物筛选是分子标记开发的重要过程, 本研究通过 77 个候选 EST-SSR 标记, 共筛选到 28 对多态性丰富的标记, 三核苷酸重复的 EST 序列中 EST-SSR 标记的开发会是下一步研究重点, 期望在其中筛选出和功能相关联的 EST-SSR 标记, 用于遗传图谱构建和相关研究。本研究获得的引物, 等位基因数分布于 2~5, 至于微卫星座位的等位基因数不高是否和样本量较少有关, 还需要进一步的验证。

凡纳滨对虾是我国最主要的养殖品种, 养殖年产量达到 100 万 t, 但是海洋动物分子生物学的研究起步较晚, 利用 GenBank 数据分析 EST-SSR 标记在凡纳滨对虾研究中尚未见报道, 本文通过大规模分析 NCBI 中的 EST 序列, 筛选得到了 77 个候选 EST-SSR 标记, 其中多态性丰富的 EST-SSR 标记有 28 对, 将为凡纳滨对虾遗传图谱构建、QTL 定位、分子标记辅助育种等工作提供有效的分子标记。

表 4 凡纳滨对虾微卫星标记成功扩增引物序列及其参数

Tab. 4 Characters and sequences of some screened microsatellite primers in *L. vannamei*

位点 locus	核心重 复序列 repeat motif	序列 (5'-3') primer sequence	预期 大小/bp expect size	实际 大小/bp observed size	退火 温度/°C annealing temperature	是否 多态 polymorphism	等位 基因数 allele number	GenBank 登录号 accession no.
<i>LvE2F</i>	(TG) ₆₂	F: TTTAGCAACACTTTTCCTT R: GGTCAAGTCCCACCAGA	361	350	46	P	5	FE158490
<i>LvE4F</i>	(TG) ₄₇	F: ATGAAGTGGCTCGTGTAAG R: ATTTGGTGGATTTAGGGA	263	260	45	P	4	FE152424
<i>LvE6F</i>	(CA) ₄₇	F: CACACGCACACGCAAACACA R: GCTAATCGAAATGAAACATG	267	263	46	P	5	FE122799
<i>LvE8F</i>	(CAGA) ₄₁	F: GGCAGCAAGAGTTCCAT R: GGTCTGAATGCTTCCTGT	299	288	50	P	4	FE069689
<i>LvE10F</i>	(CA) ₄₀	F: CACAGAGGATAGACGCAAGG R: TTCCATCAACTCGGGTGC	411	408	46	-		FE186821
<i>LvE11F</i>	(GT) ₄₂	F: CGGGGACGATCACAACAC R: CGCACGCACCCACGCACAAACA	224	223	48	P	3	FE054484
<i>LvE13F</i>	(AG) ₃₉	F: GAACCGTGCGAAAAGTCCC R: CAGCGACACGACAGAAGGA	376	320	45	P	3	FE124653
<i>LvE14F</i>	(CA) ₃₉ (CACT) ₂₀	F: TTTACTACAGGCAACAAT R: GTATGGGAGTGTAGTGC	445	440	60	-		FE085408
<i>LvE15F</i>	(AG) ₃₉	F: TTTGCCTTATTGGTTGCG R: CAGCGACACGACAGAAGGA	607	600	50	-		FE125173
<i>LvE16F</i>	(CA) ₃₉	F: GTTGAGGGAAGGTTGTGC R: CATAATAAGCAATGCCAAGA	385	380	56	P	3	FE112827
<i>LvE17F</i>	(TG) ₃₉	F: GCCAATTACGATTGTTT R: CAACATCCCCTCACAAAC	256	256	55	P	3	FE079795

续表 4

位点 locus	核心重 复序列 repeat motif	序列 (5'-3') primer sequence	预期 大小/bp expect size	实际 大小/bp observed size	退火 温度/°C annealing temperature	是否 多态 polymorphism	等位 基因数 allele number	GenBank 登录号 accession no.
LvE22F	(CA) ₃₂	F:GCACTTAGGGTTAGGTCA R:TAGCTTCTCGGCTACAAA	149	149	46	P	2	FE123558
LvE23F	(AG) ₃₁	F:TGTAAAGTAGATAGCCAAGA R:TCTGAATCCTCCACCTT	247	247	56	-		FE069314
LvE24F	(AG) ₃₀	F:AGAGGGAGGGAGGTAAAG R:TCAGTTCAGTATTGCCAAGGAGG	140	140	45	-		FE052561
LvE26F	(GA) ₄₉	F:ATTATTACAAGTCCGAGTG R:CTGAAGATGAGGGAGACG	397	397	46	P	5	FE110384
LvE27F	(GT) ₃₀	F:TAGTTTGTGCTGTCGTGGAT R:ATGAATCATAACCGCCACA	401	400	45	-		FE152778
LvE29F	(TAT) ₁₇	F:GTCAGTGTGGGAAATGG R:GGTAACAATGATGGCAAG	483	450	46	-		FE046363
LvE32F	(AG) ₂₉	F:CAGAAGTAGGCGATCAGA R:AGTATTGCCAAGGAGGTA	245	250	50	-		FE120471
LvE33F	(TC) ₂₉	F:ACATTTCCGTTAGGTCTG R:AGAAGATGAAGACGAAGA	265	265	45	-		FE124817
LvE34F	(CA) ₂₉	F:TTTTACCAATCTCCACG R:GATACATTCCTTGCTCCC	159	150	50	-		FE060358
LvE35F	(GA) ₁₆ · (AG) ₂₉	F:AGCGAGTCAGTGAGCGAGGAA R:CGCCAAACGAATGGTCTA	387	360	46	P	4	GT112971
LvE39F	(ATCT) ₂₆	F:CGTATCCTTCGCCTTCTC R:GGAAAGATGGCGTGACAT	265	250	46	P	3	FE046762
LvE40F	(CT) ₂₆	F:ACAGTCACCGTCGAAGTA R:AGGGAGGCTATAAAGGAT	185	160	46	P	2	FE093957
LvE41F	(CT) ₂₆	F:ACCCTTGCCTCATTTCGC R:CAGGCGCAAGAGTGGA	260	250	46	P	2	FE064041
LvE42F	(GAA) ₂₅	F:TCAAGAGCAACCGCAGAA R:CCAGGTCTTGGCATGGAG	433	400	45	P	2	FE148403
LvE43F	(CT) ₂₅	F:GGACCTCTGCTGTTCTA R:CCTTCTTCGCCAGTTCAT	213	200	45	P	3	FE152890
LvE47F	(GA) ₃₃	F:GTCAGTGAGCGAGGAAGC R:CGCCAAACGAATGGTCTA	406	380	50	P	3	GT112501
LvE48F	(AC) ₂₂	F:ACCGTTTGTGTTTCGTT R:TTCTCCCTCGCTTTGTAG	286	250	52	-		FE117403
LvE49F	(CA) ₂₂	F:TTTCCGAAGGCATAAACA R:CAGGCACTCATTCAAAGG	113	100	45	-		FE187271
LvE53F	(GA) ₂₂ · (GA) ₇	F:GTCAGTGAGCGAGGAAGC R:CGCCAAACGAATGGTCTA	305	300	51	P	4	GT113042
LvE54F	(AG) ₂₁	F:CTAAACCGACAGATACAA R:TAAGTCTTGTTTCAGTGGC	282	250	50	P	3	FE117290
LvE55F	(GT) ₁₂ · (GT) ₂₅	F:CTCGCTCGTCTGTTTCAC R:GAGTAACTGGGTCATCGT	289	260	48	P	3	FE080567
LvE57	(CT) ₂₀	F:TTTCCCTTTCGCATTCTCT R:ATGATTCCATCAGGTTTCG	286	260	53	-		FE079693
LvE58	(AC) ₁₉	F:CCCTTTGATAAGTGGTAA R:CTGGTACTTCCGTTTGTC	427	400	50	-		FE091877
LvE60	(GT) ₁₉	F:CGAGTCAGTGAGCGAGGAA R:GCAAGGAATCAGGCACAA	333	300	46	P	2	GT112848

续表 4

位点 locus	核心重 复序列 repeat motif	序列 (5'-3') primer sequence	预期 大小/bp expect size	实际 大小/bp observed size	退火 温度/°C annealing temperature	是否 多态 polymorphism	等位 基因数 allele number	GenBank 登录号 accession no.
LvE61	(GA) ₁₉	F:ATAAAGAAGCGAGAACGA R:CTATGGCTAGATCCGAGA	255	250	46	-		CK591498
LvE63	(AC) ₁₈	F:ACCGTTGCTTTGACCTCT R:GTGAGACCGCGTGTAGAT	272	260	50	-		FE191544
LvE64	(TG) ₁₈	F:TGACAATACTGGGAGACT R:ACATAGGAAAGCATAACAG	273	270	49	-		FE191545
LvE65	(AAT) ₁₈	F:GCTGGTTGTGACTCTAA R:CCTTGCAATGATCTGTCA	274	200	46	P	2	FE045783
LvE66	(TC) ₁₇	F:GATCACCACCATTCTTTAGCCA R:CGCTGTATGGGTCATCTGC	249	250	53	P	3	FE179615
LvE67	(AG) ₁₈	F:TTTCCTGGACAACCTCAC R:TCCCATCTTACGTTCTACT	150	150	55	P	3	CK591514
LvE70	(GA) ₁₇	F:GATGATGAGTGCCGAAGTGC R:ATTGAGCGTCGAAGTTTT	341	300	46	-		CK591492
LvE71	(TC) ₁₇	F:TTCCTCCTTGCACGATTT R:TGACTGCCCTTTCTACCA	252	300	46	P	4	CK591516
LvE72	(CA) ₁₇	F:ATCCGTCGTGAAGTTAGG R:TGTTTGTGCTTATGTGGG	208	200	47	P	3	FE153896
LvE74	(CA) ₁₆	F:GCTTCGTTGCACTGACT R:GACTGGTACTTCCGTTTG	254	250	48	P	2	FE151887
LvE75	(AG) ₁₆	F:GTTTGACTGGTACTTCCGTTTG R:CGATTGGCGAGGATTTGC	222	200	52	-		FE078051
LvE77	(TC) ₁₆	F:TTCCTACCTTCCACCTC R:GTCCACCTTCTCGCTCT	292	250	46	-		FE188370

注:F. 正向引物;R. 反向引物;P. 多态引物;- . 无多态引物。

Notes:F. forward primer;R. reverse primer;P. polymorphism;- . non-polymorphism.

参考文献:

- [1] 农业部渔业局. 中国渔业年鉴 2009 [M]. 北京: 中国农业出版社, 2009.
- [2] ILATEY F, TOSSER K G, CLOUSCARD M C, *et al.* Expressed sequenced tags for genes: a review [J]. *Genetics Selection Evolution*, 1998, 30 (6): 521 - 541.
- [3] 骆蒙, 贾继增. 植物基因组表达序列标签 (EST) 计划研究进展 [J]. *生物化学与生物物理进展*, 2001, 28 (4): 494 - 497.
- [4] WANG Y H, REN R, YU Z N. Bioinformatic mining of EST-SSR loci in the Pacific oyster, *Crassostrea gigas* [J]. *Animal Genetics*, 2008, 39 (3): 287 - 289.
- [5] TONG J, LEHNERT S A, BYRNE K, *et al.* Development of polymorphic EST markers in *Penaeus monodon*: applications in penaeid genetics [J]. *Aquaculture*, 2002, 208: 69 - 79.
- [6] MANEERUTTANARUNGRGJ C, PONGSOMBOON S, WUTHISUTHIMETHAVEE S, *et al.* Development of polymorphic expressed sequence tag derived microsatellites for the extension of the genetic linkage map of the black tiger shrimp (*Penaeus monodon*) [J]. *Animal Genetics*, 2006, 37 (4): 363 - 368.
- [7] WANG H X, LI F H, XIANG J H. Polymorphic EST SSR markers and their mode of inheritance in *Fenneropenaeus chinensis* [J]. *Aquaculture*, 2005, 249 (1 - 4): 107 - 114.
- [8] CRUZ P, MEJIA C H, PEREZ R, *et al.* Isolation and characterization of microsatellites in Pacific white shrimp *Penaeus (Litopenaeus) vannamei* [J]. *Molecular Ecology Notes*, 2002, 2 (3): 239 - 241.
- [9] BALL A O, LEONARD S, CHAPMAN R W. Characterization of (GT)_n microsatellites from native white shrimp (*Penaeus setiferus*) [J]. *Molecular Ecology*, 1998, 7 (9): 1251 - 1253.
- [10] MEEHAN D, XU Z, ZUNIGA G, *et al.* High frequency and large number of polymorphic microsatellites in cultured shrimp, *Penaeus*

- (*Litopenaeus vannamei*) [Crustacea:Decapoda] [J]. *Marin Biotechnology*, 2003, 5:311-330.
- [11] PEREZ F, ORTIZ J, ZHINAULA M, *et al.* Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei*, *Litopenaeus stylirostris*, and *Trachypenaeus birdy* [J]. *Marin Biotechnology*, 2005, 7:554-569
- [12] 张吕平, 吴立峰, 沈琪, 等. 凡纳滨对虾 *Litopenaeus vannamei* 全同胞家系的建立及生长比较[J]. *水产学报*, 2009, 33(6):942-949.
- [13] TEMNYKH S, DECLERK G, LUKASHOVA A, *et al.* Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential [J]. *Genome Research*, 2001, 11: 1441-1452.
- [14] KANTETY R V, ROTA M L, MATTHEWS D E, *et al.* Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat [J]. *Plant Molecular Biology*, 2002, 48:501-510.
- [15] 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, 13(12): 72-75.
- [16] SERAPION J, KUCUKTAS H, FENG J, *et al.* Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*) [J]. *Marine Biotechnology*, 2006, 4: 364-377.
- [17] GABOR T, ZOLTAN G, JERZY J. Microsatellites in different eukaryotic genomes: survey and analysis [J]. *Genome Research*, 2000, 10(7):967-981.
- [18] 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.
- [19] 刘必谦, 曾庆国, 骆其君, 等. 条斑紫菜 (*Porphyraezoensis*) db EST 中筛选微卫星位点及引物种间转移扩增 [J]. *海洋与湖沼*, 2005, 36(3):248-254.
- [20] 高焕, 刘萍, 孟宪红, 等. 中国对虾 (*Fenneropenaeus chinensis*) 基因组微卫星特征分析 [J]. *海洋与湖沼*, 2004, 35(5):424-431.
- [21] 徐鹏, 周令华, 相建海. 中国对虾微卫星 DNA 的筛选 [J]. *海洋与湖沼*, 2001, 32(3):255-259.

A preliminary study on microsatellite markers screening from EST sequences of *Litopenaeus vannamei*

WANG Yan-hong, HU Chao-qun*, ZHANG Lü-ping, XIA Jian-jun, REN Chun-hua

(Key Laboratory of Marine Bio-resources Sustainable Utilization, Guangdong Key Laboratory of Applied Marine Biology, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China)

Abstract: EST-SSR marker development by screening and mining the SSR repeats in the EST database of *Litopenaeus vannamei*, 161 075 EST sequences were downloaded from EST sequences in the database of GenBank. EST sequences were analyzed and the SSR was screened with SSIRT (<http://www.gramene.org/db/markers.ssrtoo>). As a result, 12 600 (7.8%) SSRs were identified from the EST resources, among which there were 10 104 (80.2%) dinucleotide, 2 036 (16.1%) trinucleotide, 336 (2.7%) tetranucleotide, 35 (0.3%) pentanucleotide and 89 (0.7%) hexanucleotide SSRs. Among the dinucleotide sequences, AG/CT repeat motif accounted for 4 820 (47.8%), 3 584 (35.4%) AC/GT repeat motif, AT/TA, CG/GC repeat motif was 1 673 (6.6%), 27 (0.2%) respectively. 77 pairs of primers were designed with Primer Premier 5.0. Fifty four primer pairs obtained expected product, with the success rate of 70.1%. And 28 of them showed polymorphism among 10 individuals, polymorphism rate was 51.9%. Number of alleles ranged from 2 to 5, and PCR product size was 140–600 bp. This implies that the polymorphism EST-SSR developed from NCBI could work in genetic mapping, gene clone and gene functional research.

Key words: *Litopenaeus vannamei*; expressed sequence tag; microsatellite; molecular marker

Corresponding author: HU Chao-qun. E-mail: cqhu@scsio.ac.cn