

应用高分辨率熔解曲线(HRM)开发大菱鲂 (*Scophthalmus maximus*)SNP 标记的研究*

刘庆明^{1,2} 李 猛^{1,2} 马爱军¹ 王新安¹ 黄智慧¹ 郭建丽¹

(1. 中国水产科学研究院黄海水产研究所 农业部海洋渔业可持续发展重点实验室 青岛市海水鱼类种子工程与生物技术重点实验室 青岛 266071; 2. 上海海洋大学 上海 201306)

提要 利用 Vector NTI Advance 11 软件分析了 NCBI 数据库中大菱鲂的 12428 条 EST 序列, 筛选出候选 SNP 位点; 应用高分辨率熔解曲线(HRM)对大菱鲂不同性状群体进行基因分型。结果表明, 522 个重叠群中共筛选出 160 多个候选 SNP 位点; 设计 56 对引物用于实验, 能扩增出目的片段的引物有 37 条, 合 45 个位点, 成功率为 66.1%; 设计探针, 能与候选 SNP 位点杂交的探针有 13 条, 杂交率为 28.9%。本实验中能成功进行基因分型的位点共有 8 个, 占杂交位点的 61.5%, 其中有 6 个为碱基替换型位点, 即 G-A 和 C-T 各占 3 个, 2 个为碱基颠换型位点, 即 T-G 和 A-T 各占 1 个。

关键词 单核苷酸多态性标记(SNP), 大菱鲂, 高分辨率熔解曲线分析, EST 序列
中图分类号 Q953

大菱鲂(*Scophthalmus maximus*)隶属于鲆科、菱鲆属, 英文名 Turbot, 又称“多宝鱼”, 是一种欧洲沿海特有的比目鱼, 属冷温性鱼类, 为世界公认的优质鱼种(雷霖霖, 2003)。目前大菱鲂已发展成为我国北方沿海工厂化养殖业的主导品种之一(雷霖霖, 2005)。随着养殖的发展, 累代养殖和近亲交配造成种质退化现象比较严重, 对大菱鲂进行遗传改良, 培育出具有优质、高产、抗逆等优良性状的新品种, 已成为发展中亟待解决的问题(雷霖霖, 2005; 邹曙明等, 2001)。本实验室自 2007 年开始采用以数量遗传学为基础的大规模家系选育技术对大菱鲂进行了遗传改良工作, 年均构建大菱鲂家系 20 到 60 个不等, 对其遗传参数进行分析, 为大菱鲂选育工作提供必要的基础依据和技术参数(马爱军等, 2008, 2009, 2010; 王新安等, 2010)。本实验室还利用家系材料开发了大菱鲂 SSR 和 SNP 等分子标记, 目的是加强对大菱鲂遗传多样性、近亲繁殖、种类和品系鉴定以及遗传连锁图谱建立的研究, 加快对数量性状位点(QTL)基因的

鉴定作为分子标记辅助选择(MAS)的研究(侯仕营等, 2011; 许可等, 2009; HOU *et al.*, 2011)。

SNP 标记是基因组中最丰富的突变类型, 它是遗传育种研究中的热门领域。Thomas 等(2008)基于 EST 开发了 304 个 SNP 标记, 构建了大西洋鲑雌性和雄性的 SNP 连锁图谱。然而 SNP 标记的开发通常需要在特定位点存在等位基因突变的已知序列信息(Schlotterer, 2004)。因此, 大量的 SNP 标记开发通常需要花费大量的时间和金钱, 特别对那些基因序列信息有限的非模式生物更是明显。目前开发大菱鲂 SNP 标记的工作才刚刚起步, 这一方面的研究尚未见报道, 具有很高的研究价值和广阔的研究空间。

近年来, SNP 检测方法和相关仪器的研发进展很快。其中, HRM 是通过 PCR 产物的熔解曲线分析检测 PCR 片段序列的微小差异的。溶解曲线的采集主要依靠饱和染料, 如 LC Green 等(Carl *et al.*, 2003)。HRM 技术主要是根据核酸分子熔解温度和熔解形状, 通过专业的分析软件对未知样品进行区分。HRM 非

* 现代农业产业技术体系专项资金(CARS-50-G01)资助; 国家 863 计划(2012AA10A408-8)资助。刘庆明, 硕士研究生, E-mail: liuqingming2@163.com

通讯作者: 马爱军, 博士, 研究员, 博士生导师, E-mail: maaj@ysfri.ac.cn

收稿日期: 2011-12-15, 收修改稿日期: 2012-03-21

标记探针基因分型技术(unlabeled probe genotyping)可以快速准确地对 SNP 等突变类型进行基因分型(Zhou *et al.*, 2004), 该技术的基本原理与筛查未知 SNP 的原理区别仅在于前者检测的熔解曲线是探针与目标 DNA 单链杂交的熔解曲线, 而后者检测的是目的片段。由于探针长度较短, 进一步提高了区分不同样品之间的分辨率(丰贵鹏等, 2010)。目前, 利用 HRM 进行突变扫描的应用非常广泛, 绝大多数应用都集中在医学领域, 科学家们发现了许多由于基因突变而导致疾病发生的现象。目前, 国内外运用 HRM 技术研究和分析水产动物 SNP 标记刚刚起步, 还有大量工作需要开展。Jiang 等(2011)利用 HRM 开发了 44 个栉孔扇贝的 SNP 分子标记, 这些位点在包含来自于 5 个自然群体的 48 个个体的群体中存在多态性。

本文中利用 Vector NTI Advance 11 软件分析 NCBI 数据库中大菱鲆的 12428 条 EST 序列, 筛出候选 SNP 位点, 再应用 HRM 对大菱鲆不同性状群体进行基因分型。确认 SNP 位点。这些 SNP 标记在大菱鲆高密度遗传图谱构建、性状作图、基因的精确定位与克隆、群体遗传结构分析以及系统发育分析等方面均具有广阔的应用前景, 这将为今后的大菱鲆的遗传改良, 培育出具有优质、高产、抗逆等优良性状的新品种研究奠定基础。

1 材料与方法

1.1 大菱鲆候选 SNP 位点的获得

1.1.1 大菱鲆 EST 序列的获得 在美国国立生物技术信息中心(NCBI)的主页, 以 *Psetta maxima* 为关键词搜索大菱鲆的 EST 数据库, 选择 FASTA 格式, 采用批量下载模式, 将全部序列保存至单一文本文档。

1.1.2 EST 序列的聚类、拼接 采用 Vector NTI Advance11 综合软件包中的 Contig Express 模块对大菱鲆 EST 序列进行聚类、拼接, 从而使符合聚类条件的若干个 EST 合并成重叠群。

1.1.3 筛选 SNP 的标准 利用拼接后的 EST 重叠群通过人工判断进行候选 SNP 的筛查。由于客观上测序错误的存在和 EST 序列数量有限的条件, 因而只在含有不少于 4 条 EST 的重叠群中寻找候选 SNP; 选择含有 SNP 位点的重叠群; 去除插入缺失型 SNP; 去除 EST 边缘多态性位点, 即去除候选 SNP 位于 reads 的最前或最后 100 个碱基的序列; 去除模糊位点; 去除候选 SNP 位点左右各 5 个碱基范围内还存在

其它差异碱基的位点; 当一种突变碱基在所在 EST 序列突变位点碱基中出现的频率大于 20%以上时, 则认为是一个候选 SNP 位点, 可以进行下一步的检测和验证; 候选 SNP 中含差异碱基 reads 数不应小于 2 等(张新叶等, 2009; 王绍宗等, 2010)。

1.2 大菱鲆候选 SNP 的验证

1.2.1 DNA 模版的制备 本实验样品来源于烟台天源水产有限公司和青岛智达海洋科技有限公司, 运用海洋动物组织基因组 DNA 提取试剂盒(离心柱型, TIANGEN)对大菱鲆 8 个家系的 10 个不同性状群体的 42 个样品提取 DNA, 样品详细资料见表 1。

表 1 开发大菱鲆 SNP 标记的样品信息
Tab.1 Sample information of developing SNP marker in turbot

序号	性状	数量	取样方式
1	体色带点、生长快	4	随机
2	白化	4	随机
3	体色带点、缺鳃盖	4	随机
4	体色不带点、生长快	4	随机
5	生长慢	4	随机
6	免疫差、生长慢	4	随机
7	免疫好、生长快	4	随机
8	免疫差、生长慢	4	随机
9	冰岛野生	7	随机
10	法国野生	3	随机

1.2.2 引物及探针设计 通过比对、分析 EST 序列, 在得到的候选 SNP 位点两端设计引物, 要求目的片段长度小于 300bp, 引物设计用 Primer5.0, 参数分析用 Oligo7.0。根据候选 SNP 位点设计探针, 要求探针的退火温度为 60—80℃, 探针 3 端采用 2 碱基错配封闭, 探针的设计采用 Oligo7.0。

1.2.3 不对称 PCR 和杂交 不对称 PCR 反应体系为 15μl, 包括 20ng/μl 大菱鲆基因组 DNA1μl, 10×PCR buffer 1.5μl, 2.5mmol/L dNTPs 1.2μl, 25mmol/L MgCl₂ 1.2μl, 5U/μl *Taq* DNA polymerase 0.075μl, 1×LC Green Plus 1.5μl, 10μmol/L 正链引物 1.2μl, 2μmol/L 负链引物 1.2μl, 加 ddH₂O 至 15μl; 设置 PCR 仪的程序为 95℃ 变性 5min; 94℃ 30s, 引物对应 *T_m* 30s, 72℃ 30s, 60 个循环; 72℃ 延伸 7min, 4℃ 保存; 每个不对称 PCR 反应体系中加入 10μmol/L 对应位点的 3' 端封闭的非标记探针 1.5μl, 设置 PCR 仪的杂交程序: 变性温度 95℃ 10min, 之后自然降至室温, 置 4℃ 保存;

1.2.4 HRM 采集数据并分型 在 Light Scanner 分析仪(Idaho, LightScanner96)上对杂交产物的高分辨

率熔解曲线进行采集,采集曲线的温度范围为 45—97℃,每摄氏度获得 10 个点,升温速度为 0.1℃/s。数据的检索和分析运用 Light Scanner 分析软件。评判三种基因型的方法与探针设计依据有关,若探针与 SNP 位点野生型完全匹配,则熔解峰对应温度较高者为野生型,对应较低者为突变纯合型,具有双峰的为杂合型(Shale *et al.*, 2007)。

2 结果

2.1 候选 SNP 的获得

本实验共得到 1457 个重叠群(Contigs)和 2738 个单一序列,在 1457 个 Contigs 中,Contig size (组成 Contig 的 EST 序列数)最大值为 411,平均值为 6.65 个 EST 序列。其中包含 4 个及其以上的 EST 序列的重叠群共有 522 个。从这 522 个重叠群中共筛选出 160 多个候选 SNP 位点,从中挑选部分位点用于本实验。图 1 显示了候选 SNPs 在 ContigExpress Project 软件检测序列拼接中的结果,C337TSNP 位点被选定为候选 SNP 位点。

```
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCTTGCGCGCCAATCAGATCATCCCCACTGC
300 TTCGAGCTCAGGGACACTGGTCGCTATGGTTTCATCCTGCGCGCCAATCAGATCATCCCCACTGC 365
```

图 1 利用 ContigExpress Project 软件检测候选 SNPs 的拼接结果

Fig.1 Testing putative SNPs splicing results by ContigExpress Project software

注:图中标注黑体下划线碱基被认定为发生突变的候选 SNP 位点

2.2 基因组 DNA 的提取

提取的 DNA 用 1.0% 琼脂糖凝胶电泳和 Thermo 公司的超微量紫外 Nanodrop2000 检测提取的基因组 DNA 的质量和浓度,图 2 为部分样品基因组 DNA 的琼脂糖凝胶电泳检测结果;提取的基因组 DNA 的浓度多在 300—400ng/μl 之间, A_{260}/A_{280} 均在 2.0 附近,无明显拖尾现象,结果表明基因组 DNA 比较完整、无污染。将基因组 DNA 稀释至约 20ng/μl 备用。

2.3 引物及探针

用设计的引物对大菱鲂 DNA 做普通 PCR,结果显示,能扩增出目的片段的引物有 37 条,合 45 个位点,成功率为 66.1%。图 3 为部分引物对大菱鲂基因组 DNA 做普通 PCR 摸索 T_m 的扩增结果。对能扩增出目的片段对应的 SNP 位点设计探针,做不对称 PCR、杂交,结果显示,能与 DNA 分子杂交的探针有 13 条,杂交率为 28.9%。表 2 为 11 个能与探针杂交的位点信息,图 4 为 Unlabeled Probes 和 PCR 产物杂交 HRM 分析结果。

2.4 基因分型

用能做出探针峰的引物和探针对群体样品进行不对称 PCR、探针杂交和基因分型。结果发现,能成功进行基因分型的位点共有 8 个,占杂交位点的 61.5%,其中有 6 个为碱基替换型位点,即 G-A 占 3 个, C-T 占 3 个, 2 个为碱基颠换型位点,即 T-G 和 A-T 各占 1 个。图 5、图 6 分别为部分位点 Unlabeled Probes 基因分型标准熔解曲线图和衍生标准熔解曲线峰图结果。

2.5 包含 SNP 位点序列的功能分析

借助 NCBI 网站 BlastX 程序分别对 8 个含有 SNP 位点 contigs 的拼接序列进行分析,据此推测其潜在功能。比对结果显示 8 个 EST-contigs 的拼接序列中,除 Contig0496 没有显著相似的功能基因匹配外,其余皆找到相关匹配,涉及的基因种类较杂,表 3 是包含 SNP 位点序列预测蛋白的基因注释信息。

3 讨论

3.1 SNP 标记开发方案

目前 SNP 标记的开发一般情况下是通过基因组测序得到的,利用测序获得的序列进行比对得到 SNP。因此,对于已经开展大规模

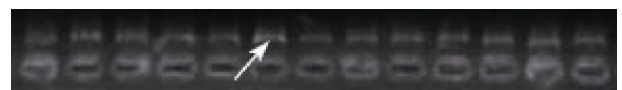


图 2 部分样品基因组 DNA 的琼脂糖凝胶检测

Fig.2 Genomic DNA agarose gel detection of part samples

注:图中箭头标注行为基因组 DNA 琼脂糖凝胶条带效果

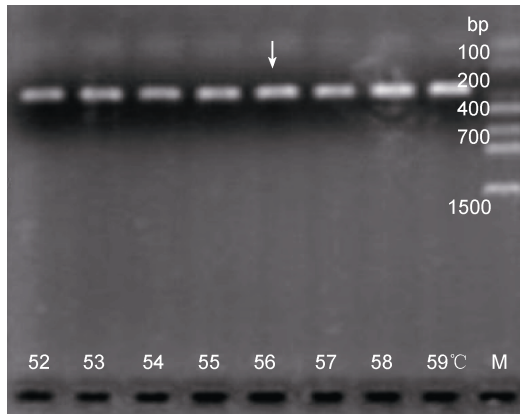


图3 普通 PCR 摸索引物最佳 T_m 的扩增结果

Fig.3 The best T_m multiplies results by common PCR
注: 图中箭头标注条带对应温度为选取最佳退火温度

模基因组测序的物种来说, 开发 SNP 常常伴随着基因组序列的拼接过程而产生。在缺乏基因组序列的情况下, 由于 EST 序列和基因组序列不完全一致, 利用 EST 序列开发 SNP 会增加 SNP 验证的难度和成本。但对于大多数非模式生物而言, 利用 EST 序列开发 SNP 标记是目前筛选 SNP 的重要途径(王绍宗等, 2010)。吴玲等(2010)利用生物信息学方法在基于 EST 序列玉米 SNP 标记开发的研究中, 通过比对分析 2018530 条 EST 序列, 拼接发掘出遍布全基因组的 80363 个 SNP 位点。在 SNP 位点两侧保守序列上设计引物, 开发出 12388 个 SNP 标记。王绍宗等(2010)利用长牡蛎已有的 EST 序列数据库, 筛选得到候选 SNP 位点共计 1140 个。通过片段长度差异等位基因

表 2 11 个能与探针杂交的位点信息

Tab.2 11 hybridization sites information with probes

SNP 位点	引物与探针(5'-3')	产物长度	SNP 类型	突变类型
G583ASNP	F: CCTTCATCCCTTCAGACC R: CCCTCTAAATCAGCGTTGTG P: TGGGTCACCTCAGCAATAATCGG	147	G-A	转换
G639ASNP	F: CCTTCATCCCTTCAGACC R: CCCTCTAAATCAGCGTTGTG P: GCCTGGGTTTTGAGCCGTGT	147	G-A	转换
C282TSNP	F: ACGACTCTGCCTTTCTTC R: CGCTCAGCCTTTATTGGAC P: CTGAGACTGAGGTGGGAAGGT	342	C-T	转换
C310TSNP	F: CAACCAGGGCATCAAGTA R: TAAAGAGCCACGAGTCATA P: GATGAAGCCGTACCGACCCGCA	236	C-T	转换
C433TSNP	F: GGTGGATGAGAACGGTAA R: AAAATTGTGAAACGATGTGC P: GAGGCAGCACTTCCGCAGAT	229	C-T	转换
G280ASNP	F: CTGTGCTCCTATGTGGCTTAC R: GGGTCGTCCTTCTGTGC P: GTTTCATGACTTAGCGAAGCGG	262	G-A	转换
A424TSNP	F: GCCGAGTAACAGACAGAAG R: TAACCTGAACTAAACGGACATT P: CAGCCATTCATGTTGCAGCGA	138	A-T	颠换
G681ASNP	F: TCTTCCTCCACGAAACCG R: GATGGCGTAGGCAAACAG P: CAAGAGGACGACGAACCCGAGAG	248	G-A	转换
C524TSNP	F: ACCAATGTGGGAAGAAGA R: TTGAAAGCAGGATAGCAG P: CGAGTGTAGACCCAGGCTTGAT	163	C-T	转换
A270TSNP	F: GCATTACCTTCTAGCCCTC R: GCAGGCTTTTATTTCCGTT P: TTCCAGTTTGAAACAGAATCACCC	344	A-T	颠换
C428TSNP	F: GCATTACCTTCTAGCCCTC R: GCAGGCTTTTATTTCCGTT P: TTCCAGTTTGAAACAGAATCACCG	344	C-T	转换

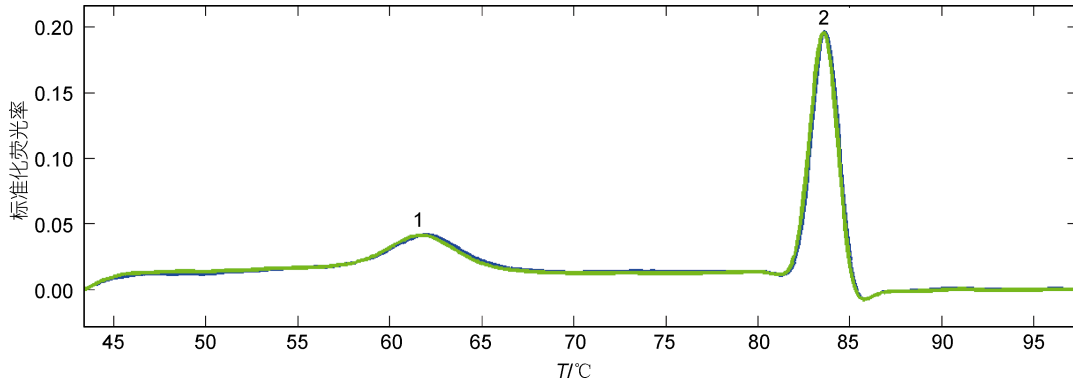


图 4 Unlabeled Probes PCR 产物杂交的 HRM 分析结果

Fig.4 HRM analysis results of Unlabeled Probes PCR products hybridization
注: 1 为探针-DNA 杂交峰, 2 为双链 DNA 峰

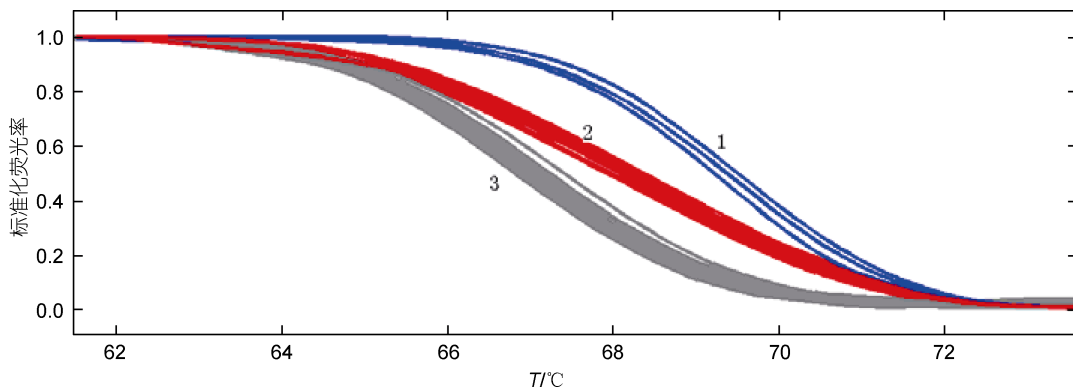


图 5 Unlabeled Probes 基因分型分析标准熔解曲线图

Fig.5 Genotyped analysis normalized melting curves results by Unlabeled Probes
注: 1 表示野生纯合基因型, 2 表示突变杂合基因型, 3 表示突变纯合基因型

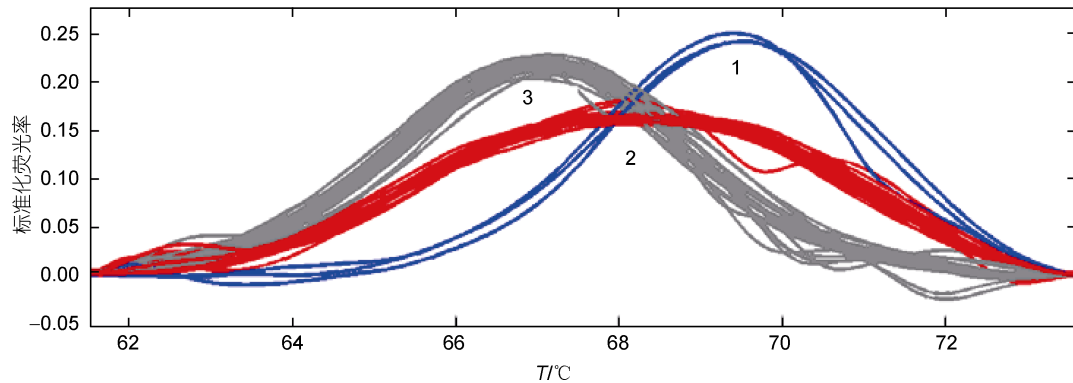


图 6 Unlabeled Probes 基因分型分析标准熔解曲线峰图

Fig.6 Derived genotyped analysis normalized melting curves peaks results by Unlabeled Probes
注: 1 表示野生纯合基因型, 2 表示突变杂合基因型, 3 表示突变纯合基因型

特异性 PCR(FLDAS-PCR)分型方法, 在一野生群体中进行检测和验证, 结果共有 17 个候选 SNP 位点显示多态性。研究表明, 通过基于 EST 数据库的 SNP 开发, 可以有效弥补某些海洋生物基因组学滞后影响 SNP 标记开发的缺陷。

HRM 是筛查 SNP、SSR 等标记及基因分型的一项新技术。实验表明所使用的饱和染料有较强的 DNA 结合能力(Cheryl *et al*, 2008), 但是在本实验过程中显示了一定程度的抑制作用。该技术在实验过程中可实现闭管操作, 降低误差的可能性(Mont-

表 3 对序列预测蛋白的基因注释信息
Tab.3 Gene annotation of predicted proteins of the contigs

重叠群名称	注释
Contig1377	有保守序列的假定蛋白(Conserved hypothetical protein)
Contig0143	泛激素 40s 核糖体蛋白 s27a (Ubiquitin-40s ribosomal protein s27a)
Contig1044	细胞色素 <i>b</i> (Cytochrome <i>b</i>)
Contig0165	类结合蛋白 2 透明质酸(Hyaluronan-binding protein2-like)
Contig0239	去氢酶副族 2 (NADH dehydrogenase subunit2)
Contig0496	—

gomery *et al.*, 2010)。探针采用两碱基错配的方法实现 3'封闭, 目的是防止 PCR 过程中探针 3'端延伸, 影响目的片段扩增。为防止 3'端封闭不完全影响扩增效果, 故本实验设计在 PCR 之后加探针。在 PCR 反应前加入矿物油, 防止等量的反应体系在 PCR 过程中挥发, 影响最终实验结果。

3.2 候选 SNP 的筛选

张新叶等(2009)对 21657 条杨树 EST 序列进行分析, 共得到 94 个候选 SNP 位点; 吴玲等(2010)对 2018530 条玉米 EST 序列比对分析, 发掘出遍布全基因组的 80363 个 SNP 位点。本实验共分析了 12428 条 EST 序列, 从得到的 522 个重叠群中筛选出 160 多个候选 SNP 位点。对比结果显示, 在大菱鲆现有的 EST 序列数量条件下, 分析得到的候选 SNP 结果较好; 相对于研究较为充分的玉米而言, 大菱鲆目前的 EST 序列较少, 这方面的工作还要进一步开展。

在本实验重叠群的分析中可以看出, 现有的 EST 序列在基因组中分布并不均匀, 较多的冗余和单一序列是候选 SNP 得率较低的原因之一。从候选 SNP 统计结果可以看出, 在碱基替代型候选 SNP 中, A/G、C/T 转换最多, 这与张新叶等(2009)对杨树相关研究的统计结果大致相同。理论上 SNP 发生有四种类型, 即转换、颠换、插入和缺失, 但实际上发生的只有两种, 即转换和颠换, 二者之比为 2:1 (王强等, 2010), 特别是 C-T 转换较多, 因为 CpG 的 C 是甲基化的, 容易自发脱氨基形成胸腺嘧啶 T (Johnson *et al.*, 2000)。可以看出候选 SNP 统计结果与理论分析一致, 这些 SNP 的具体信息对下一步的 SNP 标记真伪鉴定提供了重要依据(张新叶等, 2009)。

3.3 SNP 标记的验证

3.3.1 引物和探针的设计

由于 LightScanner 检

测突变的准确性和特异性与目的片段的大小有很大关系, 所以设计引物的时要求目的片段小于 300bp, 准确性和特异性可达到 99.6%。要求探针的退火温度 60—80℃, 使之与引物和目的产物退火温度有明显差别, 3'端封闭。人为错配碱基的设计遵循强弱错配原则, 即如果 3'端是强错配(A/G 或者 C/T), 则需要引入弱的错配(C/A 或者 G/T), 引入错配即可阻断 3'端的扩增, 反之亦然(张振, 2010)¹⁾。Shale 等(2007)在非标记探针检测异常融解峰特征的分析报告中, 测试了 3'端修饰物如 C3 spacer, 氨基修饰的 C6、磷酸化、反向 dT、3'端单碱基错配等在非标记探针实验中的封闭能力。结果显示, 没有一个 3'端修饰物是 100%有效的, 但氨基修饰的 C6、反向 dT、C3 spacer 表现出非常好的封闭效力, 磷酸则不是一个好的封闭物, 而单碱基错配则应避免作为 3'端的封闭修饰物, 这是本实验探针杂交率低的一种解释。

3.3.2 不对称 PCR 实验过程中, 用引物对大菱鲆基因组 DNA 做普通 PCR, 结果有些没有扩增产物, 有些扩增产物为大的非目的片段, 还有些扩增产物有目的片段但混有杂带。对于无明显扩增产物的引物对来说, 王绍宗等(2010)认为这些扩增产物为超长片段或者引物序列位于外显子拼接位点, 而李宏俊等(2011)则认为是这些 EST 序列本身碱基分布不匀称, 对设计的引物质量较差, 引物二聚体、发卡结构或错配较为严重; 对于后两种情况的分析, 李宏俊等(2011)认为 EST 序列为基因组 DNA 转录、剪切, 修饰后的表达序列, 所以用根据 EST 序列设计的引物, 可能横跨基因组 DNA 内部的某些内含子, 这是导致扩增产物为大的为目的片段或条带混杂的原因。

能用扩增出目的片段的引物做不对称 PCR, 途径有三条: (1) 通过限制性引物的浓度, 在 PCR 的反

1) 张 振, 2010. 基于微卫星和 SNP 标记的皱纹盘鲍遗传连锁图谱及其应用. 北京: 中国科学院研究生院博士学位论文, 64—68

应过程中,当限制性引物因量少而消耗完后,非限制性引物继续扩增而产生大量的单链 DNA。(2) 通过限制性引物的碱基数目与组成,使限制性引物的退火温度(T_m)低于非限制性引物的 T_m ,经常规 PCR 的几个循环后,将 PCR 的退火温度提高至非限制性引物的 T_m ,此时限制性引物不能结合到模板上,只有非限制性引物继续扩增产生大量单链 DNA。(3) 二次 PCR 方法,即先用等浓度的引物进行第一次 PCR 扩增,扩增出目的片段的双链(dsDNA),再以此 PCR 产物为模板,利用其中的一条引物进行二次 PCR 扩增,扩增出目的片段的单链(ssDNA)。鉴于实验操作简便,本实验选择第一种途径做不对称 PCR。反应体系中所加的非限制性引物和限制性引物的量比例至关重要,据文献中报道,其最佳比例一般是 1:50—1:100,关键因素是限制性引物的绝对量。限制性引物太多或太少,均不利于制备 ssDNA。杨泽晓等(2011)对不对称 PCR 引物比例做了摸索,结果发现限制性引物和非限制性引物的比例为 20:1 上下时,其 PCR 产物的效果最佳。本实验进行了梯度为 5:1、10:1、20:1、40:1 的引物比梯度探索,结果显示非限制性引物和限制性引物比为 10:1 或 20:1 时效果普遍较好,但各引物对本身具有特异性,还需要缩小范围进一步研究,以期达到更好的效果。

3.3.3 探针杂交 探针与单链 DNA 杂交结果显示,有些熔解曲线没有对应的探针峰,有些熔解曲线出现杂峰。前者原因为 SNP 位点周围序列质量较差或 3' 端的两碱基错配封闭不完全或杂交不稳定等。Shale 等(2007)在非标记探针检测异常融解峰特征的分析报告中,对非标记探针检测中所出现的异常融解曲线进行了研究,非标记探针 3' 端的不完全封闭导致了聚合酶介导的探针延伸,是出现异常融解曲线的决定因素。一旦部分探针封闭不完全或解封闭,它就会成为上游引物并和下游引物一起,产生一个新的扩增产物,再加上探针与其本身扩增产物一定程度上的杂交,表现在最终结果熔解曲线上则呈现出杂乱峰图。

3.3.4 基因分型 将有探针峰的位点用于群体基因分型验证,结果有些位点无探针峰多态性,分析其原因为此候选 SNP 为假阳性位点、国内养殖群体单一、取样量的局限性等;有些杂合突变没有出现双峰,可能原因为探针较短,造成双峰距离太近,致使杂合双峰不明显;有些曲线混乱,可能是杂交或离心不充分、矿物油和试剂有抑制作用等;杂峰和基因型缺失的情况,原因前文已经阐述。另外还发现了各个基因

型数量分布并不均匀,表明该样品的遗传变异程度较高(陈扣扣等,2009)。殷豪等(2011)利用高分辨率熔解曲线分析梨微卫星标记的研究中,扩增产物生成的熔解曲线复杂,无法进行有效的归类,究其原因可能是 PCR 扩增子类型复杂,除有 SSR 位点外,在序列的其它位置也可能存在一些差异,或者是所采用的退火温度不合适而导致了一些非特异性扩增,本文出现的峰图杂乱也可能是这方面的原因。

总之,用 HRM 做单核苷酸多态性分析要求精密度较高,步骤繁琐,只有到熔解曲线分析时才能看到整个实验的效果,中间缺乏检测手段,所以要求操作上,特别是做群体分析时,除了允许样品本身差异外,其它一切变量要尽可能相同,防止一切由非单碱基差异导致的熔解曲线的差异。

参 考 文 献

- 马爱军,王新安,雷霖霖等,2008. 大菱鲂(*Scophthalmus maximus*)四个不同地理群体数量形态特征比较. 海洋与湖沼, 39(1): 24—29
- 马爱军,王新安,雷霖霖等,2009. 大菱鲂(*Scophthalmus maximus*)不同生长阶段体重的遗传参数和育种值估计. 海洋与湖沼, 40(2): 187—194
- 马爱军,王新安,薛宝贵等,2010. 大菱鲂(*Scophthalmus maximus*)选育家系的构建和培育技术研究. 海洋与湖沼, 41(3): 301—306
- 丰贵鹏,李亮,2010. 高分辨率熔解曲线法的非标记探针基因分型技术. 生命的化学, 30(2): 314—317
- 王强,李卫真,张永云,2010. SNPs 技术的研究进展及应用. 四川畜牧兽医, 38(8): 31—33
- 王绍宗,李莉,元海刚等,2010. 长牡蛎(*Crassostrea gigas*)17 个 EST-SNP 标记的开发. 海洋与湖沼, 41(2): 274—281
- 王新安,马爱军,2010. 应用时间序列分析模型对大菱鲂选育 F₁ 优良家系发育的动态研究. 水产学报, 34(9): 1319—1326
- 许可,马爱军,王新安等,2009. 大菱鲂(*Scophthalmus maximus*)生长性状相关的微卫星标记筛选. 海洋与湖沼, 40(9): 577—583
- 李宏俊,梁瑜,邢坤等,2011. 紫贻贝 EST-SNP 的筛选及多态性检测. 水产学报, 35(3): 348—355
- 杨泽晓,韩雪清,王印等,2011. 猪圆环病毒 2 型不对称 PCR 检测方法的建立. 动物医学进展, 32(2): 1—5
- 吴玲,付凤玲,李晚忱等,2010. 利用生物信息学方法进行基于表达序列标签的玉米单核苷酸多态性标记的开发. 核农学报, 24(5): 968—972
- 邹曙明,李思发,蔡完其,2001. 牙鲆和大菱鲂养殖群体的分子标记和遗传变异. 中国水产科学, 7(4): 25—38
- 张新叶,宋丛文,杨彦伶等,2009. 基于 EST 序列的杨树候选 SNPs 标记分析. 华中农业大学学报, 28(6): 741—745

- 陈扣扣, 杨博辉, 郭 健等, 2009. 甘肃高山细毛羊肉毛兼用品系微卫星标记与生产性状的相关性分析. 贵州农业科学, 37(6): 116—122
- 侯仕营, 马爱军, 王新安等, 2011. 大菱鲂 4 个引进地理群体遗传多样性的微卫星分析. 渔业科学进展, 32(1): 16—23
- 殷 豪, 王彩虹, 田义轲等, 2011. 利用高分辨率熔解曲线 (HRM) 分析梨微卫星标记. 园艺学报, 38(8): 1601—1606
- 雷霖霖, 2003. 大菱鲂养殖技术. 上海: 上海科学技术出版社, 8
- 雷霖霖, 2005. 海水鱼类养殖理论与技术. 北京: 中国农业出版社, 524—541
- Carl T W, Gudrun H R, Cameron N G *et al*, 2003. High-Resolution Genotyping by Amplicon Melting Analysis Using LC Green. *Clinical Chemistry*, 49(6): 853—860
- Cheryl S T, Dominic E D, 2008. Differentiation between vaccine and wild-type varicella-zoster virus genotypes by high-resolution melt analysis of single nucleotide polymorphisms. *Journal of Clinical Virology*, 43(1): 18—24
- HOU Shiyang, MA Aijun, WANG Xinan *et al*, 2011. Isolation and characterization of 45 polymorphic microsatellite loci of turbot (*Scophthalmus maximus*) and cross-species amplification. *Chinese Journal of Oceanology and Limnology*, 29(2): 311—316
- Jiang Guodong, Li Jiqin, Li Ling *et al*, 2011. Development of 44 gene-based SNP markers in Zhikong scallop, *Chlamys farreri*. *Conservation Genet Resour*, 3(4): 659—663
- Johnson G C L, Told J A, 2000. Strategies in complex disease mapping. *Current Opinion Genetics & Development*, 10(3): 330—334
- Montgomery J L, Sanford L N, Wittwer C T, 2010. High-resolution DNA melting analysis in clinical research and diagnostics. *Expert Rev Mol Diagn*, 10(2): 219—240
- Schlotterer C, 2004. The evolution of molecular markers—just a matter of fashion. *Nat Rev Genet*, 5(1): 63—69
- Shale Dames, Rebecca L Margraf, David C Pattison, 2007. Characterization of Aberrant Melting Peaks in Unlabeled Probe Assays. *The Journal of Molecular Diagnostics*, 9(3): 290—296
- Zhou L, Myers A N, Vandersteen J G *et al*, 2004. Closed-tube genotyping with unlabeled oligonucleotide probes and a saturating DNA dye. *Clin Chem*, 50(8): 1328—1335

STUDY ON THE DEVELOPMENT OF SNP MARKERS IN *SCOPHTHALMUS MAXIMUS* USING HIGH-RESOLUTION MELTING (HRM)

LIU Qing-Ming^{1,2}, LI Meng^{1,2}, MA Ai-Jun¹, WANG Xin-An¹, HUANG Zhi-Hui¹, GUO Jian-Li¹

(1. Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences; Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture; Qingdao Key Laboratory for Marine Fish Breeding and Biotechnology, Qingdao, 266071;
2. Shanghai Ocean University, Shanghai, 201306)

Abstract The experiment mines *Scophthalmus maximus* 12428 EST sequence from NCBI database. Putative SNP were detected from assembled contigs using Vector NTI Advance 11. SNP genotyping was performed using high-resolution melting to different character populations. Primers and probes were designed using Primer5 and Oligo7, respectively. The results show that, More than 160 putative SNP were found from 522 contigs; 56 putative SNP were selected for primer design and for SNP marker development, 37 pairs of primer were successfully amplified, including 45 putative SNP site, the success rate is 66.1%; Probes which can hybrid with Putative SNP sites have 13, hybridization Rate is 28.9%; 8 sites of hybridization sites displayed polymorphisms in different character populations, accounting for 61.5% of the hybrid sites. There are 6 transition sites, namely the G-A and C-T are 3 of each. There are 2 transversion sites, namely the T-G and A-T are 1 of each.

Key words Single nucleotide polymorphism (SNP), *Scophthalmus maximus*, High-resolution melting analysis, EST sequence