

# 用 RAPD 技术筛选中国荷斯坦牛产奶量性状遗传标记

杜立新, 万海伟, 王爱华, 李宏滨

(中国农业科学院畜牧研究所, 北京 100094)

**摘要:** 选用 320 条 10 碱基随机引物在由 36 头高产(305 d 产奶量 > 8 500 kg) 和 32 头低产(305 d 产奶量 < 5 600 kg) 中国荷斯坦牛所组成的 2 个 DNA 池间进行 RAPD-PCR 扩增, 从中筛选出稳定性好、带型清晰且在 2 个 DNA 池间有明显差异的 RAPD 引物 24 条。用筛选到的 24 条引物对所有个体进行 PCR 检测, 得到了 5 个与奶牛产奶量性状相关的 RAPD 标记, 并对其中 4 个进行了克隆测序和 SCAR 标记转化以及生物信息学分析。结果表明: SCAR 标记 yield<sub>s</sub>139 与产奶量密切相关, 应用电子克隆方法已将该标记由 921 bp 延伸到 2 141 bp, 经同源性比较发现, 此标记对应于奶牛基因组中 LINEs 家族中的一个重复元件, 推测该元件附近可能存在与产奶量性状相关的 QTLs 或主效基因。

**关键词:** 奶牛; 产奶性状; 分子标记

**中图分类号:** S823.3

**文献标识码:** A

**文章编号:** 0366-6964(2005)09-0882-05

遗传因素是影响奶牛产奶量的关键因素之一。因此, 从 DNA 水平对奶牛产奶量性状进行研究, 寻找与产奶量紧密连锁的分子标记或克隆出控制产奶量的基因, 对于进行标记辅助选择及培育高产奶牛, 具有十分重要的意义。随机扩增多态性 DNA (Random Amplified Polymorphic DNA, RAPD), 是由 Williams 和 Welsh 在 1990 年各自独立发现的一种 DNA 多态检测技术<sup>[1,2]</sup>。该技术自问世以来, 以其简便、经济、快捷、且无须事先了解物种的相关分子生物学信息等特点, 迅速渗透到基因组研究的各个领域<sup>[3,4]</sup>。SCAR(Sequence Characterized Amplified Region) 标记, 即序列特异性扩增区域, 是在 RAPD 技术基础上发展起来的一种具有更高的可重复性的分子标记<sup>[5]</sup>。本研究应用 RAPD 技术分别对高产奶牛组和低产奶牛组的基因组 DNA 进行了多态性分析, 并对筛选出的与产奶量相关的标记进行克隆测序和 SCAR 标记转化以及生物信息学分析, 以期找到与奶牛产奶量性状相关的分子标记或新的候选基因, 为进行标记辅助选择及培育高产奶牛提供科学依据。

## 1 材料与方法

### 1.1 试验材料

1.1.1 奶牛血样 荷斯坦牛血样采自山东济南佳宝奶牛场。高产奶牛(305 d 产奶量 > 8 500 kg) 36 头, 低产奶牛(305 d 产奶量 < 5 600 kg) 32 头, 且全为第 1 胎产奶。颈静脉采血 10 mL, ACD 抗凝, -20 °C 保存。

1.1.2 引物及主要试剂 320 条 S 系列随机引物购自上海生工生物工程公司; *Taq* DNA 聚合酶、限制性内切酶 *Eco*R I、DNA Marker DL2 000、dNTP 及 DNA 回收试剂盒均购自日本 TaKaRa 公司; 感受态细胞大肠杆菌 DH5 $\alpha$  为北京天为时代公司产品, 质粒 pGEM-T-easy 购自美国 Promega 公司。

### 1.2 试验方法

1.2.1 正交优化法建立 RAPD-PCR 最佳反应体系 常规方法提取奶牛血液 DNA 后, 用分光光度计检测每个个体的 DNA 浓度, 取每个个体等量 DNA 分别混成高产奶牛 DNA 池和低产奶牛 DNA 池。采用 L25(5<sup>6</sup>) 正交表将模板、引物、dNTPs、Mg<sup>2+</sup>、*Taq* 酶、BSA 等 6 个因素分别设置 5 个浓度梯度, 以反应总体积 25  $\mu$ L 进行 PCR 扩增。循环参数为: 95 °C 预变性 5 min; 95 °C 变性 1 min, 36 °C 退火 1 min, 72 °C 延伸 2 min, 41 循环后, 72 °C 延伸 10 min。扩增产物在 1.5% 琼脂糖凝胶电泳, 电压为 5 ~ 6V/cm。电泳缓冲液为 0.5  $\times$  TBE, 电泳 1~ 2 h

收稿日期: 2004-08-26

基金项目: 国家奶业专项(2002BA51801-13)

作者简介: 杜立新(1956), 男, 陕西华阴人, 博士, 教授, 主要从事动物遗传资源与分子育种研究。E-mail: lxdu@263.net

后在凝胶成像系统上检测并拍照。

1.2.2 引物的筛选及 RAPD 标记的克隆测序 用 320 条 S 系列 10 bp 随机引物,分别以高产 DNA 池和低产 DNA 池为模板进行 PCR 扩增,扩增产物用 1.5% 琼脂糖凝胶电泳检测,从中筛选扩增条带多、重复性好、分辨率高且扩增条带在高、低产 DNA 池间有显著差异的引物。经 DNA 片段的回收、连接、感受态的制备、转化、质粒 DNA 的提取及重组质粒的酶切鉴定等环节后,提交北京博亚公司进行测序。

1.2.3 SCAR 标记的转化 根据测得的 RAPD 标记 DNA 序列,用 Premier 5 软件设计了 4 对对应的特异性引物进行 PCR 扩增。PCR 反应体系为 25  $\mu$ L,反应程序为:94  $^{\circ}$ C 预变性 5 min;94  $^{\circ}$ C 变性 1 min,58  $^{\circ}$ C 退火 1 min,72  $^{\circ}$ C 延伸 1 min,32 循环后 72  $^{\circ}$ C 延伸 10 min。

1.2.4 图谱分析及数据统计处理 对扩增图谱观察记录,结果中条带的有和无、强和弱统一分别用“1”和“0”表示。统计特异条带出现的频率。通过  $\chi^2$  (卡方) 检验某一条特异带在高产群体和低产群体间出现的频率差异。

1.2.5 生物信息学分析 对测序成功的 RAPD 标记在美国国立卫生研究院(NIH)国家生物学信息中心(NCBI, <http://www.ncbi.nlm.gov/Blast>) 进行序列同源性分析,并在 Bos\_Taurus\_WGS\_Trace 数据库进行同源性搜索,对搜索到的重叠群(Contigs)用 DNAMAN 5.0 软件进行拼接延伸。

## 2 结果与分析

### 2.1 不同产奶量奶牛基因组的 RAPD 多态性分析

2.1.1 RAPD 反应体系的正交优化 采用 6 因素、5 水平试验共 25 个组合对 RAPD 反应体系进行正交优化,结果如图 1 所示:2、3、4、5、8、11、15、17、18、24、25 泳道无任何扩增条带;1、9、10、16、19、21、22 泳道条带数少,且不清晰,重复性差;6、7、12、13、14、20、23 泳道带数较多,带型清晰,重复性好。从试验成本角度考虑,确定 20 泳道的试验条件为奶牛基因组 DNA RAPD-PCR 最佳反应体系:PCR 反应总体积为 25  $\mu$ L,其中包括 ddH<sub>2</sub>O 13.1  $\mu$ L、10 $\times$  Buffer 2.5  $\mu$ L、MgCl<sub>2</sub> (15 mmol/L) 2.5  $\mu$ L、dNTP (2 mmol/L) 2.5  $\mu$ L、RAPD primer 2  $\mu$ L、Template DNA 2  $\mu$ L、Taq 酶 0.2  $\mu$ L、BSA 0.2  $\mu$ L。

2.1.2 引物筛选 采用优化后的反应体系对 320 条 S 系列随机引物进行 PCR 扩增。经 3 轮筛选以



1~ 25. 分别为根据 L<sub>25</sub>(5<sup>6</sup>) 正交表设计的条件组合的扩增结果

1~ 25. The RAPD-PCR results of 25 elements component by L<sub>25</sub>(5<sup>6</sup>) Orthogonal design

图 1 奶牛基因组 DNA RAPD-PCR 正交试验扩增结果

Fig. 1 The RAPD-PCR results of bovine DNA using Orthogonal optimization test

重复验证,得到稳定性较好、带型清晰且在两 DNA 池间有明显差异的候选 RAPD 引物 24 条,分别为 S35、S73、S81、S85、S100、S131、S132、S135、S139、S145、S148、S198、S217、S270、S406、S409、S417、S425、S428、S440、S456、S457、S506、S517。

### 2.1.3 RAPD 标记检测及其与产奶量相关性分析

用筛选到的 24 条 RAPD 引物对所有个体进行 PCR 检测,根据扩增图谱统计分析后,得到了 5 个与奶牛产奶量性状相关的 RAPD 标记,分别为 S139-921、S409-414、S217-1227、S198-258、S417-J。其中,S139-921、S409-414、S417-J、S198-258 在高产奶牛个体中出现,与产奶量呈正相关。标记 S217-1227 在低产奶牛个体中出现,与产奶量呈负相关。每个 RAPD 标记与产奶量的相关程度的  $\chi^2$  检验值见表。

表 1 各个 RAPD 标记的  $\chi^2$  检验值  
Table 2 The  $\chi^2$  value of each RAPD marker

RAPD 标记	$\chi^2$ 值	P	$\chi^2_{0.05}$	$\chi^2_{0.01}$
S139-921	37.01	< 0.01		
S409-414	16.88	< 0.01		
S217-1227	23.12	< 0.01	3.841	6.635
S198-258	10.81	< 0.01		
S417-J	9.75	< 0.01		

2.1.4 SCAR 标记的转化 根据 RAPD 标记的测序结果,设计合成 SCARs 引物对所有的奶牛个体 DNA 进行 PCR 扩增验证。检测结果表明,S139-921 标记阳性率为:高产 100%,低产 25%,可以认

为是成功转化,将其命名为 yield-s139(图2)。另外4个 RAPD 标记的 SCAR 扩增,高产个体阳性率均为 100%,而低产个体阳性率也都在 70% 以上,未得到成功转化。

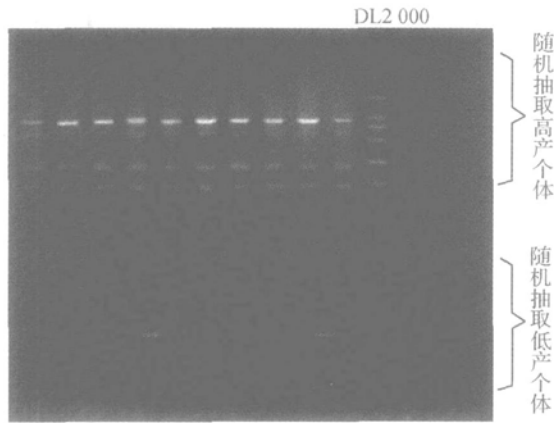


图2 RAPD 标记 S139-921 SCAR 扩增结果  
Fig 2 The transformed SCAR results of RAPD marker S139-921

2.1.5 生物信息学分析 RAPD 标记 S139-921 经 BLASTN 分析,与 GenBank 中多个牛的基因有很高的同源性,且同源区高度一致。根据这几个基因在 GenBank 的登录号,查阅基因的登录信息,经分析后发现,高同源区位于牛基因中一个长散在重复元件 Bov-B LINE/BovB 区域。推测 RAPD 标记 S139-921 对应于奶牛基因组中的一个 LINE 重复元件。

用 RAPD 标记 S139-921 序列检索 Bos\_Taurus\_WGS\_Trace 数据库后发现,该标记与 Contigs gnl1 til 380567341 与 gnl1 til 380566310 overlaps 区同源性非常高,达到 98%,说明这两个 Contigs 可以用来拼接延伸 RAPD 标记 S139-921。用 DNAMAN 5.0 软件拼接,得到一个长度为 2 141 bp 的序列,使原来的 RAPD 标记向左端延伸了 562 bp,向右端延伸了 658 bp。经对延长序列中重复元件进行的 RepeatMasker 程序检索发现,延长后得到的序列中被遮蔽的序列的确是奶牛基因组中的一个重复元件,包含一个部分 Bov-B LINE/BovB 区域,一个部分 LIMC/D LINE/L1 区域及一个短的 SINE 部分区域,进一步证实所获得的 RAPD 标记 S139-921 为一个包含部分 LINE 元件的 DNA 片段。

### 3 讨论与结论

#### 3.1 RAPD 标记与产奶量的相关性分析

本研究主要从分子水平研究奶牛的产奶量性状,因此采样时选择了同一奶牛场同一品种产奶量相差悬殊的两个极端群体,而且所有个体的年龄、产奶胎次、产奶时间及在产奶期间的生理健康情况都大致相同。这样尽可能避免了其它因素的干扰作用,力求所采样品能够准确可靠地反映出其 DNA 水平的差异。本研究用了 320 条 RAPD 引物,共扩出 2 956 条带,即大约对基因组的近 3 000 个位点进行了分析,所得结果可在一定程度上反映基因组的异同。然而,由于上述因素的限制,样本的数量比较小。

通过分析扩增图谱发现,在近 3 000 多条带中,绝大部分都完全一致,说明群体间遗传差异极小,除了泌乳性状外,其它性状在理论上是基本一致的。只有 5 条带在高、低产群体中出现极显著差异( $P < 0.01$ )。因此推测这 5 个 DNA 片段可能与产奶量有关,甚至有可能是影响产奶量的基因片段,因此可将其作为产奶量辅助选择的遗传标记。

#### 3.2 SCAR 标记的转化

SCAR 标记对反应条件不敏感,比 RAPD 标记更稳定,重复性高。将 RAPD 标记转化为特定引物的 SCAR 标记,可提高 RAPD 标记作为遗传标记的可信度,使其在辅助育种中更灵活、有效。目前,已有很多 SCAR 标记转换成功的例子<sup>[6,7]</sup>。本研究在对 RAPD 标记 S139-921、S409-414、S217-1227、S198-258 进行了克隆测序的基础上,成功地将 RAPD 标记 S139-921 转化为 SCAR 标记 yield-s139。但更多研究却存在将 RAPD 标记转化为 SCAR 标记后多态性消失的现象<sup>[8,9]</sup>。本研究中另外 4 个 RAPD 标记的转换也未获得成功。这表明将 RAPD 标记转化为 SCAR 标记并非像理论上那么容易,其成功率相对来说还是比较低的。转换失败的原因可能是产生这种 RAPD 多态性的位点发生在 RAPD 10 碱基引物退火区,当引物在 3' 端加长后,加长的碱基有足够的同源性而克服了原来 RAPD 引物的不能完全匹配,因而可以退火、延伸,使原有的 RAPD 多态性消失。因此, RAPD 标记能否得到成功转换,主要取决于增加后的引物长度、延长的碱基与模板同源性的、原 RAPD 引物与模板的错配程度以及 RCR 的反应条件。

### 3.3 RAPD 标记的延伸

由于 RAPD 标记序列通常较短, 所含信息量有限, 除本身作为遗传标记外, 我们无从知道这段序列的生物学意义。如果能将 RAPD 标记向两端延伸, 不仅可以清楚知道产生 RAPD 标记的原因, 而且还可以了解模板之间的准确差异, 充分挖掘 RAPD 标记本身所蕴藏的生物学信息, 为下一步的研究提供更多的信息含量。目前国内外尚未见有行之有效的延伸 RAPD 标记片段的方法, 而传统的反向 PCR 与染色体步行等试验手段, 成本极高, 技术复杂, 操作繁琐, 成功率低。本研究充分利用牛基因组计划的研究成果, 及丰富的生物数据库资源, 运用生物信息学手段, 通过同源性检索 Bos\_Taurus\_WGS\_Trace 数据库, 得到牛基因组中的一系列 Contigs, 然后用 DNAMAN 5.0 软件进行拼接, 使 RAPD 标记向两端的延伸得以成功实现。

### 3.4 关于同源性比较

如果待检索序列中包括一个重复元件的序列, 比如说一个 Alu 重复, 可能会出现许多错误和令人费解的结果。基因组序列中包含大量的重复序列, 典型的是一些多基因族(例如 Alus、LINEs 和人的基因组序列中的 MERs), 甚至 mRNA 中也可能含有重复序列, 几乎都在非翻译区内。因此, 重复元件在数据库序列中也非常普遍, 如果待检索序列中也有这些重复元件, 虽然这些重复元件显示相当的差异, 仍然有足够的相似性使比对具有一定的高显著性, 导致在比对中出现大量不正确的正分。因此, 正确评估搜索结果非常重要。如果非编码区匹配, 而编码区不匹配, 那么序列中就很有可能存在重复元件; 如果查询序列同大量序列匹配, 但是这些序列相互之间没有什么关系, 比对的分值又都很接近, 这样的结果同样可疑。如本研究中用 RAPD 标记 S139-921 检索就出现与许多基因序列匹配的情况, 经 RepeatMasker 程序检索分析后发现, 其的确包含一个 LINE 重复元件部分片段。

### 3.5 关于 LINEs 重复元件

LINEs 是真核生物基因组中反转录转座元件的一个独特家族, 每个元件长约 2~7 kb, 通常含有一个或几个长的 ORF(open reading frame), 与反转录病毒的反转录酶部分同源, 缺失 LTR(long terminal repeat) 序列, 可在基因组中移动, 在基因组中为  $10^4 \sim 10^5$  个拷贝。LINE 元件通常只有 10% 为全长序列, 大多数以 5' 端剪切体的形式出现, 内有

CpG island/GC box、API 和 SP1 结合位点等内部启动子元件。SP1 是增强子中的一个核心元件, 与近侧启动子 GC box 结合可激活上游异源启动子的调控元件。国内外研究表明, 许多哺乳动物的基因中都有 LINE 转座元件的插入<sup>[10,11]</sup>, 如山羊  $\beta$  酪蛋白基因 5' 端上游调控区和内含子 IV 中含有多个 SINE 和 LINE 元件, 这些元件以 5' 端剪切体的形式插入  $\beta$  酪蛋白基因的非编码区中, 内含多种增强转录激活的元件, 同时也参予由 RNA 聚合酶 III 指导的反义 RNA(antisense RNA) 的逆转录, 在转录后调控  $\beta$  酪蛋白基因的表达<sup>[12]</sup>。种种迹象表明, 散在重复元件决不是人们最初以为的那种无功能的自私基因, 而是对进化和遗传、基因表达和调控具有重要功能的成分。

### 参考文献:

- [1] Williams J G, Kubelik A R, Livak K J, et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers[J]. *Nucleic Acids Res*, 1990, 18(22): 6 531~ 6 535.
- [2] Welsh J, McClelland M. Fingerprinting genomes using PCR with arbitrary primers[J]. *Nucleic Acids Res*, 1990, 18(24): 7 213~ 7 218.
- [3] 杜立新, 曹顶国. 小尾寒羊微卫星和 RAPD 标记的研究[J]. *遗传学报*, 2003, 30(11): 1 041~ 1 044.
- [4] 王 杰, 徐金瑞, 白文林, 等. 四川几个山羊品种(群体)与岩羊 RAPD 分析[J]. *畜牧兽医学报*, 2003, 34(5): 434~ 437.
- [5] Naqvi N I, Chattoo B B. Development of a sequence characterized amplified region (SCAR) based indirect selection method for a dominant blast-resistance gene in rice[J]. *Genome*, 1996, 39(1): 26~ 30.
- [6] Hernandez P, Dorado G, Ramirez M C, et al. Development of cost-effective Hordeum chilense DNA markers: molecular aids for marker-assisted cereal breeding [J]. *Hereditas*, 2003, 138(1): 54~ 58.
- [7] Gunter L E, Roberts G T, Lee K, et al. The development of two flanking SCAR markers linked to a sex determination locus in *Salix viminalis* L [J]. *J Hered*, 2003, 94(2): 185~ 189.
- [8] Kethidi D R, Roden D B, Ladd T R, et al. Development of SCAR markers for the DNA-based detection of the Asian long-horned beetle, *Anoplophora glabripennis* (Motschulsky) [J]. *Arch Insect Biochem Physiol*, 2003, 52(4): 193~ 204.
- [9] Fernandez S, Costa A C, Katsuyama A M, et al. A

- survey of the inter- and intraspecific RAPD markers of *Eimeria* spp. of the domestic fowl and the development of reliable diagnostic tools[J]. *Parasitol Res*, 2003, 89(6): 437~ 445.
- [10] Chatterjee B, Bahadur M, Srivastava S, et al. Differential organization of a LINE-1 family in Indian pygmy field mice[J]. *Indian J Exp Biol*, 2003, 41(1): 53~ 57.
- [11] Perepelitsa Belancio V, Deininger P. RNA truncation by premature polyadenylation attenuates human mobile element activity[J]. *Nat Genet*, 2003, 35(4): 363~ 366.
- [12] Perez M J, Leroux C, Bonastre A S, et al. Occurrence of a LINE sequence in the 3' UTR of the goat alpha s1-casein E-encoding allele associated with reduced protein synthesis level[J]. *Gene*, 1994, 147(2): 179~ 187.

### Screening of the Genetic Marker for the Milk Yield Trait in Chinese Holstein by RAPD Method

DU Lixin, WAN Haiwei, WANG Aihua, LI Hongbin

(*Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100094, China*)

**Abstract:** Two bovine genomic DNA pools were composed of 36 high-milk-yield Chinese Holstein cattles (milk yield in 305 days > 8 500kg) and 32 low-milk yield Chinese Hosten cattles (milk yield < 5 600kg). 24 high polymorphic RAPD primers selected from 320 arbitrary primers were used for analyzing genetic difference between two DNA pools. 5 specific fragments were obtained from DNAs of two groups cattles. 4 of them were sequenced and converted to SCAR markers. The results of statistic analysis showed that SCAR marker yield-S139 was highly associated with milk yield trait. By electronic cloning method the marker was extended from 921 bp to 2 141 bp. As a result, the fragment was a repeat element of LINE family in bovine genome by homologous analysis. So it was reasonably to assume that some QTLs or major genes associated with milk yield trait maybe lied in this LINE region.

**Key words:** dairy cattle; milk yield trait; molecular marker

## 中国畜牧业交易会 ——中国畜牧业高端论坛

经农业部批准,中国畜牧业协会将于2005年11月4-8日在山东青岛组织召开2005中国畜牧业交易会。

为了更好地为企业服务、了解企业心声、传达政府相关政策;使企业了解最新的科研成果与行业动态、协助企业探寻市场商机、展示企业风采与市场运作成就;加快协调畜牧业产、学、研、运、销、管等多方面关系以应对市场变化,为政府相关部门制定发展规划和产业决策提供科学依据,推动畜牧行业的横向联合,共同打造畜牧业科研开发与经济运作的平台,在“2005中国畜牧业交易会”期间特举办“中国畜牧业高端论坛”,拟邀请政府相关高层领导,畜牧及相关行业企业家、专家、市场销售总监等权威人士,针对行业的难点、热点、焦点问题,开展面对面的互动式交流与探讨,搭建全国畜牧及相关行业工作者的交流平台,为畜牧及相关行业上、中、下游企业合作提供商机、展示畜牧产业各生产领域及环节企业风采,在国家实施宏观调控和进一步落实中央一号文件精神的新形势下,全面推进中国现代畜牧业。