

# 甘蓝 EST-SSR 标记的开发与应用

陈琛<sup>1,2</sup>, 庄木<sup>1,\*</sup>, 李康宁<sup>3</sup>, 刘玉梅<sup>1</sup>, 杨丽梅<sup>1</sup>, 张扬勇<sup>1</sup>, 程斐<sup>2</sup>,  
孙培田<sup>1</sup>, 方智远<sup>1</sup>

(<sup>1</sup> 中国农业科学院蔬菜花卉研究所, 北京 100081; <sup>2</sup> 青岛农业大学, 山东青岛 266109; <sup>3</sup> 内蒙古农业大学, 呼和浩特 010081)

**摘要:** 从 NCBI 数据库下载 62 567 条甘蓝相关 EST, 处理后得到 19 611 条无冗余的 EST, 对其进行 SSR 搜索, 共得到 1 219 条 SSR, 分布在 1 176 条 EST 上, 分布密度为 1/11.48 kb, 包括 273 种重复基元。在甘蓝 EST-SSR 中, 二核苷酸(353 个 SSR)和三核苷酸(423 个 SSR)占主导地位, 所占比例分别为 28.96% 和 34.70%; 出现最多的重复基元是 AG/CT, 占总数的 25.59%, 其次为 AAG/CTT (94 个, 占 7.71%)。针对 1 176 条含有 SSR 的 EST 设计合成了 978 对引物。用两份甘蓝自交系对 978 对引物进行初筛, 897 对引物有扩增产物, 共扩增出 1 026 条扩增带; 128 对引物表现出多态性, 共有 258 条多态性带 (占总带数的 25.15%)。利用 4 对多态性引物, 初步构建了中甘 11、8398、中甘 15、中甘 21 杂交新品种及其亲本的指纹图谱。以上结果说明从甘蓝相关 EST 数据库中开发的 SSR 引物有较好的可用性。

**关键词:** 甘蓝; EST-SSR; 标记开发; 指纹图谱

**中图分类号:** S 635

**文献标识码:** A

**文章编号:** 0513-353X(2010)02-0221-08

## Development and Utility of EST-SSR Marker in Cabbage

CHEN Chen<sup>1,2</sup>, ZHUANG Mu<sup>1,\*</sup>, LI Kang-ning<sup>3</sup>, LIU Yu-mei<sup>1</sup>, YANG Li-mei<sup>1</sup>, ZHANG Yang-yong<sup>1</sup>,  
CHENG Fei<sup>2</sup>, SUN Pei-tian<sup>1</sup>, and FANG Zhi-yuan<sup>1</sup>

(<sup>1</sup>Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China; <sup>2</sup>Qingdao Agricultural University, Qingdao, Shandong 266109, China; <sup>3</sup>Inner Mongolia Agricultural University, Huhhot 010081, China)

**Abstract:** A total of 62 567 ESTs in *Brassica oleracea* were downloaded from NCBI for the identification and development of SSR marker. In total, 1 219 SSRs distributed in 1 176 ESTs were identified from 19 611 non-redundant ESTs, with an average of one SSR per 11.48 kb, and included 273 SSR motifs. Analysis of SSR motifs revealed that the dinucleotides (353 SSR) and tricleotides (423 SSR) were major motifs, accounted for 28.96% and 34.70% respectively. The AG/CT was the most frequent motifs and accounted for 25.59%, followed by AAG/CTT (94 SSR, 7.71%). Based on the 1 176 SSR-containing ESTs, a total of 978 primer pairs were successfully designed and assessed validation of the amplification using two cabbage inbred lines (397 and 20-2-5). The results showed that 897 primer pairs yielded 1 026 amplification bands, of which 128 primer pairs exhibited polymorphism with 258 polymorphic bands accounted for 25.15%. SSR fingerprint was obtained for 4 cabbage F<sub>1</sub>, Zhonggan 11, 8398, Zhonggan 15 and Zhonggan 21, and their

**收稿日期:** 2009-12-07; **修回日期:** 2010-01-25

**基金项目:** 国家‘十一五’科技支撑计划项目 (2008BADB1B02); 中央级公益性科研院所基本科研业务费专项; 国家自然科学基金项目 (30700543); 国家‘863’计划项目 (2007AA10Z174, 2006AA100108); 农业部园艺作物遗传改良重点开放实验室项目

\* 通信作者 Author for correspondence (E-mail: zhuangmu@mail.caas.net.cn)

parents, and all the above materials could be discriminated by four specific EST-SSR primer pairs. The results indicated that the developed SSRs from ESTs of *Brassica oleracea* were validate and practicable.

**Key words:** cabbage; EST-SSR; marker development; fingerprinting

简单重复序列 (Simple sequence repeat, SSR) 标记技术自 1989 年建立以来, 由于多态性高, 检测容易, 数量丰富, 共显性和特异性等特点 (Powell et al., 1996), 已被广泛用于遗传图谱绘制、重要农艺性状的基因定位、生物遗传多样性、遗传进化、品种鉴定等诸多方面 (韦祖生 等, 2008)。构建 DNA 指纹图谱是 SSR 分子标记技术利用的一个重要方面, 目前在油菜 (许鲲 等, 2008)、马铃薯 (Provan et al., 1996)、番茄 (栾雨时 等, 1999)、大白菜 (李丽 等, 2009) 等多种作物已有报道。根据 SSR 标记的序列性质不同, SSR 标记可分为基因组 SSR (genomic SSR, gSSR) 和表达序列标签 SSR (expressed sequence tag SSR, EST-SSR)。由于 gSSR 分子标记开发需要构建基因组文库、制备探针和分子杂交等过程, 制约了其发展和应用。随着 GenBank 中许多重要物种的 EST 急剧增加, 使其成为一种新的和有效的 SSR 标记来源 (Cordeiro et al., 2001)。EST-SSR 来自基因的转录区, 与功能基因有紧密的连锁, 且在物种间有较高的通用性, 其开发与应用已在小麦、棉花、大豆、花生、番茄、油菜和大白菜等作物中报道。

鉴于目前公开报道可利用的甘蓝 SSR 标记有限 (<http://www.brassica.info>), 本研究中基于 NCBI 公共数据库公开报道的 6 万余条甘蓝相关 EST, 开发 EST-SSR 标记并分析其多态性, 为甘蓝类蔬菜作物及其它芸薹属蔬菜的种质资源分析、指纹图谱、遗传作图和分子标记辅助育种等提供理论依据。

## 1 材料与方法

### 1.1 材料和模板 DNA 的提取

以早熟圆球的 01-20 和晚熟扁圆球的 20-2-5 两份甘蓝自交系为初筛引物材料, 检测引物的扩增与多态性。以北早 (02-12)、中甘 11、金早生 (01-88)、084、中甘 15、01-20、8398、79-156、87-534、中甘 21、99Z521-12 等 11 份甘蓝材料 (编号依次为 1~11), 构建甘蓝杂交新品种中甘 11、8398、中甘 15 和中甘 21 的 DNA 指纹图谱。中甘 11 的亲本是北早 (02-12) (来自中国地方品种) 和金早生 (01-88) (来自加拿大), 其叶球近圆形, 外叶色深绿, 从定植到成熟约 45 d; 中甘 15 的亲本是 01-20 (来自北京早熟的自交系) 和 084 (来自 Golden Arce), 其叶球近圆形, 外叶色绿, 从定植到成熟约 55 d; 8398 的亲本是 01-20 (来自北京早熟的自交系) 和 79-156 (来自丹麦), 其叶球近圆形, 外叶色绿, 从定植到成熟约 50 d; 中甘 21 的亲本是 99Z521-12 (来自北京早熟的显性不育系) 和 87-534 (来自德国), 其叶球圆球形, 外叶色绿, 从定植到成熟约 50 d。所有试验材料均由中国农业科学院蔬菜花卉所甘蓝组提供, 于 2008 年春季种植于本所试验田, 在定植缓苗后苗期取幼嫩叶片采用改良 CTAB 法提取 DNA。

### 1.2 甘蓝 EST 来源及 EST-SSR 的发掘

甘蓝相关 EST 来自 NCBI (美国国家生物技术信息中心) 数据库 (<http://www.ncbi.nlm.nih.gov/dbEST>), 共计 62 567 条 (截止 2008 年 6 月)。

从数据库直接获取的 EST 中包含一些低质量片段 (<100 bp), 同时存在着带有少量载体序列及末端存在 polyA/T “尾巴” 的序列, 影响相关信息分析, 需要进行前处理, 除去 5' 端或 3' 端 50 bp 的 polyT 或 polyA, 直到在 50 bp 大小窗口看到的 5' 端和 3' 端 polyA 和 polyT 不超过 (T)<sub>5</sub> 或 (A)<sub>5</sub>, 得到高质量的 EST。

EST 是随机选取测序的, 因此导致同一基因重复测序的冗余现象不可避免。前处理后的 EST 通过

软件 Phrap 进行片段重叠群分析和聚类。拼接时设定的初始装配参数为: 最小匹配碱基数 (minmatch) 为 20, 最小分值 (minscore) 为 40。

利用 SSRIT 在线分析软件 (<http://www.gramene.org/db/searches/ssrtool>) 对聚类后的无冗余 EST 序列进行 SSR 搜索。SSR 筛选标准为: 单核苷酸, 二、三、四、五和六核苷酸的最少重复次数分别为 18 次、9 次、6 次、5 次、4 次、3 次以上, 所有重复类型的总核苷酸数不能少于 18 个。

### 1.3 EST-SSR 的引物设计

利用 DNASTar 的 Primer select 程序对含有 SSR 的 EST 设计引物, 并由上海生工生物工程技术服务有限公司合成。引物设计的主要参数: 引物长度控制在 17~24 bp, 预期产物长度控制在 150~350 bp。

### 1.4 SSR 分析

PCR 反应体系为 20  $\mu\text{L}$ , 其中包括: 200  $\mu\text{mol} \cdot \text{L}^{-1}$  dNTPs, 上链和下链引物各 4 pmol, 1.0 U *Taq* 酶, 1  $\times$  Buffer (内含 1.5  $\text{mmol} \cdot \text{L}^{-1}$   $\text{MgCl}_2$ ), 50 ng DNA 模板, 所用试剂均购自博迈德公司。PCR 扩增反应条件是: 94  $^{\circ}\text{C}$  预变性 4 min; 94  $^{\circ}\text{C}$  变性 30 s, 合适温度退火 30 s, 72  $^{\circ}\text{C}$  延伸 1 min, 30 个循环; 72  $^{\circ}\text{C}$  延伸 7 min。各引物的合适退火温度根据引物合成报告单而定。8%聚丙烯酰胺凝胶电泳分离 PCR 产物, 160 V 恒压。快速银染法染色。在相同迁移率位置上, 有带记为 1, 无带记为 0, 缺失记为 9。为保证数据准确、可靠, 每块胶板均由两人独立记录, 然后比对, 有争议的数据作缺失处理。

## 2 结果与分析

### 2.1 甘蓝 EST-SSR 出现的频率和特点

从 NCBI 下载 62 567 条甘蓝类作物 EST, 经过处理后得到 19 611 条无冗余 EST, 总长度为  $1.4 \times 10^7$  bp, 经搜索共挖掘到 1 219 个 SSR, 分布在 1 176 条无冗余的 EST 上, 发生频率 6.0% (含有 SSR 的 EST 与无冗余 EST 数之比), 平均分布距离为 11.5 kb。

表 1 甘蓝 EST 中 SSR 出现频率  
Table 1 Occurrence of SSRs in ESTs of *Brassica oleracea*

| 重复类型<br>Type of repeat | 基元种数<br>Number of motif type | 占总数的比例/%<br>Proportion in all motifs | SSR 数<br>SSR number | 占全部 SSR 比例/%<br>Proportion in all SSRs | 平均距离/bp<br>Mean distance |
|------------------------|------------------------------|--------------------------------------|---------------------|--|--------------------------|
| 单核苷酸 Mononucleotide    | 2                            | 0.73                                 | 106                 | 8.70                                   | $1.32 \times 10^5$       |
| 二核苷酸 Dinucleotide      | 4                            | 1.47                                 | 353                 | 28.96                                  | $4.00 \times 10^4$       |
| 三核苷酸 Trinucleotide     | 18                           | 6.59                                 | 423                 | 34.70                                  | $3.31 \times 10^4$       |
| 四核苷酸 Tetranucleotide   | 17                           | 6.23                                 | 24                  | 1.97                                   | $5.83 \times 10^5$       |
| 五核苷酸 Pentanucleotide   | 20                           | 7.33                                 | 24                  | 1.97                                   | $5.83 \times 10^5$       |
| 六核苷酸 Hexanucleotide    | 182                          | 66.67                                | 251                 | 20.59                                  | $5.58 \times 10^4$       |
| 七核苷酸 Seven nucleotide  | 22                           | 8.06                                 | 29                  | 2.38                                   | $4.83 \times 10^5$       |
| 八核苷酸 Eight nucleotide  | 6                            | 2.20                                 | 7                   | 0.57                                   | $2.00 \times 10^6$       |
| 九核苷酸 Nine nucleotide   | 2                            | 0.73                                 | 2                   | 0.16                                   | $7.00 \times 10^6$       |
| 合计 Total               | 273                          | 100                                  | 1219                | 100                                    | $1.15 \times 10^4$       |

注: 平均距离=无冗余 EST 总长度 / SSR 数量。

Note: Mean distance = Total length of non-redundant ESTs / SSR numbers.

从表 1 可以看出, 在所有甘蓝 EST-SSR 中, 二核苷酸和三核苷酸为优势重复类型, 共占 EST-SSR 总数的 63.66%, 三核苷酸重复类型所占的比例最大 (423 个, 占 SSR 总数的 34.70%), 二核苷酸重复类型次之 (353 个, 占 28.96%), 其余重复类型按所占比例高低依次为六核苷酸重复 (20.59%)、单核苷酸重复 (8.70%)、七核苷酸重复 (2.38%), 四核苷酸重复和其它重复类型所占比例均小于 2%。

从 SSR 分布的密度来看, 甘蓝 EST 中平均每 11.5 kb 就出现一个 SSR, 不同重复类型的 SSR 覆盖的平均距离不一样, SSR 数越多其覆盖的平均距离越小 (表 1)。

共观察到 273 种重复基元 (表 2 为比例大于 1% 的重复基元), 出现最多的重复基元是 AG/CT (312 个, 占总数的 25.59%), 其次是 AAG/CTT (94 个, 占 7.71%) 和 A/T (90 个, 占 7.38%), 其余重复类型所占比例较小 (均不足 5%)。

二至六核苷酸平均基元长度变化范围差异很大, 二核苷酸的长度变化范围最大 (18~102 bp), 五核苷酸长度范围最小 (20~30 bp); 从平均长度看, 二核苷酸最长 (24.13 bp), 其次是五核苷酸 (23.83 bp)、四核苷酸 (22.67 bp)、六核苷酸的 (21.25 bp)、三核苷酸 (20.96 bp) (表 3)。

2.2 甘蓝 EST-SSR 标记的扩增分析

对搜索得到的 1 219 个甘蓝类作物 EST-SSR, 应用 DNASTar 中的 Primer select 程序设计并合成了 978 对引物 (命名为 BoE001 ~ 978), 以两份熟性和球形不同的甘蓝材料 (397 和 20-2-5) 对所获得引物检测。结果表明: (1) 81 对引物没有扩增产物, 可能是由于模板 DNA 来自基因组包含内含子, 或所设计的引物跨越 mRNA 剪切位点等原因造成的; (2) 897 对引物有扩增带, 占总引物的 91.82%, 共扩增出 1 026 条带, 其中 128 对引物有多态性, 共产生 258 条多态性带, 多态性引物占总引物的 13.09%, 表明这套引物对甘蓝材料扩增是有效的。

2.3 4 个甘蓝新品种 SSR 指纹图谱的构建

从两份甘蓝材料 (397 和 20-2-5) 表现多态性的 128 对引物中, 选择 55 对条带清晰、非特异扩增带较少的引物, 对中甘 11、8398、中甘 15、中甘 21 及各自的亲本等 11 份甘蓝材料进行扩增检测。结果获得 BoE002、BoE222、BoE225 和 BoE916 等 4 对多态性较好的引物 (表 4), 扩增产物条带 3~5

表 2 甘蓝 EST 中主要重复基元  
Table 2 The major repeat motifs in ESTs of *Brassica oleracea*

| 重复类型<br>Repeat type    | 重复基元<br>Repeat motif | 数量<br>Number | 发生频率/%<br>Frequency | 所占比例/%<br>Proportion |
|------------------------|----------------------|--------------|---------------------|----------------------|
| 单核苷酸<br>Mononucleotide | A/T                  | 90           | 0.46                | 7.38                 |
|                        | C/G                  | 16           | 0.08                | 1.31                 |
| 二核苷酸<br>Dinucleotide   | AC/GT                | 16           | 0.08                | 1.31                 |
|                        | AG/CT                | 312          | 1.59                | 25.59                |
|                        | AT/TA                | 24           | 0.12                | 1.97                 |
| 三核苷酸<br>Trinucleotide  | AAC/GTT              | 29           | 0.15                | 2.38                 |
|                        | AAG/CTT              | 94           | 0.48                | 7.71                 |
|                        | ACC/GGT              | 18           | 0.09                | 1.48                 |
|                        | AGA/TCT              | 50           | 0.25                | 4.10                 |
|                        | AGC/GCT              | 23           | 0.12                | 1.89                 |
|                        | AGG/CCT              | 43           | 0.22                | 3.53                 |
|                        | AGT                  | 21           | 0.11                | 1.72                 |
|                        | ATC/GAT              | 31           | 0.16                | 2.54                 |
|                        | ATG/CAT              | 17           | 0.09                | 1.39                 |
|                        | CAG/CTG              | 17           | 0.09                | 1.39                 |

注: 比例小于 1% 的基元未列出。

Note: The repeat motifs with the percentage less than 1% were not listed here.

表 3 甘蓝 EST-SSR 基元重复长度  
Table 3 The repeat length of SSRs in ESTs of *Brassica leracea*. /bp

| 重复类型<br>Repeat type     | 变化范围<br>Variation | 平均长度<br>Average |
|-------------------------|-------------------|-----------------|
| 二核苷酸<br>Dinucleotide    | 18 ~ 102          | 24.13           |
| 三核苷酸<br>Trinucleotide   | 18 ~ 75           | 20.96           |
| 四核苷酸<br>Tetranucleotide | 20 ~ 36           | 22.67           |
| 五核苷酸<br>Pentanucleotide | 20 ~ 30           | 23.83           |
| 六核苷酸<br>Hexanucleotide  | 18 ~ 42           | 21.25           |

表 4 构建甘蓝指纹图谱的 4 对 EST-SSR 引物  
Table 4 4 EST-SSR primers used for fingerprinting in cabbage

| 引物<br>Primer | 重复基元<br>Repeat motifs | EST 编号<br>EST code | 上链引物 (5'-3')<br>Forward primer (5'-3') | 下链引物 (5'-3')<br>Reverse primer (5'-3') | 退火温度/<br>℃ Tm | 产物大小/bp<br>Product size |
|--------------|-----------------------|--------------------|--|--|---------------|-------------------------|
| BoE002       | (TCA)7                | 150909620          | CGTCACGGTGGCGCTTTATTTT                 | TAGCGGCAGCGACGTGGAGAAC                 | 60.0          | 284                     |
| BoE222       | (CGATGA)3             | 150914869          | TTGCTGAAGAAGTTGGGATTGAGG               | CGGCTTTTCTTTTCCTTCTGC                  | 55.0          | 184                     |
| BoE225       | (ATC)7                | 150918267          | ATCTGATGTGGAGCCTGTTGTT                 | CCTTACGATCTGGGAGTGAAAA                 | 50.0          | 171                     |
| BoE916       | (TCTCAT)3             | 75973414           | AACTGAGACTGAAAAAGACAAACA               | TTCAGCATCAATCAACTTATCAAT               | 49.3          | 166                     |

不等，分子量 140~280 bp，共扩增出多态性带 16 条，扩增重复性好（图 1）。利用这 4 对引物初步构建了中甘 11、8398、中甘 15 和中甘 21 等 4 个甘蓝新品种的 SSR 指纹图谱（图 2）。

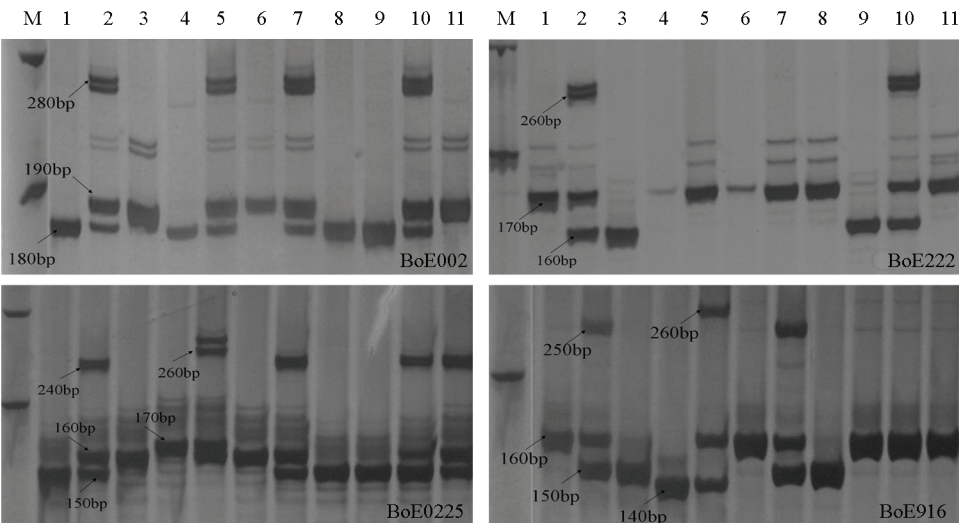


图 1 4 个 SSR 引物对 4 个甘蓝杂交新品种及其亲本的扩增多态性结果  
1~11 依次为甘蓝材料北早，中甘 11、金早生、084、中甘 15、01-20、8398、79-156、87-534、中甘 21 和 99Z521-12。

Fig.1 Amplified bands of four polymorphic primers in 4 cabbage hybrids and their parents  
1~11 are cabbage samples Beizao, Zhonggan 11, Jinzaosheng, 084, Zhonggan 15, 01-20, 8398, 79-156, 87-534, Zhonggan 21 and 99Z521-12 respectively.

|              | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|--------------|---|---|---|---|---|---|---|---|---|----|----|
| BoE002-280bp | — | + | — | — | + | — | + | — | — | +  | —  |
| BoE002-190bp | — | + | + | — | + | + | + | — | — | +  | +  |
| BoE002-180bp | + | + | — | + | + | — | + | + | + | +  | —  |
| BoE222-260bp | — | + | — | — | — | — | — | — | — | +  | —  |
| BoE222-170bp | + | + | — | + | + | + | + | + | — | +  | +  |
| BoE222-160bp | — | + | + | — | — | — | — | — | + | +  | —  |
| BoE225-260bp | — | — | — | — | + | — | — | — | — | —  | —  |
| BoE225-240bp | — | + | — | — | — | — | + | — | — | +  | +  |
| BoE225-170bp | — | — | — | + | + | — | — | — | — | —  | —  |
| BoE225-160bp | — | + | + | — | + | + | + | — | — | +  | +  |
| BoE225-150bp | + | + | — | — | — | — | + | + | + | +  | +  |
| BoE916-260bp | — | — | — | — | + | — | — | — | — | —  | —  |
| BoE916-250bp | — | + | — | — | — | — | + | — | — | —  | —  |
| BoE916-160bp | + | + | — | — | + | + | + | — | + | +  | +  |
| BoE916-150bp | — | + | + | — | — | — | + | + | — | —  | —  |
| BoE916-140bp | — | — | — | + | + | — | — | — | — | —  | —  |

图2 4个甘蓝杂交种及其亲本的SSR指纹图谱  
+表示有泳带；—表示无泳带；1~11分别对应甘蓝材料北早、中甘11、金早生、084、中甘15、01-20、8398、79-156、87-534、中甘21和99Z521-12。

Fig.2 SSR Fingerprint of 4 cabbage hybrids and their parents  
+: bands; —: No bands; Lanes 1~11 represent the cabbage samples Beizao, Zhonggan 11, Jinzaosheng, 084, Zhonggan 15, 01-20, 8398, 79-156, 87-534, Zhonggan 21 and 99Z521-12 respectively.

### 3 讨论

#### 3.1 甘蓝 EST-SSR 的分布特征

对 19 611 条无冗余甘蓝相关 EST 进行搜索, 在 1 176 条 EST 上挖掘到 1 219 个 SSR, 含 SSR 序列占无冗余 EST 序列总数的 6.0%, 高于大白菜 10.4% (葛佳 等, 2005), 与 EST 文库中大约高达 2% 至 11% 的 EST 序列中含有 SSR 的推测相符合 (Cordeiro et al., 2001)。甘蓝 EST-SSR 的基元类型丰富 (273 种), 频率高, 密度较大 (平均距离 11.48 kb), 高于拟南芥 (14.9 kb)、小麦 (15.6 kb)、棉花 (20.0 kb) (李永强 等, 2004), 水稻 (11.8 kb)、玉米 (28.3 kb) 和大豆 (23.8 kb) (Gao et al., 2003), 但低于大白菜 (10.3 kb) 和油菜 (4.34 kb) (葛佳 等, 2005; 李小白 等, 2007)。这种差异与物种的真实 SSR 信息 (本研究中近半数 EST 来自甘蓝类蔬菜中的芥蓝)、EST 的数量、搜索 SSR 软件的算法与参数设定等多种因素有关。

在甘蓝 EST-SSR 中, 最多的重复类型是三核苷酸 (423 个 SSR), 占总数的 34.70%, 这与已报道的甘蔗、棉花、水稻、玉米、小麦、大白菜、番茄和拟南芥的结果 (Cardle et al., 2000; 王静毅等, 2008) 一致。AAG/CTT (94 个, 占 EST-SSR 总数的 7.71%) 是甘蓝三核苷酸重复中出现频率最高的类型, 这与大白菜 (葛佳 等, 2005)、油菜 (李小白 等, 2007)、拟南芥 (Cardle et al., 2000)、大豆 (Gao et al., 2003) 等中的情形相一致, 本研究结果进一步验证了在双子叶植物中 AAG/CTT 重复丰度很高的推测 (Gao et al., 2003)。

在甘蓝 EST-SSR 所有的基元重复类型中, AG/CT (312 个, 占总数的 25.59%) 所占比例最高。这与大白菜 (84.2%)、大麦 (52%) 和小麦 (45%) 的结果 (Thiel et al., 2003; Nicot et al., 2004) 近似, 也与 Morgante 等 (2002) 对多种不同植物综合分析后所报道的结果相符, 但与猕猴桃中频率最高的二核苷酸重复序列为 CT (Fraser et al., 2004) 和番茄中 AT 为频率最高 (Areshchenkova & Ganai, 2002) 的结果不一致。GC 重复基元在多数植物中很难见到或稀有, 在甘蓝二核苷酸重复中 GC 仅出现了一次, 在大白菜二核苷酸重复中无 GC (葛佳 等, 2005), 小麦的报道中也得到了类似的结果 (Nicot et al., 2004)。

#### 3.2 甘蓝 EST-SSR 标记的应用性

本研究对获得的 978 对甘蓝 EST-SSR 引物, 用 397 和 20-2-5 进行扩增初筛时, 91.82% 的引物有扩增带, 其中 13.09% 的引物在 397 和 20-2-5 间表现多态性。EST 的高度保守性是导致其多态性较低的重要因素 (Zhang et al., 2005), 对无多态性的 EST-SSR 扩增产物可先用不同的限制性内切酶处理, 然后电泳分离酶切产物, 利用 SSR-CAPs 方法来提高 EST-SSR 引物的多态性 (McCallum et al., 2001; Iwata et al., 2001)。

以 SSR 分子标记技术构建的 DNA 指纹图谱已广泛的应用于玉米 (Smith, 1997)、小麦 (李根英 等, 2006)、油菜 (许鲲 等, 2008)、马铃薯 (Provan et al., 1996)、番茄 (栾雨时 等, 1999)、大白菜 (李丽 等, 2009) 等多种作物。中甘 11、8398、中甘 15 和中甘 21 是我国当前春甘蓝栽培中广泛应用的主栽品种, 本研究通过对多态性甘蓝 EST-SSR 引物的扩增筛选, 仅用 4 对引物就能区分这 4 个甘蓝新品种及其各自的亲本, 初步建立了它们的 EST-SSR 指纹图谱。引物的多态性信息量 (PIC) 是衡量引物适合程度的重要指标, 反映引物揭示信息量的多少, 越高包含的信息量越大 (许鲲 等, 2008)。构建甘蓝新品种指纹图谱的引物 PIC 值由高到低的顺序为 BoE225 (0.69)、BoE916 (0.65)、BoE0002 (0.57)、BoE222 (0.48), 平均 PIC 值为 0.60, 进一步表明这套引物是有效的。

EST-SSR 来自基因的编码区, 与功能基因有紧密的连锁, 这有可能对决定重要表型性状的等位基因进行直接鉴定 (Chen et al., 2001)。此外, EST 的高度保守性使得 EST-SSR 且在物种间有较高的通

用性 (Gupta et al., 2003; 忻雅 等, 2006)。本研究开发的甘蓝 EST-SSR 标记, 已用于大白菜重要性状的遗传连锁图谱构建 (汪骞 等, 2009), 其在甘蓝类蔬菜作物及其它芸薹属蔬菜的种质资源分析、遗传作图和分子标记辅助育种等应用研究正在进行中。

## References

- Areshchenkova T, Ganai M W. 2002. Comparative analysis of polymorphism and chromosomal location of tomato microsatellite markers isolated from different sources. *Theor Appl Genet*, 104: 229 – 235.
- Cardle L, Ramsay L, Milbourne D, Macaulay M, Marshall D, Waugh R. 2000. Computational and experimental characterization of physically clustered simple sequence repeats in plants. *Genetics*, 156: 847 – 854.
- Chen X, Salamini F, Gebhardt C. 2001. A potato molecular-function map for carbohydrate metabolism and transport. *Theor Appl Genet*, 102(2): 284 – 295.
- Cordeiro G M, Casu R, McIntyre C L, Manners J M, Henry R J. 2001. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Science*, 160: 1115 – 1123.
- Fraser L G, Harvey C F, Crowhurst R N, De Silva H N. 2004. EST-derived microsatellites from *Actinidia* species and their potential for mapping. *Theor Appl Genet*, 108: 1010 – 1016.
- Gao Li-feng, Tang Ji-feng, Li Hong-wei, Jia Ji-zeng. 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Molecular Breeding*, 12: 245 – 261.
- Ge Jia, Xie Hua, Chui Chong-shi, Hong Jian-ming, Ma Rong-cai. 2005. Analysis of expressed sequence tags (ESTs) derived SSR markers in Chinese cabbage (*Brassica campestris* L. ssp. *pekinensis*). *Journal of Agricultural Biotechnology*, 13 (4): 423 – 428. (in Chinese)
- 葛 佳, 谢 华, 崔崇士, 洪剑明, 马荣才. 2005. 大白菜表达序列标签SSR标记分析. *农业生物技术学报*, 13 (4): 423 – 428.
- Gupta P K, Rustgi S, Sharma S, Singh R, Kumar N, Balyan H S. 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol Genet Genomics*, 270 (4): 15 – 23.
- Iwata H, Ujino-Ihara T, Yoshimura K, Nagasaka K, Mukai Y, Tsumura Y. 2001. Cleaved amplified polymorphic sequence markers in sugi, *Cryptomeria japonica* D. Don, and their locations on a linkage map. *Theor Appl Genet*, 103: 881 – 895.
- Li Gen-ying, Susanne Dreisigacker, Marilyn L. Warburton, Xia Xian-chun, He Zhong-hu, Sun Qi-xin. 2006. Development of a fingerprinting database and assembling an SSR reference kit for genetic diversity analysis of wheat. *Acta Agronomica Sinica*, 32(12): 1771 – 1778. (in Chinese)
- 李根英, Susanne Dreisigacker, Marilyn L. Warburton, 夏先春, 何中虎, 孙其信. 2006. 小麦指纹图谱数据库的建立及SSR分子标记试剂盒的研发. *作物学报*, 32(12): 1771 – 1778.
- Li Li, He Wei-ming, Ma Lian-ping, Liu Pang-yuan, Xu Hai-ming, Xu Jia-bing, Zheng Xiao-ying. 2009. Construction chinese cabbage (*Brassica rapa* L.) core collection and its EST-SSR fingerprint database by EST-SSR molecular markers. *Genomics and Applied Biology*, 28 (1): 76 – 88. (in Chinese)
- 李 丽, 何伟明, 马连平, 刘庞源, 徐海明, 徐家柄, 郑晓鹰. 2009. 用EST-SSR分子标记技术构建大白菜核心种质及其指纹图谱库. *基因组学与应用生物学*, 28 (1): 76 – 88.
- Li Xiao-bai, Zhang Ming-long, Cui Hai-rui. 2007. Analysis of SSR information in EST resource of oilseed rape. *Chinese Journal of Oil Crop Sciences*, 29 (1): 20 – 25. (in Chinese)
- 李小白, 张明龙, 崔海瑞. 2007. 油菜EST资源的SSR信息分析. *中国油料作物学报*, 29 (1): 20 – 25.
- Li Yong-qiang, Li Hong-wei, Gao Li-feng, He Bei-ru. 2004. Progress of simple sequence repeats derived from expressed sequence tags. *Journal of Plant Genetic Resources*, 5(1): 91 – 95. (in Chinese)
- 李永强, 李宏伟, 高丽锋, 何蓓如. 2004. 基于表达序列标签的微卫星标记 (EST-SSRs) 研究进展. *植物遗传资源学报*, 5(1): 91 – 95.
- Luan Yu-shi, An Li-jia, Huang Bai-qu, Zhang Wen-jun, He Meng-yuan. 1999. Using simple sequence repeat probe for DNA fingerprints in tomato cultivars. *Acta Horticulturae Sinica*, 26(1): 51 – 53. (in Chinese)
- 栾雨时, 安利佳, 黄百渠, 张文俊, 何孟元. 1999. 用SSR探针进行番茄品种的DNA指纹分析. *园艺学报*, 26(1): 51 – 53.

- McCallum J, D Leite, M Pither-Joyce, Havey M J. 2001. Expressed sequence markers for genetic analysis of bulb onion (*Allium cepa* L.). *Theor Appl Genet*, 103: 979 - 991.
- Morgante M, Hanafey M, Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nature Genetics*, 30: 194 - 200.
- Nicot N, Chiquet V, Gandon B, Amilhat L, Legeai F, Leroy P, Bernard M, Sourdille P. 2004. Study of simple sequence repeat (SSRs) markers from wheat expressed sequence tags (ESTs). *Theor Appl Genet*, 109: 800 - 805.
- Powell W, Machray G C, Provan J. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Science*, 7: 215 - 222.
- Provan J, Powell W, Waugh R. 1996. Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). *Theor Appl Genet*, 92: 1078 - 1084.
- Smith J S C, Chin E C L, Shu H, Smith O S, Wall S J, Senior M L, Mitchell S E, Kresovich S, Ziegler J. 1997. An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. *Theor Appl Genet*, 95: 163 - 173.
- Thiel T, Michalek W, Varshney R K, Graner A. 2003. Exploiting EST database for the development and characterization of gene-derived SSR markers in barley (*Hordeum vulgare* L.), *Theor Appl Genet*, 106: 411 - 422.
- Wang Jing-yi, Chen Ye-yuan, Liu Wei-liang, Wu Yao-ting. 2008. Development and application of EST-derived SSR markers for bananas (*Musa nana* Lour.), *Hereditas*, 30(7): 933 - 940. (in Chinese)
- 王静毅, 陈业渊, 刘伟良, 武耀廷. 2008. 香蕉EST-SSRs的标记与开发, *遗传*, 30 (7): 933 - 940.
- Wang Qian, He Jiang-ming, Lin Liang-bin, Zhuang Mu, Wang Yan, Wang Xiao-wu, Wu Jian. 2009. the Development of backcross introgression lines (BILs) and genetic analysis for *Brassica campestris*. *Acta Horticulturae Sinica*, 36 (9): 1305 - 1310. (in Chinese)
- 汪 骅, 和江明, 林良斌, 庄 木, 王 艳, 王晓武, 武 剑. 2009. 大白菜回交导入系群体构建及遗传分析. *园艺学报*, 36 (9): 1305 - 1310.
- Wei Zu-sheng, Xia Zhi-qiang, Li Kai-mian, Wang Wen-quan. 2008. Analysis of genetic diversity of cassava genepool by EST-SSR markers. *Chinese Journal of Tropical Crops*, 29 (3): 304 - 309. (in Chinese)
- 韦祖生, 夏志强, 李开绵, 王文泉. 2008. 木薯种质库遗传多样性的EST-SSR标记. *热带作物学报*, 29 (3): 304 - 309.
- Xin Ya, Cui Hai-rui, Zhang Ming-long, Yao Yan-ling, Lu Mei-zhen, Jin Ji-qiang, Lin Rong-shao, Cui Shui-lian. 2006. Transferable EST-SSR markers for Chinese cabbage. *Chinese Journal of Cell Biology*, 28: 248 - 252. (in Chinese)
- 忻 雅, 崔海瑞, 张明龙, 姚艳玲, 卢美贞, 金基强, 林容杓, 崔水莲. 2006. 白菜EST-SSR标记的通用性. *细胞生物学杂志*, 28: 248 - 252.
- Xu Kun, Zhang Dong-xiao, Wu Xiao-ming, Lu Guang-yuan, Liu Feng-lan, Chen Bi-yun, Gao Gui-zhen. 2008. SSR analysis of new winter rapeseed cultivars in the national field tests. *Chinese Journal of Oil Crop Sciences*, 30 (1): 29 - 34. (in Chinese)
- 许 鲲, 张冬晓, 伍晓明, 陆光远, 刘凤兰, 陈碧云, 高桂珍. 2008. 国家冬油菜区试新品种SSR指纹图谱构建与遗传关系分析. *中国油料作物学报*, 30 (1): 29 - 34.
- Zhang Pingzhi, Dreisigacker S, Melchinger A E, Reif J C, Mujeeb Kazi A, Van Ginke M, Hoisington D, Warburton M L. 2005. Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheat and their backcross-derived lines using SSR markers. *Molecular Breeding*, 15(1): 1 - 10.