

## 多位点整合的转基因棉花后代分子鉴定

马小波<sup>1,2</sup>, 王芙蓉<sup>1</sup>, 张传云<sup>1</sup>, 刘国栋<sup>1</sup>, 高俊平<sup>1,2</sup>, 张军<sup>1\*</sup>

(1. 国家棉花遗传改良中心山东分中心, 山东棉花研究中心, 济南 250100 ;2. 山东师范大学, 济南 250014)

**摘要:** GFP47 是一个利用双元载体 pPZP111 构建的绿色荧光蛋白基因的简单质粒载体 pPZP-GFP 经农杆菌介导转化陆地棉 (*Gossypium hirsutum* L.) 获得的 T<sub>0</sub> 株系, PCR 和 Southern 杂交分析证实其不仅整合了多个拷贝的正常的 T-DNA, 而且有紧邻 T-DNA 左边界的载体骨架序列的整合。对来源于 GFP47 的 33 个 T<sub>1</sub> 个体的 PCR 分析表明, T<sub>1</sub> 后代中可以分离只有正常 T-DNA 整合的个体。对 T<sub>1</sub> 群体中不同基因型个体的 Southern 杂交分析结果显示, 该多拷贝整合的 T<sub>0</sub> 转化体至少在棉花基因组中有四个插入位点, 不同的 Southern 带型暗示不同插入位点的 T-DNA 结构和排布有很大的不同。至少有一个位点整合了只包括目的基因和选择标记基因在内的正常 T-DNA 结构。

**关键词:** 转基因棉花; 多位点整合; T-DNA; 载体骨架; PCR; Southern 杂交分析。

**中图分类号:** S562.035.03 **文献标识码:** A

**文章编号:** 1002-7807(2008)01-0009-05

## Molecular Characterization on Offsprings from a Multiple-sites Integration Transgenic Cotton (*Gossypium hirsutum* L.)

MA Xiao-bo<sup>1,2</sup>, WANG Fu-rong<sup>1</sup>, ZHANG Chuan-yun<sup>1</sup>, LIU Guo-dong<sup>1</sup>, GAO Jun-ping<sup>1,2</sup>, ZHANG Jun<sup>1\*</sup>

(1. Shandong Branch of National Cotton Genetic Improvement Center, Cotton Research Center of Shandong Academy of Agricultural Sciences, Jinan 250100, China; 2. College of Life Sciences, Shandong Normal University, Jinan 250014, China)

**Abstract:** To investigate the pattern and organization of foreign gene integration in transgenic cotton, PCR and Southern blot analysis on T<sub>0</sub> plant, and the corresponding T<sub>1</sub> population of a transgenic cotton (*Gossypium hirsutum* L.) with multiple-integration sites derived from a simple construct pPZP-GFP via *Agrobacterium*-mediated transformation, and the corresponding T<sub>1</sub> population were carried out in this paper. The seeds harvested from GFP47 T<sub>0</sub> and genomic DNA isolated from T<sub>0</sub> plants were introduced from Plant Biotechnology lab of Temasek Life Sciences Laboratory, National University of Singapore. T<sub>1</sub> plants were planted in the greenhouse for DNA isolation and thereafter molecular analysis. PCR results showed that the T<sub>0</sub> plant was integrated not only with normal T-DNA including target gene (*gfp*) and selection marker gene (*nptII*), but also with vector backbone fragment outside the left border of the T-DNA. Besides confirmed the PCR results, Southern blot analysis on T<sub>0</sub> plant indicated that the T<sub>0</sub> plant was integrated with multi-copy of T-DNA. PCR characterization on 33 T<sub>1</sub> individuals showed that some individuals with only normal T-DNA structures were segregated from the population. Further analysis with Southern blot on these individuals with different genotype identified by PCR showed that there were, at least, 4 loci with different integration patterns in the cotton genome. The desirable transgenic event with only single copy of normal T-DNA integration could be

**收稿日期:** 2006-09-18 **作者简介:** 马小波 (1981-), 男, 在读硕士研究生; \* 通讯作者, scrczj@saas.ac.cn

**基金项目:** 国家转基因专项基金项目 (JY03-B-05-02); 山东省“三 O 工程”种质资源项目棉花子专题

segregated from the offsprings of this multiple-sites integration transgenic cotton.

**Key words:** transgenic cotton; multiple-sites integration; T-DNA; vector backbone; PCR; southern blot

基于根癌农杆菌 (*Agrobacterium tumefaciens*) 所具有的将 DNA 片段转移并整合到植物基因组中的能力发展起来的植物遗传转化系统, 使植物分子生物学及植物基因工程的研究进入了一个飞速发展的时期<sup>[1-2]</sup>。能够从农杆菌的 Ti 质粒上转移并整合到植物基因组中的一段 DNA, 称为 T-DNA (transfer DNA), 其两端各有一段 25bp 的不完全正向重复序列, 即 T-DNA 边界序列 (T-DNA borders), 分别为右边界 (right border RB) 和左边界 (left border LB)。早期的研究认为只有两个边界之间的 T-DNA 能够转移并整合到植物基因组中, 但是最近几年的许多研究已经证实 T-DNA 以外的载体骨架序列经常会整合到植物基因组中<sup>[3-7]</sup>。所转基因在植物基因组中的不同整合方式不仅决定其能否在遗传转化体内的稳定遗传和表达<sup>[8-10]</sup>, 载体骨架序列的整合也是转基因植物生物安全性考虑的一个重要方面。

基因工程的手段可以通过引入外源基因解决许多常规育种所不能企及的问题<sup>[11]</sup>; 转基因抗虫和抗除草剂的棉花已经在全球的主要产棉国商业化<sup>[12-13]</sup>。但是对于转基因棉花中外源基因在整个基因组中的排布形式及载体骨架序列的整合等许多与外源基因的遗传稳定性和生物安全性密切相关的问题却尚未进行过深入的研究。本研究利用一个通过农杆菌介导获得的整合了多个拷贝外源基因并有载体骨架序列整合的棉花 T<sub>0</sub> 转化体及其后代个体, 通过 PCR、Southern 杂交, 对 T-DNA 和载体骨架不同区段在棉花基因组中的整合与排列方式进行了系统的研究, 对于多拷贝外源基因在棉花基因组的整合及遗传行为进行了分析。

## 1 材料和方法

### 1.1 实验材料

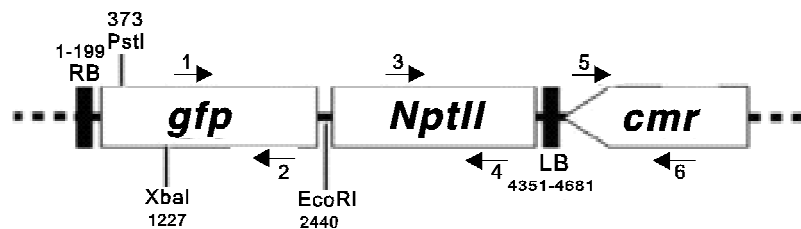
研究所用实验材料 GFP47T<sub>0</sub> 代植株, 采用农杆菌介导的转化方法以陆地棉 coker312 为原始材料获得, GFP47 的 T<sub>0</sub> 植株 DNA 和 T<sub>1</sub> 代种子引自新加坡国立大学 Temasek Life Sciences Laboratory 的棉花生物技术实验室。T<sub>1</sub> 代群体在山东棉花研究中心温室种植, 用于遗传分析并分单株收获自交种子。

### 1.2 棉花基因组 DNA 的分离纯化

基因组 DNA 的提取参照 CTAB 法<sup>[14-15]</sup>, 采棉花生长旺盛期幼叶分离纯化基因组 DNA 用于 PCR 和 Southern 杂交分析。

### 1.3 PCR 反应

PCR 引物由 invitrogen 公司合成, 根据遗传转化所用的双元载体 pPZP-GFP (图 1) 的序列设计了 3 对引物 (表 1), 引物 1 和 2 扩增 *gfp*, 引物 3 和 4 扩增 *nptII*, 5 和 6 扩增 LB 外侧的载体序列。PCR 反应体系: 20~50 ng 模板 DNA, 1U 的 Taq 酶, 0.4 μM 的引物, 1× PCR Buffer (10 mmol·L<sup>-1</sup> Tris-HCl (pH8.3), 50 mmol·L<sup>-1</sup> KCl), 160 μmol·L<sup>-1</sup> 的 dNTP, 2.5 mmol·L<sup>-1</sup> 的 Mg<sup>2+</sup>。反应程序为: 95℃ 预变性 3 min, 35 个循环 [95℃ 变性 20 s, 58℃~65℃ (温度视所用的引物组合而定) 退火 30 s, 72℃ 延伸 1 min (延伸时间根据不同引物组合所扩增的片段大小而定)], 72℃ 再延伸 10 min。扩增产物用 1% 浓度的琼脂糖电泳, goldview 染色, 紫外检测并照相。



*gfp* 为绿色荧光蛋白基因, *nptII* 为新霉素磷酸转移酶 II 基因, *cmr* 为抗氯霉素基因是位于左边界 (LB) 外的载体序列; 1 和 2, 3 和 4, 5 和 6 是 PCR 引物, 分别用于扩增这 3 个不同的基因。

图 1 双元载体 pPZP-GFP 结构

Fig. 1 Structure of binary vector pPZP-GFP

表 1 所用的 PCR 反应引物  
Table 1 PCR primers used in this study.

引物序号	方向与在载体上位置	引物序列(5'-3')
1	F1251	ATGGGTAAGGGGAGAAGAAGCTTTTCACTGG
2	R1792	TGGTCTGCTAGTTGAACGCTTCCATC
3	F3354	CTATTCGGCTATGACTGGGCACAACA
4	R4031	GTAAAGCACGAGGAAGCGGTCAGC
5	F4804	CAGTGGCGGTTTTTCATGGCTTCTTGT
6	R5321	AGAAACTGCCGGAATCGTCGTGGTA

#### 1.4 Southern 杂交

采用 Roche 公司的地高辛核酸检测系统,分别利用引物 1/2、3/4 和 5/6,以质粒 pPZP-GFP 为模板,以含地高辛-11-dUTP 的 dNTPs 为反应底物,PCR 标记获得 *gfp*、*npt II* 和 *cmr* 三个探针;利用不同的限制性内切酶酶切各转基因棉花植株的基因组 DNA,分别以三个探针在尼龙膜上杂交,检测杂交信号。

## 2 结果与分析

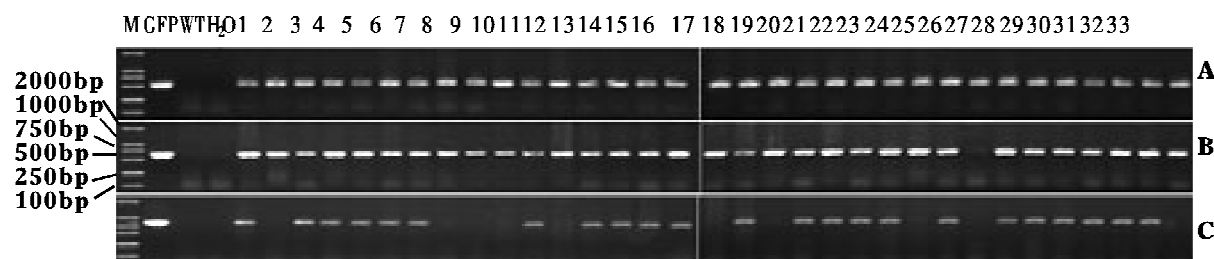
### 2.1 T<sub>0</sub> 植株中目标基因和载体片段的整合

分别用引物组合 1/2、3/4、5/6 对 GFP47 T<sub>0</sub> 植株基因组 DNA 进行 PCR 分析。发现其不仅整合了正常的 T-DNA 的拷贝包括 *gfp*、*npt II* 基因序列,而且有紧邻 T-DNA 左边界的载体骨架序列 *cmr* 基因的整合;分别以相应的探针对 *Xba I* 消化的基因组 DNA 进行 Southern 杂交分析(图 3),发现其整合了多个拷贝的 *gfp* 和 *npt II*

但仅有一个拷贝的 *cmr*,可能存在多个插入位点或者复杂的 T-DNA 结构。单纯由 T<sub>0</sub> 代的 PCR 和 Southern 分析的结果还不足以揭示这种复杂的整合型式(Complex T-DNA integration patterns)。

### 2.2 T<sub>1</sub> 群体中不同基因片段 PCR 的分离分析

对来源于 T<sub>0</sub> 代的 33 个 T<sub>1</sub> 代植株个体的 PCR 结果显示全部的植株有 *gfp* 基因,32 株个体含有 *npt II* 基因;只有 1 株没有检测到 *npt II*;大约有 75% (25/33) 的植株检测到 *cmr* 基因,有大约 1/4 的个体没有检测到 *cmr*(图 2)。PCR 结果清楚地表明 GFP47 的 T<sub>1</sub> 代群体中可以分离出没有载体序列整合的个体,暗示 GFP47T<sub>0</sub> 基因组中整合的不同位点中可能具有不同的结构。根据 T<sub>0</sub> 的 Southern 结果中 *cmr* 仅有一个杂交条带来看,可能只有一个位点整合了具有载体序列的外源 DNA 结构。



A、B 和 C 分别是 *gfp* 基因、*nptII* 基因和 *cmr* 基因的 PCR 结果, M: DNA 分子量标准; GFP: 二元载体 pPZP-GFP; WT: 野生型陆地棉 Coker312; H<sub>2</sub>O: 空白对照; 泳道 1~33: 33 株 T<sub>1</sub> 代植株。

图 2 T<sub>1</sub> 代植株个体的 PCR 结果

Fig. 2 PCR results of T<sub>1</sub> individuals

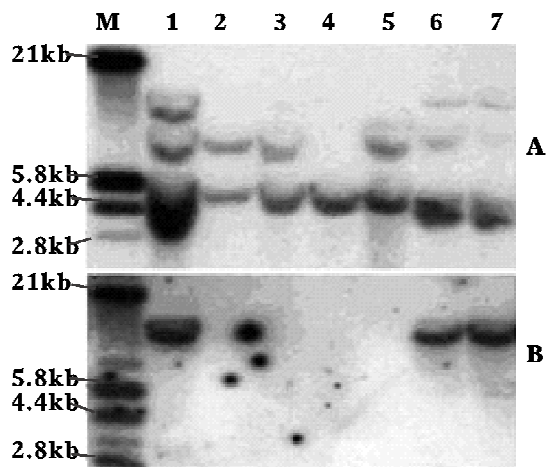
### 2.3 T<sub>1</sub> 群体中 Southern 带型的分离分析

2.3.1 T<sub>1</sub> 群体中不同整合位点的分离分析。选择了部分经 PCR 证实没有载体序列整合和有载体序列整合的 T<sub>1</sub> 个体,利用限制性内切酶 *Xba I* 消化基因组 DNA,分别以 *gfp*、*cmr* 为探针进行 Southern 杂交(图 3),由于 *Xba I* 的单一识别位

于 *gfp* 阅读框之前(图 1),可以大体揭示整合的拷贝数,而且如果是同一个整合位点,上述两个探针应该是杂交在同一个条带上。图 3 不仅证实了只有一个位点整合了具有载体序列的外源 DNA 结构(第 6、7 泳道),而且发现不含载体序列的不

同个体具有两种不同的带型(第 2、3、5 泳道与第 4 泳道),说明除了具有载体序列整合的一个位点以外,至少还有两个不同的整合位点,其中至少有一个位点是整合了单拷贝的正常的 T-DNA。

**2.3.2 不同整合位点外源 DNA 的可能结构。**为进一步明确 GFP47 各种不同整合位点的结构,利用另外一个在载体质粒上具有单一切点的限制性内切酶 *EcoRI* (切点位于两个基因表达框之间(图 1))消化基因组 DNA,分别以 *gfp* 和 *nptII* 为探针进行 Southern 杂交,发现 *gfp* 探针出现四种不同的带型,图 4 中给出了不同的带型的个体,其中第 2、4 泳道各不相同,而三者都没有载体序列的整合。由此推测原始  $T_0$  植株中具有 3 个独立的没有载体序列整合的插入位点,且位于



限制性内切酶 *XbaI* 消化基因组 DNA,杂交探针 A 为 *gfp*,B 为 *cmr*。M:地高辛标记的分子量标准;泳道 1:GFP47 $T_0$ ;泳道 2~7:不同的  $T_1$  代个体。

图 3 整合位点分离的 Southern 分析

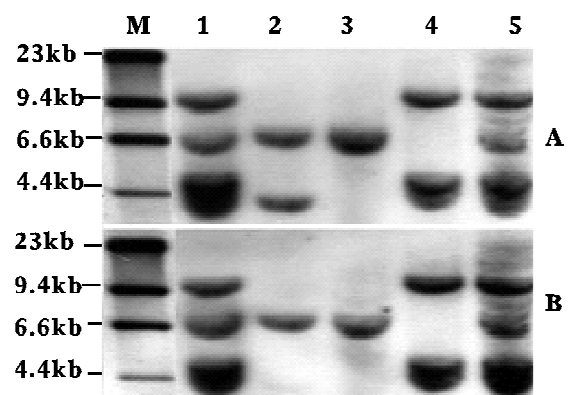
Fig. 3 Southern analysis on segregation of integration sites

### 3 结论与讨论

农杆菌介导的转基因植物中的外源基因存在不同的结构与排列方式——包括单位点整合的单拷贝的简单结构、单位点整合的串联重复和单拷贝或多拷贝的多位点整合的复杂结构以及外源基因的重排等<sup>[16]</sup>。而不同的整合方式可能又对外源基因的表达产生不同的影响,以遗传工程技术为基础的分子育种的成败与否,关键在于转基因能否在遗传转化体内稳定遗传和表达<sup>[17]</sup>。

利用绿色荧光蛋白基因构建的简单双元载体 pPZP-GFP 的农杆菌转化陆地棉,不仅在转化过程中可以利用正筛选(紫外检测荧光)和负筛选(卡那霉素的筛选系统)进行转化体的鉴定筛选,

不同的染色体上,全部应该为 4 个位于不同染色体上的整合位点。第 5 泳道是分离群体中与第 1 泳道( $T_0$ )相同的个体,推测该个体是所有 4 个位点的杂合体。比较 *gfp* 和 *nptII* 两个探针的杂交带型发现,第 2 泳道的个体有两个 *gfp* 杂交带,却仅有一个 *nptII* 的杂交带,可能是一个正常的 T-DNA 结构与另外一个缺失了 *nptII* 的 T-DNA 串联而成的复杂结构;第 3 泳道是一个正常的仅有一个拷贝整合的 T-DNA 结构。而第 4 泳道的个体很可能是一个串联重复的结构,也有可能是两个位点的杂合体。对于分离出的各种不同整合位点的个体的详细结构以及插入位点棉花基因组的结构信息的研究正在进行中。



限制性内切酶 *EcoRI* 消化基因组 DNA,杂交探针 A 为 *gfp*,B 为 *nptII*。M:地高辛标记的分子量标准;泳道 1:GFP47 $T_0$ ;泳道 2~5:不同的  $T_1$  代个体。

图 4 整合位点结构的 Southern 分析

Fig. 4 Southern analysis on structure of integration sites

同时简单的 T-DNA 结构非常有利于转基因植物中的结构和插入位点的分子生物学分析。本研究中的  $T_0$  代 GFP47 是具有多拷贝、多位点整合的转基因棉花,研究证实其不仅整合了正常的 T-DNA 的拷贝,也有紧邻 T-DNA 左边界载体骨架序列的整合,而且存在重排的 T-DNA 结构。

载体序列的整合在转基因植物中不仅影响目的基因的表达<sup>[8-10]</sup>,而且也是转基因植物生物安全性的重要方面。一般转基因过程中,一旦鉴定出具有载体序列整合的  $T_0$  个体就丢弃以保证以后顺利进行安全性评价。本研究发现,有载体序列存在的多位点整合的转基因棉花后代中可以分离出正常符合要求的转基因系,另外,Southern

分析发现有复杂整合位点的存在(图 4),其中第 2 泳道的个体可能是一个正常的 T-DNA 串接了一个缺失了 *nptII* 基因的不完整的 T-DNA,由于两个 T-DNA 共用一个选择标记基因,因而能够通过正负两个筛选系统而被选择下来。在 PCR 结果中有一个个体没有 *nptII* 基因的存在,很可能是在同一个细胞中有多个转化事件发生,其它的整合已经使转化的棉花细胞具有了抗卡那霉素的能力,而使该没有选择标记基因的片段整合在独立的位点上。由此推测,多位点转基因植株的获得可能源自一个细胞中多个独立的转化事件,但这些转化事件的发生可能是有次序的。

因此,有必要对一些多拷贝的  $T_0$  代植株进行细致的研究,以在分离群体中选择可能的有利用价值的后代。同时,对于多拷贝的具有复杂的整合位点结构的转基因个案进行深入研究,对探讨农杆菌介导的植物遗传转化中的一些重要机制如整合与重排、外源基因的整合对植物基因组本身的影响等重要问题具有重要的价值。

外源基因的整合是一个相当复杂的问题。虽然目前人们已研究出许多有效的转化体系可将外源基因转入受体细胞中,但是转入的外源基因能否整合,整合的位置以及整合的状态等均难以控制。虽然全球已经有大面积的转基因棉花在种植,却没有发现棉花遗传转化中 T-DNA 结构以及整合机制的报道。对于利用一个多位点整合的转基因棉花后代中的外源基因整合、插入位点的数目及变化进行初步的研究。希望对于 T-DNA 整合的机制有较为深刻的了解,并期望对降低 T-DNA 复杂结构以及载体骨架的转移而提高棉花转基因的效率提供理论依据。

致谢:本研究采用仅用于研究的转基因棉花 GFP47 的  $T_0$  植株 DNA;种子由新加坡国立大学 Temasek Life Sciences Laboratory(淡马锡生命科学实验室)提供,谨致谢忱。

#### 参考文献:

[1] NEWELL C A. Plant transformation technology. Developments and applications [J]. Mol Biotechnol, 2000, 16(1): 53-65.  
 [2] TOPPING J F, Wei W, Clarke M C, et al. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Application in T-DNA tagging[J]. Methods Mol Biol, 1995, 49: 63-76.  
 [3] PODEVIN N, De Buck S, De Wilde C, et al. In-

sights into recognition of the T-DNA border repeats as termination sites for T-strand synthesis by *Agrobacterium tumefaciens* [J]. Transgenic Res, 2006, 15(5): 557-571.

- [4] RAMANATHAN V, Veluthambi K. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA[J]. Plant Mol Biol, 1995, 28(6): 1149-1154.  
 [5] YIN Zhong-chao, Wang Guo-liang. Evidence of multiple complex patterns of T-DNA integration into the rice genome[J]. Theoretical and Applied Genetics, 2000, 100(3-4): 461-470.  
 [6] WENCK A, Czako M, Kanevski I, et al. Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation[J]. Plant Mol Biol, 1997, 34(6): 913-922.  
 [7] KONONOV M E, Bassuner B, Gelvin S B. Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome; evidence for multiple complex patterns of integration[J]. Plant J, 1997, 11(5): 945-957.  
 [8] IGLESIAS V A, Moscone E A, Papp I, et al. Molecular and cytogenetic analyses of stably and unstably expressed transgene loci in tobacco[J]. Plant Cell, 1997, 9(8): 1251-1264.  
 [9] JAKOWITSCH J. Molecular and cytogenetic characterization of a transgene locus that induces silencing and methylation of homologous promoters in trans [J]. Plant J, 1999, 17(2): 131-140.  
 [10] MATZKE A J, Matzke M A. Position effects and epigenetic silencing of plant transgenes [J]. Curr Opin Plant Biol, 1998, 1(2): 142-148.  
 [11] JOHN M E, Keller G. Characterization of mRNA for a proline-rich protein of cotton fiber[J]. Plant Physiol, 1995, 108(2): 669-676.  
 [12] 刘方,王坤波,宋国立. 中国棉花转基因研究与应用[J]. 棉花学报, 2002, 14(4): 249-253.  
 [13] 王仁祥. 中国转基因抗虫棉的应用及发展对策[J]. 棉花学报, 2003, 15(3): 180-184.  
 [14] PATERSON A H, Brubaker C L, Wendel J F. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis[J]. Plant Mol Biol, 1993, 11: 122-127.  
 [15] 王芙蓉,张军,刘任重,等. 海岛棉 DNA 导入陆地棉栽培品种获得变异种质的初步遗传分析[J]. 中国农业科学, 2005, 38: 1528-1533.  
 [16] 曾凡锁,詹亚光. 转基因植物中外源基因的整合特性及其研究策略[J]. 植物学通报, 2004, 21: 565-577.  
 [17] 华志华,黄大年. 转基因植物中外源基因的遗传学行为[J]. 植物学报, 1999, 41: 1-5. ●