

摘 要

通过构建三个双元载体的质粒 pDsE、pDsG 和 pAc 和农杆菌介导的 T-DNA 转基因,发展了一个带有基因捕获器和增强子捕获器的 *Ac/Ds* 转座子系统作为插入序列标签来研究水稻基因组。主要研究结果如下:

1. 通过农杆菌介导的 T-DNA 转基因,共获得 1188 株独立可育的转基因水稻植株。其中来自日本晴的转基因水稻为 526 株包括 NAc(转 pAc)51 株, NE(转 pDsE)361 株, NG(转 pDsG)114 株;来自中花-11 号的转基因水稻为 662 株,包括 CAc(转 pAc)176 株, CE(转 pDsE)320 株, CG(转 pDsG)166 株。潮霉素抗性试验表明 T-DNA 在水稻染色体中平均插入位点数为 1.3 个;通过 Southern 杂交分析, T-DNA 在水稻基因组中的平均拷贝数为 1.8 个。
2. T-DNA 插入到水稻基因区的 GC%含量分布与插入到基因组 DNA 的分布基本一致。通过对 126 个 T-DNA 插入位点两侧 500bp 水稻序列 GC 含量分析,表明插入位点处水稻基因组 DNA 平均 GC 含量为 42.5%。83%的 T-DNA 插入子落在水稻 GC 含量为 30-50%的基因组区域。有 50 个 T-DNA(40%)插入到水稻的基因区。
3. T-DNA 在水稻基因组 DNA 的插入,既有准确的插入,又有不准确的插入。在 182 个 T-DNA 与水稻的结合序列析中,只有 69 个序列(38%)中没有 filler-DNA 的存在;而存在 filler-DNA 的序列有 113 个,占总数的 62%,长度从 1bp 到数百 bp 不等。通过对 61 个大于 50bp 的 filler-DNA 分析,发现来自 T-DNA 内部的 DNA 序列是 filler-DNA 的主要来源,总共有 49 个占总数的 80%;有 8 个(13%)是来自于 T-DNA 以外的双元载体主干序列;另有 4 个(7%)与水稻序列和载体序列都没有同源性。在 158 个右边界旁邻序列中有 77 个保持了特征序列 TGA;而左边界序列的变化比较大,没有发现在某一特定核苷酸断裂的现象。在 T-DNA 左边界与水稻基因组 DNA 直接相联的序列中发现有 1-8 个碱基的同源序列存在。另外还发现有 16 例 T-DNA RB-LB 的直接串联。
4. *Ds* 元件附近的基因组 DNA 结构对跳跃频率有影响,PCR 分析表明,来自 11 个 *Ds* 亲本总共 20 个 F₂ 群体其 *Ds* 跳跃频率变化范围为 0-40%;而亲本 *Ds* 元件拷贝数对 *Ds* 的跳跃频率影响不大。Southern 杂交结果表明, *Ds* 在 F₂ 群体中的重新插入频率和独立插入频率都保持在 70%以上。有 *Ds* 插入但没有 *Ac* 元件而稳定插入的 F₂ 植株频率变化范围在 14-33%之间。
5. *Ds* 在水稻基因组 DNA 中有插入到基因区的倾向。在 29 个 *Ds* 插入位点旁邻

序列的分析中, 有 17 个(59%)插入到了基因区, 其中有 6 个在启动子区域, 7 个在外显子区域, 有 4 个在内含子区域。另外还有 10 个序列位于非基因区域。*Ds* 插入位点的两侧会形成 8bp 水稻碱基的正向重复。

6. 在 *Ds* 独立插入的 F_2 水稻植株的叶、根、花和种子中, 增强子捕获器的 GUS 活性的检出率为 28%, 基因捕获器的 GUS 活性检出率为 22%。

关键词: 水稻, T-DNA, filler-DNA, *Ac/Ds*, 转座子, 基因捕获器, 增强子捕获器

Abstract

By constructing three binary vectors (pDsE, pDsG and pAc) and T-DNA transformation mediated by *Agrobacterium*, we developed an *Ac/Ds* system harboring gene trap and enhancer trap for gene tagging in rice. The following results are obtained:

1. Through *Agrobacterium*-mediated T-DNA transformation, 1188 independent fertile transgenic rice lines were produced. In *japonica Nipponbare*, 526 lines were obtained including 51 NAc lines transformed with pAc, 361 NE lines transformed with pDsE and 114 NG lines transformed with pDsG. In *japonica Zhonghua-11*, 662 lines were obtained including 176 CAc lines transformed with pAc, 320 CE lines transformed with pDsE and 166 CG lines transformed with pDsG. Hygromycin resistance tests revealed that transgenic plants contain an average of 1.3 loci of T-DNA inserts. Southern blot analysis of transgenic plants indicated that they contain average 1.8 T-DNA copies in rice genome.
2. The distribution of GC content of rice gene region inserted by T-DNA was very similar with that of rice genome DNA. Through GC content analysis of 500bp rice genomic DNA flanking insertion site in 126 lines, we found that the average GC content of the insertion site was 42.5% and 83% of the T-DNA insertion lines locate in a genomic domain ranging from 30% to 50% in terms of GC content. Fifty (40%) out of the total 126 T-DNA insertion lines were in the gene region of rice.
3. Both precise insertion and imprecise insertion existed when T-DNA integrated into rice genomic DNA. Among total 182 T-DNA/rice DNA junctions, only 69 (38%) were found with no filler-DNA, whereas the remaining 113 (62%) contained filler-DNA from 1bp to several hundred base pairs. By analysis of 61 filler-DNA larger than 50bp, we found that 49 (80%) were from the sequence within T-DNA. Eight (13%) of them were from backbone sequence of binary vector out of T-DNA. In addition, 4 (7%) had no similarity with rice or vector sequences. Three signature nucleotides TGA of right border repeat were kept in 77 out of 158 T-DNA insertions. No specific nucleotide where the T-DNA left border broken was found when T-DNA integrated into rice genomic DNA. Microhomologous fragments from 1 to 8bp were observed in the sequences next

to the left border of T-DNA insertions. Sixteen direct ligation of T-DNA right border to left border were observed in our experiments.

4. The structure of genomic DNA at insertion site affected *Ds* transposition frequency greatly. *Ds* excision frequency varied from 0% to 40% in 20 F_2 populations derived from 11 different *Ds* parents with PCR analysis. It was found that *Ds* copy number had little influence on *Ds* transposition frequency. Southern blot analysis revealed that more than 70% of *Ds* element reinserted into rice genomic DNA and most of them (>70%) were independent in F_2 population. 14-33% of *Ds* reinsertion F_2 plants were stable without *Ac* element.
5. *Ds* is inclined to integrate into gene region of rice genomic DNA. Seventeen out of total 29 (59%) flanking sequences of *Ds* elements were identified to be within gene region of rice genomic DNA, including 6 in promoter, 4 in exon and 7 in intron. Another 10 flanking sequences of *Ds* elements were proved to be within intergenic regions. 8bp duplication of rice genomic DNA appeared at insertion site of *Ds* element.
6. Among the F_2 plants with independent *Ds* insertions, approximately 28% of enhancer trap and 22% of gene trap inserts displayed GUS activity in leave, roots, flowers and seeds.

Key words: rice (*Oryza sativa* L.), T-DNA, filler-DNA, *Ac/Ds*, transposon, gene trap, enhancer trap

致 谢

本研究是在导师吴平教授悉心指导下完成的,研究倾注着导师无数的心血。在此谨向导师表示最诚挚的谢意!导师开阔的眼界、广博的知识、敏锐的思维、严谨求实的治学态度、对事业和人生信念的执著追求、百折不挠永不停息的奋斗精神时刻激励着我,催我奋进,将使我受益终生。

在攻读博士学位论文期间,承蒙傅承兴老师、蒋德安老师、朱诚老师、易可老师、吴运荣老师、姜华武博士、李靖老师、陈汉民老师、刘非燕老师的多方帮助和指导;感谢中国水稻研究所的孙宗修老师、傅亚萍老师在转基因技术上的帮助和提供 pDsBAR1300、pUBITS 载体、感谢农业与生物技术研究室的毛碧增老师提供 pCambia1300 载体;感谢傅向东博士提供 pAHC25 质粒;感谢课题组成员张帆、陈双燕、汪明怡、周洁、汪少敏、徐敏、段瑞君同学的大力协助;感谢侯兴亮、刘洪家、佃蔚敏、吴忠长、毛传藻、郑炳松、杨玲、王首锋、杜黎明、张磨龙、朱世华、黄幅、齐晓朋、王芳、樊叶杨等同学在平时学习和工作中给我的帮助;感谢所有关心帮助过我的同学和朋友。

在此我要特别感谢我的妻子叶纨芝对我的理解与支持,感谢我的父亲、母亲、岳父、岳母、哥哥、姐姐的关心与爱护,是他们使我能够全身心地投入到学习和研究当中。

金 维 正

2003 年 5 月于杭州华家池

缩 略 词 表

Abbreviation

缩写	英文	中文
ABRC	The Arabidopsis Biological Resource Center	拟南芥生物资源中心
<i>Ac/Ds</i>	Activator/Dissociation	激活子/分离子
AD	Arbitrary degenerate primer	随机简并引物
ADH	alcohol dehydrogenase	乙醇脱氢酶
ALS	acetolactate synthase	乙酰乳酸合成酶
AS	Acetosyringone	乙酰丁香酮
BAC	Bacterial artificial chromosomes	细菌人工染色体
bp	base pair	碱基对
<i>CodA</i>	cytosinedeaminase	胞嘧啶脱胺酶
DNA	Deoxyribonucleic acid	脱氧核糖核酸
<i>En/I</i>	<i>Enhancer/ Inhibitor</i>	增强子/抑制子
EST	Expressed sequence tag	表达序列标签
GFP	green fluorescent protein	绿色荧光蛋白
GUS	β -Glucuronidase	β -D 葡萄糖醛糖苷酶
HPT	Hygromycin phosphotransferase	潮霉素磷酸转移酶
HYG	Hygromycin	潮霉素
IAAH	indoleacetamide hydrolase	吲哚乙酰胺水解酶
IPCR	Inverse PCR	反向 PCR
LUC	luciferase	荧光素酶
NAA	naphthalene acetic acid	萘乙酸
NAM	naphthalene acetamide	萘乙胺
NOS	Nopaline synthase	胭脂碱合成酶
NPT II	Neomycin phosphotransferase	新霉素磷酸转移酶
OsEXP	Expansin	水稻膨胀素
PAC	P1-derived artificial chromosomes	P1 人工染色体
PEG	Polyethylene glycol	聚乙二醇
PPT	phosphinothricin	草丁膦
TAIL-PCR	Thermal asymmetric interlaced PCR	热不对称交错 PCR
T-DNA	Transferred DNA	转移 DNA
TE	Transposon element	转座子元件
Ti	Tumor-inducing	肿瘤诱导 (Ti 质粒)
TIR	Terminal inverted repeat	末端倒转重复
X-Gluc	5-bromo-4-chloro-3-indolyl- β -D glucuronide	5-溴-4-氯-3-吲哚 β -D 葡萄糖醛酸
YAC	Yeast artificial chromosomes	酵母人工染色体

第一章 文献综述

1 T-DNA 作为插入序列标签的研究进展

1.1 农杆菌介导的 T-DNA 遗传转化

T-DNA 作为转基因载体用于植物的遗传转化始于 20 世纪八十年代。农杆菌介导的 T-DNA 转基因首先成功于双子叶植物,随着技术的进步与突破,特别是在单子叶植物的转化技术获得突破成熟之后,利用 T-DNA 作为载体在单子叶植物中的遗传转化也成为一种常规技术。因为这种方法简便、成本低、稳定性好而受到广泛的青睐。目前 T-DNA 用作外源目的基因导入植物的载体,已经成为植物遗传育种、品种改良的一种有效先进的工具。

因为 T-DNA 在转基因植物基因组中比较随机的分布,所以无论是在双子叶还是单子叶植物中都被已经用作研究植物本身基因功能一种很好的插入序列标签、大大加快了植物基因功能的研究。

1.1.1 T-DNA 的发现与应用

Smith和Townsend发现植物冠瘿瘤病(crown gall disease)是农杆菌引发的,之后对农杆菌的研究就广泛的开展起来了。冠瘿瘤病曾在法国、东欧和澳大利亚的葡萄树、核果树和坚果树上有大面积的发生,造成很大的危害。White和Braun发现冠瘿瘤组织能在正常细胞不能生长的无激素培养基上无限制生长并且不需要细菌的存在;Link和Eggers发现冠瘿瘤组织能产生生长素;发现细胞分裂素在冠瘿瘤组织中大量存在。Braun提出“肿瘤诱导因子”(tumor-inducing principle)假说,尽管这一假说对这一因子的本质不清楚,但推测染色体外存在一种遗传物质DNA,能使处于静止状态的植物细胞转变为分裂活跃的细胞。

随着分子生物学技术的出现,Schilperoort等发现无菌培养的冠瘿瘤含有细菌来源的DNA,尽管在当时这一结论是有争议的。但到70年代,科学家发现用高温培养的方法可以使农杆菌失去致瘤能力,并分离鉴定出一种与致瘤能力完全相关的Ti质粒(tumor-inducing plasmid)(Van Larebeke等1974, Van Larebeke等1975, Zaenen等1974),这使得研究的范围缩小到Ti质粒上,最终通过Southern杂交的方式确定了转移到植物细胞中的只是Ti质粒的一部分即T-DNA (Transferred DNA)(Chilton等1977, Chilton等1978, Depicker等1978)。以上内容摘自Zupan等(2000)的综述。

进一步研究发现, T-DNA 两端左右边界为 25bp 的正向重复序列, 即边界序列 (border sequence), 分别称之为左边界(left border, LB)和右边界(right border, RB), 边界序列为保守序列(5'-TGACACGATATATTGGCGGGTAAAC-3'), 在 T-DNA 右边界外还有一个 24bp 的超级驱动序列 (overdrive sequence) (TAACTCGCTGTGTATGTTTGTGTTG), 对 T-DNA 的转移起增强子的作用(见王关林 1998, p140)。

利用农杆菌介导的 T-DNA 转基因成功报道于 1983 年, Murai 等将种子贮藏蛋白基因成功转到向日葵和烟草并得到表达, 之后利用农杆菌介导的 T-DNA 转基因在双子叶植物中就成为了一种常规方法。

1.1.2 农杆菌介导的 T-DNA 转基因技术在单子叶植物中的研究进展

农杆菌介导的 T-DNA 转基因在单子叶植物上取得突破之前, 主要采用直接转化法。直接转化法就是通过处理的裸露 DNA 直接导入植物细胞, 实现基因转化的一种技术。按技术原理来分, 转基因技术通常可分为物理法和化学法两种。化学法又可分为聚乙二醇(PEG)介导基因转化法、脂质体介导基因转化法; 物理法又可分为电激法(electroporation)介导基因转化法、超声波介导基因转化法、显微注射(microinjection)介导基因转化法、激光微束介导基因转化法、基因枪轰击(microprojectile bombardment)基因转化法。另外还有萌发种子电泳法、花粉管通道法、种子浸泡法等。

其中 Zhang 和 Wu(1988)首先报道了采用 PEG 介导法得到第一例转基因水稻植株; Toriyama 等(1988)和 Zhang 等(1998)两个研究小组分别报道了水稻原生质体电激法得到的转基因水稻植株。但是两者都需要原生质体的培养, 原生质体的培养本身就是一个复杂而且困难的方法, 而且原生质体的再生频率也不高, 因此这两种方法在应用上都受到一定的限制。基因枪法由于对转化受体的要求不高, 而且转化的效率高, 因此方法一经出现, 就受到广泛的应用, 并先后用这一方法在大豆(Wang 等 1998)、玉米(Gordon-Kamm 等 1990)、水稻(Christou 等 1991, 朱冰等 1996)、小麦(Vasil 等 1992)、黑麦(Somers 等 1992)等农作物上获得了转基因植株。

农杆菌介导 T-DNA 在单子叶植物中的转基因在 20 世纪 90 年代取得突破。Chan 等(1992, 1993)用水稻幼胚为材料获得一些转基因水稻植株, 并对一株水稻进行了遗传分析, Southern 杂交证明转基因 T-DNA 能够遗传到后代中。Hiei 等(1994)在水稻中的转基因工作则以无可辩驳的证据表明农杆菌不但可以转化单子叶植物而且还可以和双子叶植物一样有效。他们采用水稻成熟胚的盾片愈伤组织, 用乙酰丁香酮(Acetosyringone, AS)诱导农杆菌(LBA4404, EHA101)感染愈伤

组织,使转化频率大大提高,并获得了大量的转基因水稻植株。他们对众多转基因水稻植物进行了详细的遗传分析,还有就是获得了充分的分子证据,包括 Southern 杂交和 T-DNA 边界序列信息的获得。之后, Dong 等(1996)在爪哇稻中转化成功, Aldemita 等(1996)转化籼稻也获得了成功, Rashid 等(1996)在籼稻中转化取得了同粳稻中一样高的转化效率。这些工作表明农杆菌介导的 T-DNA 转基因在单子叶植物中已经完全的确立起来了。

1.2 T-DNA 在植物基因组 DNA 中整合的机制

T-DNA 整合到植物基因组中具体机制至今不清。T-DNA 在植物中的插入被认为是随机的,它可以插入到任何一条植物染色体,但是插入位点通常具有以下特点:①T-DNA 优先整合到转录活跃的植物基因位点;②T-DNA 偏爱插入到 A-T 丰富区;③植物 DNA 的靶位点与 T-DNA 有小的同源性(microhomology) (Brunaud 等 2002, Mayerhofer 等 1991, Tinland 1996, Szabados 等 2002)。目前,主要有以下几种关于 T-DNA 整合到植物基因组 DNA 中的模型。

1.2.1 双链断裂修复模型

双链断裂修复模型首先是植物靶 DNA 产生双链断裂,解旋的靶 DNA 与双链 T-DNA 退火,单链部分被 3'→5'外切酶或修复系统中单链特异的核酸内切酶去除;然后进行修复连接,并由此导致了 DNA 的整合和缺失。靶 DNA 和 T-DNA 缺失程度由靶 DNA 与 T-DNA 退火时形成的短的 DNA 互补顺序的位置所决定 (Mayerhofer 等 1991)。

1.2.2 单链缺口修复模型

单链缺口修复模型是在靶 DNA 上形成一个缺刻,然后由于部分解旋或 5'→3'外切酶消化形成缺口,单链 T-DNA 靠近缺口,由于两链存在短的 DNA 互补顺序,于是退火形成异质双链,未配对的 3'和 5'单链部分被除去, T-DNA 末端与靶 DNA 连接,在靶 DNA 链的互补链产生缺刻,以游离的 3'DNA 末端为引物修复合成第二条 T-DNA 链 (Mayerhofer 等 1991)。

1.2.3 T-DNA 串联重复的形成

T-DNA 整合到植物基因组中不单单以单个 T-DNA 的形式,在很多情况下是以两个或两个以上的 T-DNA 串联形式整合到基因组 DNA 中的。串联的形式有 RB-RB、RB-LB 及 LB-LB 等种类。Buck 等(1999)认为这种 T-DNA 的串联是在 T-DNA 整合到植物基因组 DNA 之前形成的,他们的理由是因为在串联的 T-DNA

交接处没有发现有植物基因组 DNA 的存在。

更为复杂的是多个 T-DNA 的串联插入, Krizkova 和 Hroudal(1998)发现在两个相对完整 T-DNA 串联重复的交接还有 T-DNA 的成份存在, 他们认为这种整合方式可能是 T-DNA 整合时发生的, 首先由一条 T-DNA 与植物染色体 DNA 通过几个互补碱基配对结合, 然后另外两条 T-DNA 链通过几个碱基的互补配对分别与第一条 T-DNA 的 3' 和 5' 端结合, 游离出的 T-DNA 单链分别与染色体 DNA 3' 端和 5' 端结合, 通过对游离的 T-DNA 降解及以 T-DNA 为模板的 DNA 合成而使多个 T-DNA 整合到植物基因组中。

1.3 T-DNA 作为插入序列标签在植物功能基因组学研究中的应用

1.3.1 T-DNA 在双子叶植物功能基因组学研究中的应用

T-DNA 作为插入序列标签在双子叶植物中的应用主要以在拟南芥中为例。拟南芥植株小, 种子小而多 (一株植株能产生大约 10000 粒种子) 生长周期短, 基因组小而成为双子叶植物的模式植物。

T-DNA 插入突变体在拟南芥中早在 20 世纪 90 年代初就受到关注, 并开始建立大规模的插入突变体库 (Feldmann 1991, Fortsthoefel 等 1992, Koncz 等 1992)。为了全世界科研人员研究的方便, 早在 1991 年 9 月在美国俄亥俄州立大学就建立了拟南生物资源中心 ABRC (The Arabidopsis Biological Resource Center), 以收集、保存、分发各种突变体。现在 ABRC 已经建立了一个庞大的拟南芥插入突变体库 (包括很大一部分来源于 T-DNA 插入突变体), 并且还在以每年数万的数量在增长。通过拟南芥的网站 <http://www.arabidopsis.org> 可以获得拟南芥 T-DNA 插入突变体。在拟南芥中, 已经有大量的基因功能以 T-DNA 为插入序列标签得到研究。

1.3.2 T-DNA 在单子叶植物功能基因组学研究中的应用

T-DNA 作为插入序列标签在单子叶植物功能基因组学研究中的应用比双子叶植物中的研究要迟得多。水稻作为单子叶植物是禾本科的模式植物, 其大规模的插入突变体库见于 2000 年 Jeon 和 Jeong 等在 2003 年的报道。在我国也有多家单位正在开展建立大规模的插入突变体库 (朱正歌 等 2001)。但是至今没有报道利用 T-DNA 作为插入序列标签在水稻中克隆新的基因。不过相信在广大科研工作者的努力下, 必定能很好的利用 T-DNA 这一工具来研究水稻的基因功能。

2 *Ac/Ds* 转座子系统作为插入序列标签的研究进展

2.1 *Ac/Ds* 转座子系统的研究历史

转座子元件 TE(transposon element)首先是由 Barbara McClintock 在 20 世纪三、四十年代研究玉米籽粒颜色变化时发现的,但一直没有受到人们的重视,直到六、七十年代在大肠杆菌 *gal* 和 *lac* 操纵子发现了可以转移的插入序列(Insertion sequence),McClintock 的转座子理论才被人们所重新认识和接受。

转座子在 DNA 结构上具有一些共同的特点:一、在转座子的两端,存在末端倒转重复序列(terminal inverted repeat, TIR),它们是转座发生所必需的结构,而且至关重要。二、转座子编码转座酶(transposase),其功能是促进转座子的转座发生。三、受体 DNA 上很短的一段靶序列,由于转座子的插入,会在转座子的两侧形成正向重复序列,靶序列的正向重复序列长短由转座子决定,每一类转座子造成的正向序列重复其碱基数目是确定的。

玉米中的 *Ac/Ds* 转座子系统(activator/dissociation)是研究的比较多的一个系统。其中 *Ac* 是自主性(autonomous)成分, *Ds* 是非自主性(nonautonomous)成分。*Ac* 全长 4565bp,具有 11bp 的末端倒转重复序列 TCAGGGTGA AAA。自主性 *Ac* 元件还编码一个的转座酶基因(*Ac* transposase),具有完整结构的 *Ac* 元件能够发生自主性转座。*Ds* 在结构上缺乏编码完整转座酶基因的 DNA 序列,但是保留了其它如末端倒转重复序列发生转座作用所必须的结构,因此 *Ds* 不能发生自主性的转座,但是在 *Ac* 提供转座酶的情况下同样能够发生转座。*Ac/Ds* 转座子在玉米、拟南芥和水稻基因组中插入位点处能形成 8bp 的正向重复序列,当 *Ac* 或 *Ds* 从插入位点割离(excision),割离的发生有可能是精确地发生在转座子末端倒转重复序列两侧,也有可能是不精确的割离,通常会使原来的 8bp 重复序列的一侧或两侧缺少几个碱基,这样果会留下 *Ac/Ds* 曾经在此插入过的足迹(footprint)。

在玉米中,还存在有其它的转座子系统如 *En/I* 系统(*Enhancer/ Inhibitor*) (Peterson 1953),又称为 *Spm/dSpm* 系统(*Suppressor/mutator*) (McClintock 1954)。自主性 *En/Spm* 元件长 8.3kb,编码两个基因 *TnpA* 和 *TnpD*,它们是转座所必需的。非自主性 *I/dSpm* 元件通常是有缺陷的 *En/Spm* 元件,无法自身进行转座。这一转座子系统的 TIR 为 13bp,插入靶位点形成 3bp 的正向重复序列,与 *Ac/Ds* 一样,不精确的割离会留下足迹(Pereira 等 1985)。

利用 *Ac/Ds* 转座子系统克隆的第一个基因是玉米的 *bronze* 基因。*Ac/Ds* 作为外源转座子系统也已经成功的应用在矮牵牛、拟南芥、番茄、烟草等植物基因的克隆中(Aarts 等 1993, Chuck 等 1993, Jones 等 1994, Whitham 等 1994)。特

别是在拟南芥中许多具有重要功能的基因都是以 *Ac/Ds* 转座子系统作为序列标签进行克隆的。

2.2 *Ac/Ds* 转座子在突变体库构建中的应用

Ac/Ds 转座子系统在突变体构建中的应用形式可以有多种多样, 按 *Ac* 与 *Ds* 是否构建在一起, 能否自主性发生转座, 一般分成二元系统与一元系统, 下面就以这两种系统来说明 *Ac/Ds* 在突变体库构建中的应用。

2.2.1 二元系统

二元系统是将 *Ac* 与 *Ds* 分开构建到双元载体 T-DNA 中。自主性的 *Ac* 元件可以通过基因工程技术将转座发生所必须的 5'和 3'末端切除, 因此不会发生转座, 而只提供完整的转座酶基因。*Ds* 元件含有转座发生所必须的 TIR 结构, 而不含有转座酶基因, *Ds* 元件的内部则可以做各种修饰, 如加报告基因、抗生素基因、除草剂基因等。

在制作突变体库的时候, 首先是获得 *Ac*-T-DNA 和 *Ds*-T-DNA 的转基因亲本植株, 然后通过杂交的方式产生大量的 F_1 种子, F_1 自交产生 F_2 , 在 F_2 代中对跳跃的转座子进行选择。由于 *Ac* 转座酶的存在, *Ds* 就不会稳定。为了获得稳定的 *Ds* 插入子, 可以通过遗传分离的方法选择不带 *Ac* 的 *Ds* 插入植株。

二元系统另一个优点是在杂交前选择 *Ac* 与 *Ds* 起始植株的拷贝数, 并且通过杂交获得大量的转座频率稳定的同一 F_1 种子, 这样通过少数几个起始转基因植株就可以获得大量的 *Ds* 插入突变体(Sundaresan 等 1995)。Izawa 等(1997)首先利用二元系统来创建水稻突变体, Nakagawa (2000) 等对二元系统在水稻中的跳跃规律作了比较仔细的分析。

2.2.2 一元系统

一元系统是将 *Ac* 转座酶基因与 *Ds* 转座子元件构建在同一个载体中, 无须通过如二元系统杂交的步骤, 方便之处就是可以在转基因植物 T_1 代中可以分离筛选到 *Ds* 跳跃的植株。

一元系统又可以分为 *Ac* 自主转座系统与 *Ac/Ds*-T-DNA 系统。*Ac/Ds*-T-DNA 系统是将 *Ac* 转座酶基因置于 *Ds* 转座元件的外部, 当 *Ds* 跳到另一位置时, 可以获得与 *Ac* 转座酶分离而稳定的 *Ds* 插入子, 这一点与二元系统相类似。

Ac 自主转座系统是利用 *Ac* 元件的自主转座特性(含有完整的转座结构与转座酶基因, 转座酶基因置于转座子的内部)。它的特点是在后代可以进行连续的跳跃而不能得到稳定的突变体。Enoki 等(1999)利用自主性转座子 *Ac* 在水稻

中做了详细的分析,他们认为 *Ac* 自主转座系统可以应用于水稻功能基因组学的研究当中。

2.3 *Ac/Ds* 转座子割离、插入的检测方法

Ac/Ds 转座子在植物基因组发生的跳跃,既可以发生在体细胞中,这样的跳跃称之为体细胞跳跃(somatic transposition),转座子的跳跃还可以发生在生殖细胞中,这种跳跃方式称之为生殖细胞跳跃(germinal transposition)。在体细胞中发生的跳跃是不能通过遗传作用传到下一代的,转座只有发生在生殖细胞中才能传到下一代中。因为 *Ac/Ds* 系统是将转座子元件割离后再插入到基因组中的,因此对于体细胞跳跃的检测,只要设计特定的 PCR 引物通过扩增得到特定大小的 PCR 产物就可以获得证实 (Suzuki 等, 2001)。而对于在生殖细胞中发生的转座作用,因为其后代是通过遗传而得到的转座插入子,因此植株的所有细胞都是带有转座插入子的,而在原来的转座子供体位点(full donor site, FDS)由于转座子的割离而完全成为空的供体位点(empty donor site, EDS),对于这种转座方式可以由多种方法对其进行检测。一种方法是通过 Southern 杂交的方法,由于是通过遗传的形式得到的转座子插入,因此转座子在所有细胞基因组中插入位置是一样的,杂交显示为表现出与供体位点不同位置的特定条带;也可以通过对空位点 EDS 进行 Southern 杂交检测,显示出单一的 EDS 条带说明跳跃发生在生殖细胞当中。

更为方便的是,研究人员通过将选择一个选择标记基因的启动子与编码区用转座子元件分隔,当没有发生转座时,选择标记基因不表现为抗性,当转座子从原来的位置跳跃后,则启动子就能够正常启动编码区基因,使选择标记基因得到正常表达而获得对特定底物的抗性。通过生殖细胞跳跃其后代获得的抗性是整个植株都表现为抗性的,而如果是体细胞跳跃获得的抗性,只能表现为局部的抗性,并且这种抗性是不能遗传的。

用作转座发生的的标记基因有很多种,一般选用抗除草剂基因和抗抗生素基因,目前已经应用的有: 1、 C^r -ALS 基因是一种对绿磺隆(chlorsulfuron)除草剂具有抗性的乙酰乳酸合成酶(acetolactate synthase, ALS)基因,绿磺隆除草剂的作用机理是抑制支链氨基酸缬氨酸、异亮氨酸、亮氨酸的生物合成,其作用的靶标酶是支链氨基酸合成途径中第一个酶乙酰乳酸合成酶 ALS (Fedoroff 和 Smith, 1993); 2、STP^r(链霉素抗性)基因 (Jones 等 1989); 3、HPT II 基因(hygromycin phosphotransferase gene),潮霉素磷酸转移酶基因 (van Enckevort 等, 2001) 等。有时也有人采用荧光素酶基因作为检测跳跃发生的标记基因。Charnig 等(2000)在构建一个诱导型转座子系统时就采用了荧光素酶基因 LUC

(luciferase)作为标记基因,在其与启动子之间插入了诱导型转座子的元件,当转座子被诱导转座后,荧光素酶基因发生表达,通过荧光的检测就可以确定转座子是否在诱导的条件下发生转座;Greco等(2001)采用绿色荧光蛋白基因GFP作为*Ds*跳跃发生的标记基因。

2.4 影响 *Ac/Ds* 转座子跳跃的因素

Ac 转座酶基因表达水平的高低对 *Ds* 转座子跳跃的频率与发生的时间都有很大的影响。Bancroft 和 Dean(1993)认为 *Ac* 转座酶表达量高则 *Ds* 转座发生的频率就高,反之转座酶基因表达量低则转座发生的频率就低,表明转座发生的频率与 *Ac* 转座酶的表达存剂量效应。他们的实验结果来自于纯合的 *Ac* 转座酶基因及杂合的转座酶对 *Ds* 发生转座影响的研究。但是并不所有的结果都是这样的,Koprek等(2000)在大麦中 *Ac/Ds* 系统中的实验中,采用强启动子 CaMV35S 和弱启动子 *Ac* 本身的启动子分别使转座酶基因得到不同强度的表达,结果表明无论是强启动子还是弱启动子对 *Ds* 发生转座的频率无显著性差异。

转座酶基因的高表达还被证明与转座发生在发育早期阶段是关联的,(Bacells 和 Coupland 1994, Keller 等 1993, Koprek 等 2000, Long 等 1993, Nawagata 等 2000),因此采用强启动子如 CaMV35、UBI 等作为转座酶基因的启动子会使转座酶基因的高表达,导致发生大面积区域发生转座,并且转座后的重新插入的 *Ds* 元件有可能通过生殖细胞而传至下一代,并且下一代绝大部分中得到的 *Ds* 插入子都插入在同一个位置,因此来自于同一 F_1 的 F_2 植物中只需要保留一个 *Ds* 插入子,显然这样的 *Ds* 插入方式对于建立大规模的 *Ds* 插入突变体库是不利的。而采用相对弱的启动子如 *Ac* 转座酶本身的启动子,则可以使转座发生在植物发育比较迟的阶段,这样在来自于同一 F_1 的 F_2 群体中就可能出现很多互不相同的 *Ds* 插入子植株,这对于建立 *Ds* 插入突变体库是有利的。

2.5 *Ac/Ds* 转座子在植物基因组中插入方式及其对策

Ac/Ds 转座子在发生跳跃后重新插入到植物的基因组中有可能插入到与起始位点相连锁的位置,也有可能插入到与起始位点不相连锁的位置,了解这种插入规律对于选择基因突变的策略有很大的帮助。*Ac/Ds* 转座子被认为有很大的倾向插入到与起始位点相连锁的位置,这种倾向在玉米(Dooner 和 Belachew 1989)、番茄(Healy 等 1993)、烟草(Dooner 等 1991)、拟南芥(Bancroft 和 Dean 1993b, Keller 等 1993)中都已报道,在单子叶植物的大麦与水稻中也有类似的报道。在这样一种插入情况下,对于 *Ac/Ds* 原始位点在某个基因的附近,就可以利用转座子的这种特性来使附近的这个基因进行插入失活。但是这种插入方式对于建立大规模

的插入突变体库也是相当不利的,这需要要有很多分散于不同染色体不同位点的起始 *Ac/Ds* 植株。

在拟南芥中还发现 *Ac/Ds* 转座子还有一种倾向是插入到染色体端粒的核仁组织区 NOR(nucleolus organizers), 这就是所谓的插入热点 (hot spot), 这种插入热点的存在对于建立突变体库也是不利的, 但是这种内在的因素很难用序列的改变来消除。

2.6 *Ac/Ds* 转座子系统在水稻中的应用

Ac/Ds 转座子系统应用到水稻中作为插入序列标签的是日本人 Izawa 等 (1991), 他们采用电穿孔法将 *Ac* 转座子导入到水稻的原生质体中, 然后再生获得转基因水稻再生植株, 在愈伤组织阶段转座活性表现活跃, 在不同基因组位置的 *Ac* 转座子表现出不一样的转座性。Nawagata 等 (2000) 对 *Ac/Ds* 的跳跃规律 (包括 *Ds* 割离频率、重新插入频率、独立插入频率等) 作了比较详细的研究。为了研究的方便, 研究人员采用不同的构建方式将 *Ac/Ds* 转座子系统应用到水稻功能基因组学的研究之中 (Chin 等 1999, Enoki 等 1999, Greco 等 2001, Izawa 等 1997)。相对拟南芥来说, 水稻中应用 *Ac/Ds* 转座子系统的研究还远远不够深入, 具体来讲, 还没有建立与拟南芥一样高效筛选跳跃、独立插入的系统, 还没有建立一个稳定插入的转座子插入突变体库, 这些都是今后努力的方向。

3 捕获器系统在植物基因功能研究中的应用

近年来, 采用报告基因随机插入到基因组中来研究基因的功能, 报告基因通常构建成三种形式: 增强子捕获器、启动子捕获器和基因捕获器, 它们都能响应基因组上的顺式作用元件而得到表达 (Springer 1999)。

3.1 增强子捕获器、启动子捕获器和基因捕获器

增强子捕获器 (图1B), 是在报告基因前接上一弱启动子, 通常是含有 TATA 盒和转录起始位点, 它本身不能表达或表达很弱, 但当插入到基因组增强子元件的附件时, 就能被激活而得到表达。因为它没有方向的要求而且增强子在较远距离的位置仍然能起作用, 因此增强子中报告基因的表达率会比较高。

启动子捕获器和基因捕获器是含有一个没有启动子的报告基因, 它只有在插入到转录单位并且方向正确的时候才能得到表达。启动子捕获器 (图1C) 只有在插入到基因的外显子的时候与被插入基因形成融合基因时才能得到转录表达, 而基因捕获器 (图1D) 则是在报告基因前加上一个或多个剪接的受体序列, 当插入到基

因内含子的时候,通过基因的供体位点与捕获器的受体位点发生剪接而转录成融合基因而使报告基因得到表达。由于形成的是融合基因,因此还有可能得到被插入基因在亚细胞中的定位信息。

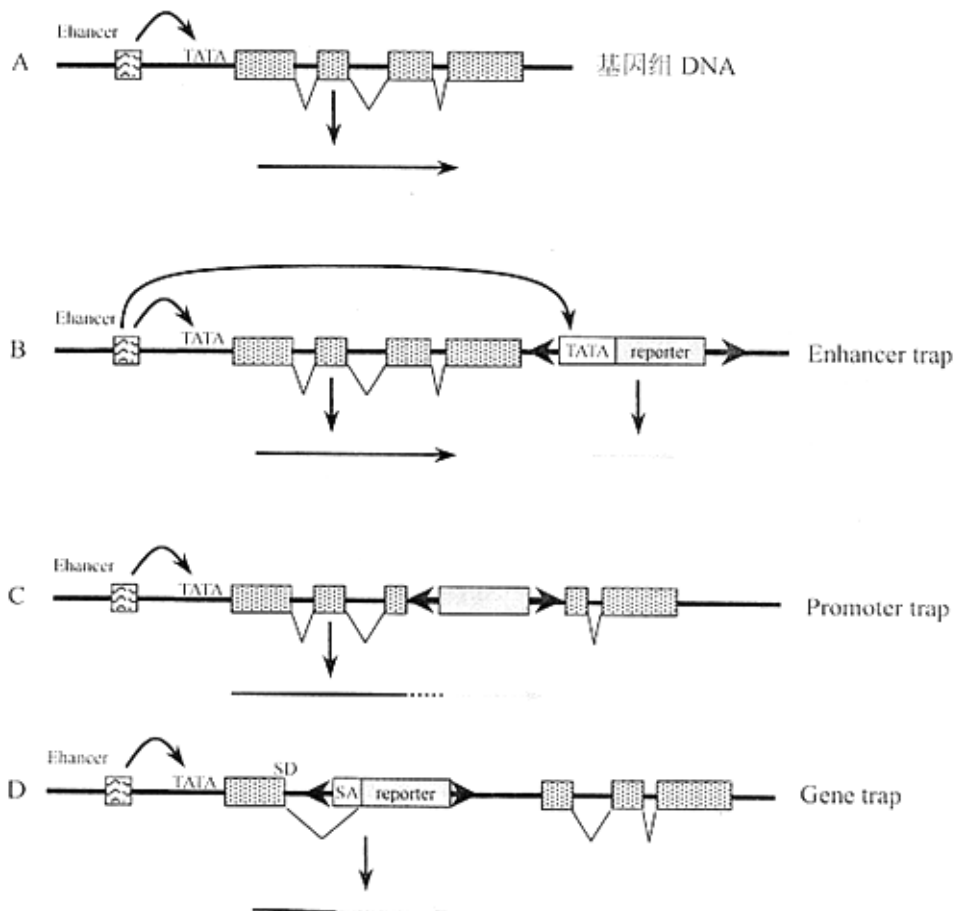


图 1 增强子捕获器、启动子捕获器和基因捕获器的示意图, 摘自 Springer (1999)。

Figure 1. Structure of Enhancer, Gene, and Promoter Trap Vectors.

(A) A generic chromosomal gene with exons (boxes) and introns (lines).

(B) Enhancer trap construct. The minimal promoter of the reporter gene (TATA) is activated by a chromosomal enhancer element, resulting in expression of the reporter gene.

(C) Promoter trap construct. The promoterless reporter gene can be expressed when insertion occurs in an exon so as to result in a transcriptional fusion.

(D) Gene trap construct. The promoterless reporter gene contains splice acceptor (SA) sequences. Expression of the reporter gene occurs upon its insertion into an intron. Splicing from the chromosomal splice donor (SD) site to the SA sequence results in creation of a transcriptional fusion. Arrows in each panel represent the transcripts that are produced as a consequence of insertion.

O'Kane和Gehring(1987)首先在果蝇研究中应用增强子捕获器系统,他们采用*lacZ*作为报告基因,在*lacZ*前接了一个组成型P元件的弱启动子,通过P元件介导将捕获器导入到果蝇基因组中,通过对*lacZ*活性染色而检测基因的表达模式,通过这种方法在果蝇中找到了许多细胞和组织特异表达的基因。之后在小鼠的研究中也采用了捕获器的方法(Allen等1988, Kothary等1988),同时捕获器随机导入到基因组中也由原来的转化方法改进到用转座子的方法来进行,并且还应用了质粒营救的方法来克隆插入位点的基因组序列(Bier等1989, Grossniklaus等1989, Wilson等1989)。由于增强子捕获器很容易得到表达,因此在大基因组且基因密度低的生物中应用不是很合适,而采用基因捕获器或增强子捕获器的方法则更为有效一些(Skarnes 1993, Brennan和Skarnes 1999)。

植物中应用的第一代基因捕获器是采用不带启动子的抗生素基因通过通过农杆菌介导的 T-DNA 转入到烟草中,只有当抗生素基因与烟草基因发生融合表达时,抗相应抗生素并且能再生的植株才能对插入的基因进行研究(André 等 1986, Teeri 等 1986)。第二代捕获器是采用了完整的抗生素基因表达盒加报告基因的方法, GUS 作为报告基因使得目的基因的时空表达可以通过简单的组织化学染色而得到研究(Kertbundit 等 1991, Topping 等 1991)。之后转座子作为捕获器的载体也应用于植物基因的研究中(Fedoroff 和 Smith 1993, Sundaresan 等 1995)。

水稻中应用捕获器作为研究工具相对滞后,因为水稻在同时期转化的技术还不成熟。Chin等(1999)首先应用*Ac/Ds*转座子系统作为基因捕获器的载体来研究水稻基因的时空表达; Jeon等(2000)则采用了T-DNA作为增强子捕获器载体来研究水稻基因的功能。

3.2 捕获器中的报告基因

GUS 基因

GUS 基因 β -Glucuronidase 来自细菌,是在植物中一个很常用的报告基因,它的特点决定了它适合于做为一个报告基因。GUS 基因很稳定,当它与其它蛋白质成为融合基因的时候,仍然能保持相当好的活性(Jefferson 等 1987); GUS 活性可以通过组织化学染色,其酶的底物有很多种,而且都已经商品化; GUS 活性的检测非常灵敏,甚至可能灵敏到单细胞水平(Jefferson 等 1987, Lindsey 等 1993)。但 GUS 基因作为报告基因也有缺点,组织化学染色都是破坏性的,不能进行活体分析;它的底物都很昂贵,如果大规模筛选则成本很高。

绿色荧光蛋白 GFP

绿色荧光蛋白 GFP(green fluorescent protein)来自水母,经过修饰改造后的

GFP 才能在植物中进行有效的应用(Haseloff 和 Amos 1995, Siemering 等 1996, Haseloff 等 1997)。GFP 可以在激发光下直接进行观察,而不象 GUS 需要使用昂贵的底物,因此检测相对就比较便宜。GFP 作为报告基因的另外一个优点就是它可以用来活观察,而不会对植物产生损伤,这样还能对同一植物的同一部位进行多次不同时间段的观察。另外 GFP 还能与目的基因融合而对目的基因产物进行亚细胞定位研究(Chiu 等 1996, Sheen 等 1995)。Haseloff 等(1997)对 GFP 基因的修饰使检测的灵敏度大大升高,使得 GFP 成为一种很好的植物报告基因。

其它报告基因体系还有如玉米的 R 基因家族的 *Lc* 基因,它是 Myc 样转录因子,调节花青素的生物合成,*Lc* 基因的表达会导致花青素的积累,致使呈现出紫色(Lloyd 等 1992, Goldsbrough 等 1996),这种系统最大的好处就是不需要用昂贵的化学底物与检测设备,还能进行活体检测,适合于进行大规模的田间筛选。但是它也有自己的缺点,*Lc* 基因的高表达会导致拟南芥毛状体(trichome)的形成。

3.3 反向选择标记

为了方便选择与原供体位点不相连锁的 *Ds* 转座子插入子,或者是为了选择与 *Ac* 转座酶相分离的转基因植株,就选用了反向选择标记。

IAAH (indoleacetamide hydrolase), 吲哚乙酰胺水解酶是由 Klee 等(1987)在矮牵牛花转基因中开始研究的。Sundaresan 等(1995)则成功的应用到拟南芥中作为 *Ds* 跳跃的反向选择标记。IAAH 基因能将萘乙胺(naphthalene acetamide, NAM)水解成萘乙酸(naphthalene acetic acid, NAA),萘乙酸能促进植株根系的生长,形成毛状根。将 IAAH 置于 *Ds* 转座子外部,当 *Ds* 转座子割离插入到基因组的另外部位,如果不是同一条染色体或是距离足够远的位置,则在后代的分离中,*Ds* 元件就有可能与原来的供体位点相分离。在加入底物的培养基中,只有 *Ds* 元件的植株因为不能将 NAM 分解成 NAA,而保持根的正常形态;而如果在同一植株中既含有 *Ds* 元件,又含有 IAAH 基因,则因为 IAAH 基因将 NAM 分解成 NAA 而造成根系形态的不正常,从而达到分离的目的。

CodA(cytosinedeaminase),胞嘧啶脱胺酶能将对真核细胞无毒的5-氟胞嘧啶(5-fluorocytosine)转变成对真核细胞有毒的5-氟尿嘧啶(5-fluorouracil),当带有 *CodA* 转基因的植株在含有5-氟胞嘧啶的培养基上时,在幼苗期就会停止生长而死去(Stougaard 1993)。因此可以利用 *CodA* 基因的这一特性作为反向选择标记基因,其选择的思路与 IAAH 一样。

链霉菌细胞色素 P450 基因的产物 CyP105A1(cytochrome P450SU1),能将一种除草剂前体磺酰脲(R7402)分解成对植物有毒害作用的除草剂,因此可以用作

反向选择标记(O'Keefe 等 1994)，其选择的思路与 IAAH 一样。

4 克隆 DNA 序列标签旁邻序列的方法

4.1 TAIL-PCR

Thermal asymmetric interlaced (TAIL-) PCR 的发明首先是用来克隆 P1 和 YAC 克隆中插入片段的末端序列(Liu 和 Whitterier 1995)，随后证明这一方法用于克隆 T-DNA 插入位点两侧的拟南芥基因组序列、甚至在基因组序列背景更加复杂的小麦中也是很有有效的(Liu 等 1995)。这种 PCR 要求利用已知的 DNA 序列设计一套半巢式的特异性引物(specific primer) SP1-3，结合一条短的随机简并引物，通过试用不同的退火温度、或高低退火温度的交替使用，使得只有特异引物才能很好地扩增出条带来，一般在进过二轮或者三轮后就能够扩增得到目的条带(图 2)。TAIL-PCR 的特点是对 DNA 质量与数量的要求不高，在 PCR 前或后不需要对样品进行特殊处理，省时省力，特异扩增的 PCR 产物可以直接用来测序。TAIL-PCR 已经成为一种应用最为广泛的克隆 T-DNA 或转座子插入突变体的两侧的基因组序列方法。采用这种方法已经在拟南芥中建立起多个插入位点旁邻序列的数据库。

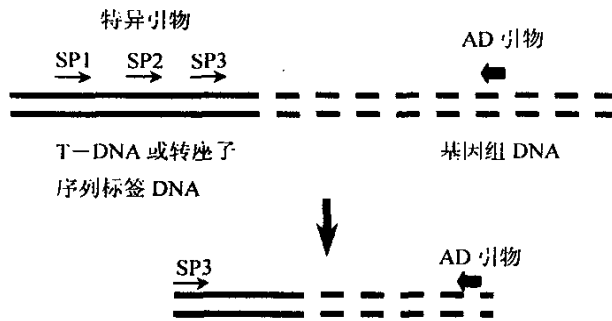


图 2 TAIL-PCR 扩增插入位点旁邻基因组序列的原理示意图。

Figure 2. Schematic representation of amplification of genome sequence flanking insertion site with TAIL-PCR.

4.2 IPCR

图 3 显示了 Inverse PCR 的原理和过程，首先是利用插入标签的序列信息选择合适的限制性内切酶的组合适(例如用 *TaqI*)将插入序列和基因组 DNA 切成合适的片段，再用 T4 连接酶将片段环化，通过特异性嵌套式引物(RB1+LB1，

RB2+LB2) 的两次扩增得到特异的性 PCR 产物。

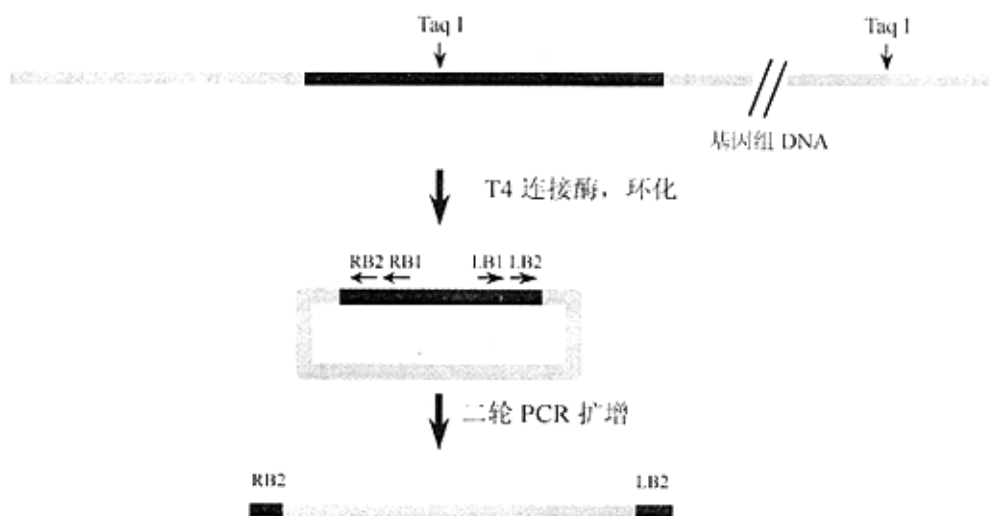


图 3 IPCR 扩增插入位点旁邻基因序列的原理示意图。

Figure 3. Schematic representation of amplification of genome sequence flanking insertion site with IPCR.

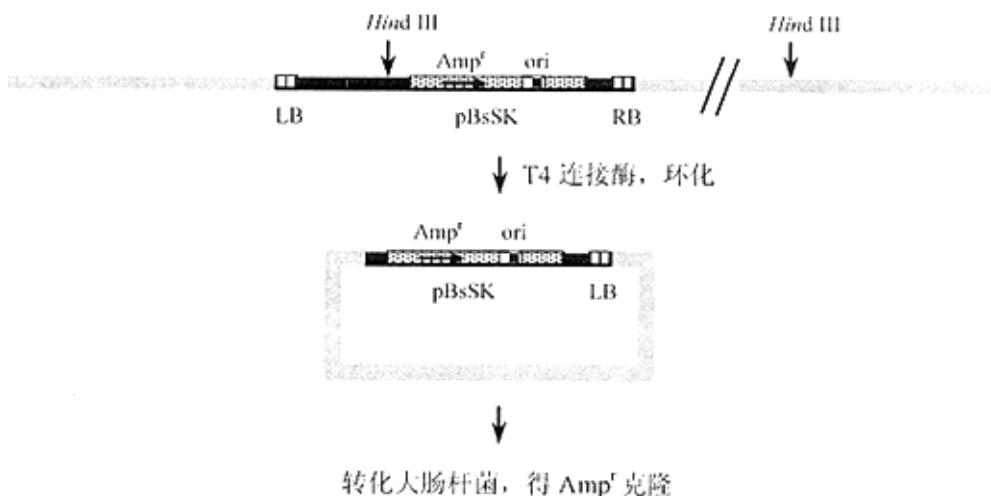


图 4 质粒营救法克隆插入位点旁侧基因组序列的原理示意图。

Figure 4. Schematic representation of cloning genome sequence flanking insertion site with plasmid rescue.

4.3 质粒营救法

图 4 为我们构建的双元载体以通过质粒营救(plasmid rescue)的方法克隆 T-DNA 插入旁侧的水稻基因组 DNA 片段。质粒营救法是将大肠杆菌质粒 pBsSK

插入到 T-DNA 序列标签中通过农杆菌介导的和 T-DNA 转基因一起转入到水稻基因组中, 然后利用适当的限制性内切酶 *Hind* III 在不破坏质粒结构的条件下酶切基因组 DNA, 通过 T4 连接酶将带有水稻基因组 DNA 片段的质粒环化, 转化大肠杆菌以长出抗性克隆就能得到含有插入位点基因组旁邻序列的质粒。

第二章 *Ac/Ds*-T-DNA 原初插入位点的分析

1 材料与方法

1.1 材料

水稻: 粳稻中花-11 号、日本晴。

菌株: 大肠杆菌 DH5 α , 农杆菌菌株 EHA105。

质粒: pCAMBIA1300, CD3-125, CD3-126, pAHC25, pDsbar1300, pUBITs。

工具酶: *Hind* III、*Sac* I、*Pst* I、*Xba* I、*Bam*HI、*Eco*R I、T4 DNA 聚合酶、T4 DNA 连接酶、CIAP 碱性磷酸酶。

Taq DNA polymerase。

PCR 产物克隆载体: pUC-T 载体。

1.2 方法

分子生物学操作按分子克隆手册(萨母布鲁克 等 1995)和精编分子生物学实验指南(奥斯伯 等 1998)进行。

1.2.1 载体构建

二元载体 pDsE、pDsG 与 pAc 的构建。

pDsE 的构建: CD3-125 用 *Xba*I 和 *Hind* III 酶切切除 NPT II 基因表达盒, 用 *Hind* III 酶切 pDsbar1300 酶切切取 bar 基因表达盒, 两者均用 T4 DNA 聚合酶补平, 用 T4 DNA 连接酶连接, 得 CD3-125bar 质粒。用 *Sac* I 酶切 CD3-125bar 切取 *Ds* 元件, 插入 pCAMBIA1300 的 *Sac* I 位点, 得 pDsE。

pDsG 的构建: 除了 CD3-126 替代 CD3-125 外, 其它方法与步骤同 pDsE 的构建一样。

pAc 的构建: pCAMBIA1300 和 pAHC25 两者都用 *Eco*R I 和 *Bam*HI 双酶切, 前者回收载体部分, 后者回收 GUS-Nos 片段, 连接转化后得 pGUSNos1300; pGUSNos1300 和 pUBITs 分别用 *Sac* I 和 *Bam*HI 双酶切, 前者回收载体, 后者回收 Ts 片段, 两者连接得 pTsNos; pTsNos 用 *Bam*HI 单酶切, 用 T4 DNA 聚合酶补平, pAHC25 用 *Pst* I 酶切回收 UBI 启动子片段, T4 DNA 连接酶补平, 然后 UBI 启动子片段与补平后的 pTsNos 连接得 pAc。

构建完成后的载体线性示意图如图 1 所示, 质粒图见附录一。

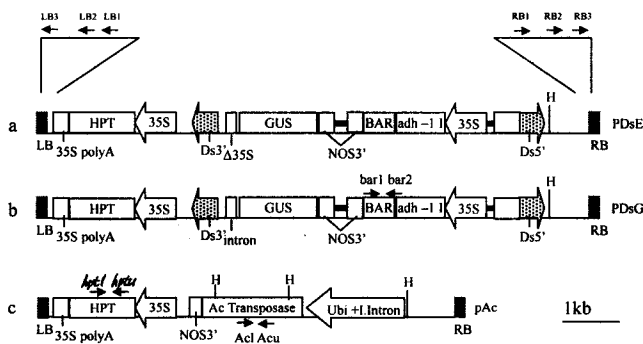


图 1 双元载体 pDsE、pDsG 和 pAc 的线性示意图

Figure 1. Schematic representation of the T-DNA containing the *Ds* element harboring gene trap (DsE) (a) and enhancer trap (DsG) (b), and *Ac*TPase (c), as well as selectable markers associated with each. All three vectors are based on pCambia1300. Abbreviations: LB, T-DNA left border; RB, T-DNA right border; HPT, hygromycin phosphotransferase; 35S, CaMV35S promoter; NOS3', 3'-end signal of nopaline synthase; GUS, β -glucuronidase; 35S polyA, CaMV35S polyA signal; BAR: phosphinothricin acetyltransferase, provides resistance to Basta; Ds5', 5' termini of dissociation from maize, Ds3', 3' termini of dissociation from maize, Δ 35S, CaMV 35S mini promoter; intron, intron of gene for G-protein α -subunit from *Arabidopsis*; UBI/E/I, maize ubi-1 promoter/Exon/Intron; *Ac* transposase, transposase of activator from maize. bar1, bar2, hpt1, hpt2, hpt3, acu, ac1 and LB1, 2, 3 and RB1, 2, 3 are primers used in PCR analysis and TAIL-PCR.

1.2.2 水稻转基因

水稻转基因采用 Hiei(1994) 的方法步骤。转基因的受体材料为水稻中花-11 号和日本晴幼胚诱导的愈伤组织; 转基因用双元载体 pAc、pDsE 和 pDsG 三种质粒; 农杆菌采用 EHA105; 50mg/L 潮霉素用作筛选物质, 再生的水稻苗长至 20cm 高时移栽到水稻溶液培养中 7-10 天, 在水稻栽培季节时移栽到土壤中自然生长, 在非水稻栽培季节中, 生长在日温 30℃/夜温 20℃、光周期为 12 光照/12 小时黑暗生长, 至结实收获。

1.2.3 T₁ 代转基因水稻的潮霉素抗性分析

消毒种子后, 播种到水稻培养液(蒋德安和曾广文 1998)的固体培养基上 (pH5.5, 琼脂浓度为 0.7%), 光周期为 14 小时光照/10 小时黑暗, 发芽后 10 天考察水稻的潮霉素抗性。

1.2.4 转基因水稻的 PCR 检测

检测 HPT 基因的引物设计为 hptu (5'-cggctcgcggaggctatggatg-3')和 hptl (5'-gcttctcgcggcgatttgta-3'), 扩增产物为 599bp。反应采用 20μl 体积(50ng DNA, 100μM dNTP, 0.5μM hptu 和 hptl, 1.5mM MgCl₂, 50mM KCl, 0.1% Triton X-100, 10mM Tris-HCl, pH9.0, 1 unit Taq DNA polymerase); 热循环条件为: 94℃ 4min; 94℃ 20sec, 62℃ 30sec, 72℃ 1min, 35 cycles; 72℃ 5min。

检测 BAR 基因的引物设计为 barl (5'-atgagcccagaacgacgcc-3')和 bar2 (5'-tcagatctcggtagcgggca-3'), 扩增产物大小为 354bp, 反应采用 20μl 体积(50ng DNA, 100μM dNTP, 0.5μM barl 和 bar2, 1.5mM MgCl₂, 50mM KCl, 0.1% Triton X-100, 10mM Tris-HCl, pH9.0, 1 unit Taq DNA polymerase); 热循环条件为: 94℃ 4min; 94℃ 20sec, 60℃ 30sec, 72℃ 1min, 35 cycles; 72℃ 5min。

检测 *Ac* 转座酶基因的引物为 Acu (5'-agtgtcctgcgaatcctcaat-3')和 Acl (5'-gtctctcgcgcctcactt-3'), 扩增产物大小为 486bp, 反应体系及 PCR 扩增条件同 BAR 基因检测的一样。

1.2.5 转基因水稻的 Southern 杂交分析

水稻总 DNA 的大量提取法卢扬江(1992)的方法, 5μg 总 DNA 用 *EcoR* I 或 *Hind* III 酶切完全, 在 0.7%的琼脂糖凝胶中电泳分离, 转移到带正电的尼龙膜 (Hybond-N+, Amersham)上, 用 ³²P 放射性同位素标记的 HPT 基因(与亲本杂交)。*Hind* III 在载体上的分布如图一所示, 在 HPT 基因和 *Ds* 上没有酶切位点, 杂交

的条带数反应 T-DNA 或 *Ds* 的拷贝数。

1.2.6 T-DNA 旁邻序列的克隆与测序

采用 Tail-PCR 的方法(Liu 等, 1995)来克隆 T-DNA 插入子两侧的序列。左右边界的特异引物设计采用 DNASTAR 软件 PRIMERSELECT 4.05 程序, 右边界引物用 RB1: 5'-gcaccgatcgcccttcccaaca-3', RB2: 5'-ccaacagttgcgcagcctgaa tg-3', RB3: 5'-ccaacagttgcgcagcctgaatg-3', 左边界引物用 LB1: 5'-tcgtccgagggcaaa gaaataga-3', 5'-taaaatccagatccccgaatta-3', 5'-ccgcaatgtgtatttaagttgtc-3', 随机引物采用如同(Liu 等 1995)中的引物。AD1, NTCGA(G/C)T(A/T)T(G/C)G(A/T)GTT; AD2, (G/C)TTGNTA(G/C)TNCTNTGC; AD3, (A/T)GTGNAG(A/T)ANCANAGA。反应条件参照 Liu 等(1995)的方法进行。特异扩增得到的第三轮 PCR 产物经凝胶电泳回收后克隆到 pUC-T 载体中, 用 M13 测序引物及发玛西亚测序试剂盒(DYEnamic ET dye terminator kit)作反应在 MEGABASE1000 测序仪上测序。

1.2.7 T-DNA 旁邻序列分析

在水稻基因组中的插入位置, 数据库 GenBank 中所有的水稻序列 nr (所有 GenBank+RefSeq Nucleotides+EMBL+DDBJ+PDB 序列, 不包含 EST、STS、GSS 和阶段 0、1、2 的 HTGS 序列), gss (genome survey sequence, include) and htgs (high throughput genomic sequence), 本地分析在 DNASTAR (version 5.0)中进行。核酸序列的同源分析按相似性>95% 或者 score>300 为准, 对于有些片段比较短的序列, 按人工判断是否为水稻序列。

2 结果与分析

2.1 T-DNA 插入的转基因水稻

本研究中我们采用农杆菌 EHA105 介导的水稻 *Ac/Ds*-T-DNA 转基因。共用了三种双元载体的质粒 pDsE、pDsG 和 pAc, 选用潮霉素基因 HPT 为选择标记, 总共获得转基因可育水稻 1188 株。其中来自日本晴的转基因水稻为 526 株包括 Nac(转 pAc)51 株, NE(转 pDsE)361 株, NG(转 pDsG)114 株, 来自中花-11 号的转基因水稻为 662 株, 包括 CAc(转 pAc)176 株, CE (转 pDsE) 320 株, CG (转 pDsG) 166 株。

2.2 转基因水稻的 PCR 检测

对于所有的转基因植株, 都可以采用 HPT 基因的扩增进行检测。图 2 是表

明对中花-11 号的转基因水稻的 PCR 检测, 1-19 是对 19 株独立的转基因水稻植株, CK 为未经过转基因的正常中花-11 号水稻植株, 扩增得到的 PCR 产物与设计产物 HPT 基因 599bp 片段的大小一致, 结果表明采用设计的 HPT 基因引物作为检测转基因水稻其特异性是非常好的。在对转基因水稻的 PCR 检测中, HPT 的阳性率在 90% 以上, 说明我们的转基因系统是非常有效的。

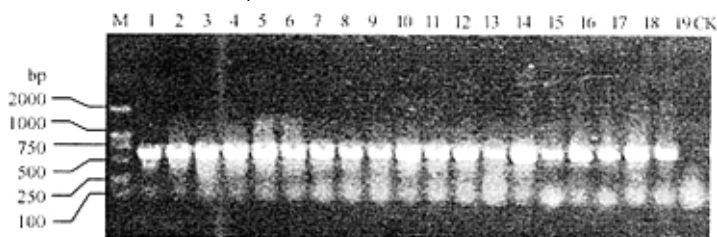


图2 转 *Ac/Ds*-T-DNA 水稻植株 HPT 基因的 PCR 检测。

Figure 2. PCR identification of HPT gene in *Ac/Ds*-T-DNA transformation rice plants.

转 pDsE 和 pDsG 的转基因水稻还可以采用 BAR 基因的 PCR 扩增进行检测。图 3 是表示转基因水稻扩增 BAR 基因检测转基因水稻, 结果在阴性对照中也出现了一个非特异扩增的条带, 大约为 500bp, 虽然在转基因水稻中扩增出 BAR 基因的与阳性对照质粒一样的特异条带, 但由于非特异条带的干扰, 所以在以后的检测中就不采用 BAR 基因片段的扩增作为检测的依据。

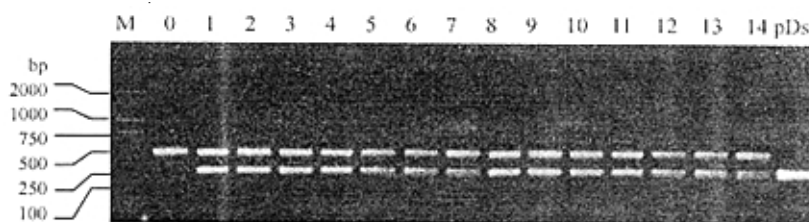


图3 转 *Ds*-T-DNA 水稻植株 BAR 基因的 PCR 检测。

Figure 3. PCR identification of BAR gene in *Ds*-T-DNA transformation rice plants.

0, Zhong-11 as negative CK; 1-14, *Ds*-T-DNA transformation rice plants, pDs, pDsE plasmid as positive CK.

转 pAc 的转基因植株可采用扩增 *Ac* 转座酶片段来进行进一步的验证。图 4 是对 22 株独立的转 pAc 水稻植株, CK 为未经过转基因的正常中花-11 水稻植株, 扩增得到的 PCR 产物与设计产物 486bp 大小一致, 结果表明采用设计的 *Ac* 转座酶基因引物作为检测转 pAc 水稻其特异性能保证准确的检测。与 HPT 基因扩增结果相比较发现, 凡是扩增到 *Ac* 转座酶基因片段的, 都能扩增到 HPT 基因片段, 说明 HPT 基因与 *Ac* 元件很好的保持完整性转入到了水稻基因组 DNA 之中。

中。

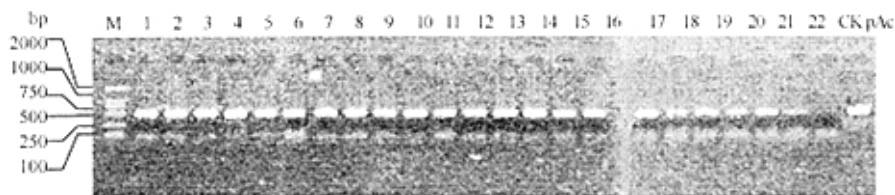


图 4 转 *Ac*-T-DNA 水稻植株 *Ac* 转座酶基因片段的 PCR 检测。

Figure 4. PCR identification of *Ac* transposase gene fragment in *Ac*-T-DNA transformation rice plants.

2.3 *Ac/Ds*-T-DNA 转基因水稻的拷贝数分析

选取 30 个转 *Ac*-T-DNA, 72 个转 *DsE*-T-DNA 和 22 个转 *DsG*-T-DNA 总共 124 个独立 T_1 代中花-11 号转基因水稻进行潮霉素抗性的试验。其中有 92 个株系 Hyg^r 表现为 3:1 分离, 说明水稻基因组 DNA 的一个位点上插入有 T-DNA, 另外 32 个株系在水稻染色体互不连锁的不同位点上插入有二个或二个以上的 T-DNA (见表 1)。124 独立株系的 T-DNA 在水稻染色体上插入位点的平均数为 1.3 个, 这一结果与 Jeon(2000)报道的在水稻中平均为 1.4 个 T-DNA 插入位点数相近。

通过 Southern 杂交分析, 在总共 48 个 T_0 代转基因水稻植株中有 26 个只有单拷贝的 HPT 基因, 12 个水稻植株中含有 2 个拷贝的 HPT 基因, 还有 10 转基因水稻植株含有 3 个或 3 个以上的 HPT 基因拷贝。这个结果表明大约有一半的转基因水稻植株只含有一个拷贝的 T-DNA 插入。通过 Southern 杂交分析, HPT 基因的平均拷贝数为 1.8 个。Jeon(2000)报道的在水稻中平均为 2.2 个 GUS 基因和 2.1 个 HPT 基因插入拷贝数

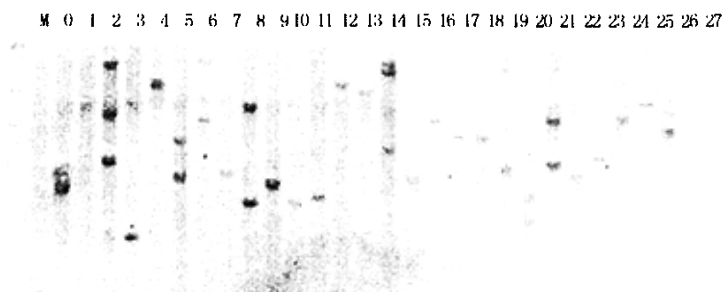


图 5 *Ac/Ds*-T-DNA 转基因水稻的 Southern 杂交分析。

Figure 5. Southern blot analysis for HPT copy number in *Ac/Ds*-T-DNA transformation rice plants. 0: Zhonghua-11, 1-27 transgenic rice line.

表 1 转基因水稻 T-DNA 插入拷贝位点数的分析
Table 1. Estimation of T-DNA loci numbers in transgenic rice

Hygromycin resistance in T1 generation					
Transfromants		resistant	sensitive	genetic loci	χ^2
1	Ac04	16	6	1	0.0
2	Ac22	17	11	1	2.3
3	Ac24	18	10	1	1.2
4	Ac25	20	5	1	0.1
5	Ac26	9	4	1	0.0
6	Ac30	12	8	1	1.7
7	Ac31	17	6	1	0.0
8	Ac32	19	8	1	0.1
9	Ac39	14	8	1	1.0
10	Ac45	19	3	1	1.0
11	Ac46	31	10	1	0.0
12	Ac48	28	8	1	0.0
13	Ac50	13	12	1	5.9
14	Ac52	25	9	1	0.0
15	Ac53	20	4	1	0.5
16	Ac54	7	3	1	0.0
17	Ac55	13	5	1	0.0
18	Ac57	14	3	1	0.2
19	Ac60	12	5	1	0.0
20	Ac63	6	5	1	1.5
21	Ac65	9	3	1	0.1
22	Ac68	4	2	1	0.0
23	Ac69	7	4	1	0.3
24	Ac70	13	7	1	0.6
25	Ac71	17	4	1	0.1
26	Ac74	11	4	1	0.0
27	Ac77	14	7	1	0.4
28	Ac80	16	4	1	0.1
29	CE011	17	7	1	0.1
30	CE013	14	8	1	1.0
31	CE016	14	4	1	0.0
32	CE022	23	6	1	0.1
33	CE026	23	10	1	0.3
34	CE027	22	8	1	0.0
35	CE028	19	7	1	0.0
36	CE029	13	3	1	0.1
37	CE031	15	5	1	0.1
38	CE034	14	7	1	0.4
39	CE035	25	11	1	0.3
40	CE037	34	10	1	0.0
41	CE038	20	7	1	0.0
42	CE042	22	9	1	0.1
43	CE045	20	4	1	0.5
44	CE050	16	6	1	0.0
45	CE053	23	9	1	0.0
46	CE055	12	5	1	0.0
47	CE056	14	7	1	0.4
48	CE059	19	7	1	0.0
49	CE061	18	5	1	0.0
50	CE063	17	5	1	0.0
51	CE064	26	13	1	1.0
52	CE065	29	8	1	0.1
53	CE066	19	4	1	0.4
54	CE067	21	4	1	0.7
55	CE068	26	7	1	0.1
56	CE070	30	7	1	0.4
57	CE071	32	8	1	0.3
58	CE072	22	9	1	0.1
59	CE076	19	4	1	0.4
60	CE078	33	10	1	0.0
61	CE079	11	7	1	1.2
62	CE080	32	15	1	0.9
63	CE081	21	7	1	0.0
64	CE082	19	7	1	0.0

Table 1. continued

Hygromycin resistance in T1 generation					
Transfromants		resistant	sensitive	genetic loci	χ^2
65	CE087	14	5	1	0.0
66	CE088	11	3	1	0.0
67	CE090	12	9	1	2.7
68	CE091	15	3	1	0.3
69	CE093	22	16	1	5.1
70	CE098	15	3	1	0.3
71	CE099	17	11	1	2.3
72	CE102	14	5	1	0.0
73	CE104	15	10	1	2.3
74	CE105	24	12	1	0.9
75	CE106	19	10	1	0.9
76	CE107	19	9	1	0.4
77	CE108	17	8	1	0.3
78	CE109	23	15	1	3.5
79	CE120	10	10	1	5.4
80	CG02	21	6	1	0.0
81	CG07	16	4	1	0.1
82	CG11	17	8	1	0.3
83	CG23	19	4	1	0.4
84	CG26	23	5	1	0.4
85	CG29	23	6	1	0.1
86	CG30	24	13	1	1.5
87	CG32	36	17	1	1.1
88	CG33	19	5	1	0.1
89	CG38	31	6	1	1.1
90	CG39	20	6	1	0.0
91	CG40	13	4	1	0.0
92	CG55	14	3	1	0.2
93	Ac01	24	2	2	0.0
94	Ac56	21	3	2	0.7
95	CE007	28	4	2	1.2
96	CE008	11	2	2	0.6
97	CE009	20	2	2	0.0
98	CE014	37	3	2	0.0
99	CE015	29	5	2	2.8
100	CE018	23	1	2	0.0
101	CE030	23	2	2	0.0
102	CE039	27	3	2	0.2
103	CE046	19	2	2	0.0
104	CE048	27	4	2	1.3
105	CE051	23	4	2	2.1
106	CE058	32	1	2	0.2
107	CE060	20	1	2	0.0
108	CE064	30	3	2	0.1
109	CE073	20	3	2	0.8
110	CE074	28	3	2	0.2
111	CE075	34	1	2	0.2
112	CE077	28	5	2	3.1
113	CE110	18	2	2	0.1
114	CE112	20	1	2	0.0
115	CG03	16	2	2	0.1
116	CG08	24	3	2	0.4
117	CG16	16	2	2	0.1
118	CG21	22	3	2	0.6
119	CG25	15	2	2	0.2
120	CG34	23	4	2	2.1
121	CG43	33	2	2	0.0
122	CG47	25	3	2	0.3
123	CE019	25	0	2	0.8
124	CG27	28	0	2	1.0
Average				1.3	

The number of loci was determined by hygromycin-resistant ratio that give the smallest χ^2 value.

$$\text{For 1 copy, } \chi^2 = \frac{(|A - 3a| - 2)^2}{3n}, \text{ for 2 copies, } \chi^2 = \frac{(|A - 15a| - 8)^2}{15n}.$$

相比为较小。而通过 T_1 代转基因水稻的潮霉素抗性分析获知 T-DNA 在基因组中的平均插入位点数为 1.3 个。出现 T-DNA 插入位点数比 HPT 基因实际拷贝数要少这样的结果数可能是因为 T-DNA 在插入到植物基因组 DNA 时发生串联重复。

表 2 HPT 基因在 T_0 代转基因水稻中的拷贝数
Table 2. Statistical of copy number of HPT gene in T_0 transgenic plants

序号	转基因植株	HPT 拷贝数
1	CAC46	1
2	CAC52	1
3	CE037	1
4	CE070	1
5	CE080	1
6	CE301	1
7	CE307	1
8	CE308a	1
9	CE310	1
10	CE312	1
11	CE313	1
12	CE314	1
13	CE326	1
14	CE327	1
15	CG038	1
16	CG201	1
17	CG203a	1
18	CG204	1
19	CG206	1
20	CG207b	1
21	CG229	1
22	CG235	1
23	CG241	1
24	CG242	1
25	CG245	1
26	CG246	1
27	CE059	2
28	CE072	2
29	CE078	2
30	CE201	2
31	CE202b	2
32	CE308b	2
33	CE323	2
34	CE324	2
35	CE325	2
36	CE328	2
37	CG203b	2
38	CG207a	2
39	CE302	3
40	CE309a	3
41	CE311	3
42	CE315a	3
43	CE319	3
44	CG240	3
45	CE014	4
46	CE202a	5
47	CE315b	5
48	CE309b	6
average		1.8

2.4 转基因水稻 T-DNA 插入位点处基因组 DNA 的分析

为研究 T-DNA 在水稻基因组 DNA 中的插入规律,通过 TAIL-PCR 的方法扩增得到 T-DNA 两侧的水稻基因组序列,为了得到更多的 T-DNA 边界的序列特征,将特异扩增的 PCR 产物克隆到 pUC-T 载体中,然后进行测序。总共得到有效序列 182 条。

在 182 条序列中,检测到有水稻序列的有 126 条。对插入位点两侧各 250bp(总共 500bp)的水稻序列 GC 含量进行分析(图 6)。插入位点的水稻基因组 DNA 平均 GC 含量为 42.5%,这一结果与水稻的基因组平均 GC%含量 43.3%的值比较接近。有超过 83%的 T-DNA 插入子落在水稻 GC%含量为 30-50%的基因组区域,而只有很少的一部分 T-DNA 插入到基因组 GC 含量高于 60%或低于 30%的区域。从 T-DNA 插入位点 GC 含量与水稻基因组 GC 含量的分布比较来看, T-DNA 在水稻基因组中的插入是随机的,有可能插入到基因组 DNA 的任意区域。

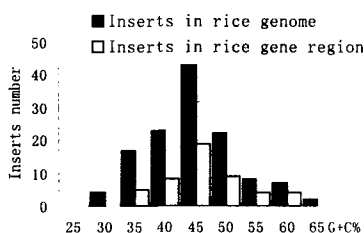


图 6 *Ac/Ds*-T-DNA 插入位点处两侧 500bp 水稻 DNA 序列的 GC%分布。

Figure 6 Distribution of GC% content of 500bp DNA sequence flanking insertion site

在总共 126 个分析的 T-DNA 插入子中,有 50 个 T-DNA 插入到水稻的基因区,插入到基因区的频率为 40%。从图 6 中可以看出, T-DNA 插入到水稻基因区的 GC%含量的分布与基因组 DNA 的分布基本一致。表 3 列出了部分插入到水稻基因区的 T-DNA 插入植株,关于所有 T-DNA 插入子在基因组 DNA 的具体位置见(<http://www.genomics.zju.edu.cn/ricetdna.html>)。T-DNA 在水稻基因组 DNA 插入处的特点详细分析见另一博士学位论文,其中有详细的阐述,在此不展开讨论。

表 3 T-DNA 在水稻基因中的插入
Table 3. T-DNA insertion in rice gene

Line	GenBank accession Number	Chr.	Score value	Insertion position	Similarity to gene or putative gene	Position in gene
CE5-1	AP003353	1	141	93536	phospholipase-like protein	intron
CE22-1	AP003247	1	254	34129	receptor protein kinase -like	last intron
CE27-1	AP003204	1	438	25633	putative 3-dehydroquinate dehydratase	third intron
CE29-4	AP003256	1	287	71774	putative dyenin light chain, A13g48360	promoter up214b
CE32-1	AP003196	1	531	15534	hypothetical protein	third intron
CE39-1	AP002868	1	113	42921	unknown protein	first intron
CE42-9	AP003416	1	580	19814	hypothetical protein	second intron
CE33-1	AL606441	4	234	91990	predicted gene	intron
CE70-4	AP004643	4	581	29302	glucose-6-phosphate isomerase	promoter
CE72-1	AP003846	7	825	73309	hypothetical protein	promoter
CE87-2	AL662957	4	698	46029	putative glutamate receptor	promoter
CE78-7	AP003380	1	589	93300	unknown e protein	Second intron
CE79-6	AP005555	9	452	2533	putative wall-associated kinase 2	promoter
CE90-3	AC126221	3	547	42341	glyoxylate aminotransferase 2 homolog	promoter
CE429	AP003995	7	315	66620	unknown protein	exon
CA601-1	AP004037	2	1298	117222	putative reverse transcriptase	exon
CA620-2	AL662957	4	613	16603	predicted gene	intron
CA621-1	AC121366	5	486	73601	putative retroelement	exon
CA627-1	AP003928	8	256	130371	predicted gene	exon
CA635-2	AC090482	10	575	67094	putative DHHC-type zinc finger protein	intron
CA636-1	AP005656	9	1128	103643	pyrophosphate-fructose-6-phosphate 1-phosphotransferase	exon
CA648-1	AC115687	3	385	30879	predicted gene	intron
CA652-1	AF377947	3	672	16974	unknown protein	exon
CA653-3	AP003853	1	414	50521	predicted gene	exon
CG02-5	AC060755	10	511	68975	ribosomal protein L2	fifth intron
CG34-2	AP004837	2	509	39909	putative gypsy-type retrotransposon protein	exon
CG38-3	AF377947	3	680	16584	unknown protein	exon
CG42-2	AP002855	1	531	79621	hypothetical protein	promoter

2.5 T-DNA 左、右边界在基因组 DNA 中的插入特点分析

2.5.1 T-DNA 左、右边界重复序列的缺失

T-DNA 在整合到植物基因组的时候通常保留右边界重复序列的最后 3 个碱基 TGA 特征序列。在 182 条序列中,检测到 T-DNA 右边界有 158 条,其中保留有完整特征序列 TGA 的有 77 例 (49%), 约占总数的一半, 保留 TG 的有 22 例, 保留 T 的有 13 例, 这三种插入情况合在一起占的比例为 71%。剩下的 46 例都有不同程度的碱基缺失情况, 详细见图 7。说明 T-DNA 右边界通常是以精确的插入方式整合到水稻基因中的, 但是在某些况下也以不精确的插入方式整合到基因组 DNA 中。

T-DNA	right border
AGTTTAACTATCAGTGT	ttgacaggatataattggcggglaaacctaagagaaaagagcgttta
AGTTTAACTATCAGTGT	tga*****77
AGTTTAACTATCAGTGT	tg*****22
AGTTTAACTATCAGTGT	t*****13
AGTTTAACTATCAGTGT	*****11
AGTTTAACTATCAGTGT	*****5
AGTTTAACTATCAGTGT	*****2
AGTTTAACTATCAGT	*****5
AGTTTAACTATCAG	*****5
AGTTTAACTATCA	*****4
AGTTTAACTATC	*****2
AGTTTAACTAT	*****1
AGTTTAACTA	*****2

其它 9 例

图 7 T-DNA 右边界重复序列的缺失。

Figure 7. Deletion of nucleotides in repeat sequence of T-DNA right border.

	T-DNA	left border
pDsE	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTACACCACAATATATCCTGCCA	
Cac52	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATGggtccaacggaacaagggtggcgctc	
Cac53	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTACACCACgccaataagtgtgt	
Cac63	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCgcccatttcaaaccttgccaagatcatgcg	
Cac64	TCCGCAATGTGTTATTAAGTTGTCTAAGCcectctcttccacgcttgcctggcaatctcgctc	
CE002	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTggttagaagacatgttaatttg	
CE011	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTACACCACAATATATCCTGCCA	
CE042	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTACGGCAtggtattctgctccac	
CE053	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTACACCACAATATATCCTCtccc	
CE087	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTAtgttgattcaggtacagttc	
CE090	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTGTTTlggttagaagacatgttaattt	
CG008	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATTNGTTTAtgttgattcaggtacagttc	
CG016	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATT-GTTTACACCACAATtgcctgggagga	
CG038	TCCGCAATGTGTTATTAAGTTGTCTAAGCGTCAATcggtccaacggaacaagggtggcgct	
CG042	TCCGCAATGTGTTATTAAGT-GTCTAAGCGTCAATT-GTTTACAggttcatgttagccttgc	
CG043	TAagtggtgaatccgcacctctggcaaccgggtgaagggttatctctatgaactgtgcgtca	

图 8 T-DNA 左边界重复序列的缺失。

Figure 8. Deletion of nucleotides in repeat sequence of T-DNA left border. Uppercase letters represent T-DNA, lowercase letters represent rice DNA.

在 182 条序列中, 检测到左边界的有 40 条, 其中 RB-LB 串联的有 16 条 (同时包含了左边界与右边界)。通过对 40 个 T-DNA 左边界序列的分析发现, T-DNA 左边界碱基的变化比较大, 没有发现某一个碱基处断裂的比率比较高。图 8 中列出了部分 T-DNA 插入子左边界 T-DNA 序列重复序列的缺失情况。

2.5.2 T-DNA 左边界小同源性序列的分析

通过对 10 个 T-DNA 左边界与水稻基因组 DNA 直接相联的序列分析, 发现 T-DNA 左边界与水稻基因组 DNA 的接合处有小同源性序列(microhomology)的存在, 同源性的碱基序列的长度为 1-8bp, 它们被认为在 T-DNA 以不规则同源重组的方式整合到植株基因组 DNA 时起着重要的作用 (图 9)。

	T-DNA	小同源性序列	水稻 DNA
NE143	...TGTGTTATTAAGTTGTCTAAGCGTCAATTTG	<i>TTTATGCG</i>	...TGTGTTATTAAGTTGTCTAAGCGTCAATTTG
CG8-6	...CAATGTGTTATTAAGTTGTCTAAGCGTCAA	<i>TTTATGCG</i>	...CAATGTGTTATTAAGTTGTCTAAGCGTCAA
NE651	...CGTTAATTCAGTACATTAACGTCGCA	<i>ATGTCG</i>	...CGTTAATTCAGTACATTAACGTCGCA
NE455	...TGTGTTATCAATTTGTCTAAGCGTCAATTTG	<i>TTTATGCG</i>	...TGTGTTATCAATTTGTCTAAGCGTCAATTTG
CE53-1	...TCTAAGCGTCAATTTGTTTACACCAATATA	<i>TCTCTCT</i>	...TCTAAGCGTCAATTTGTTTACACCAATATA
NE830	...CCGAATTAATTCGGCGTTAATTCAGTACATTA	<i>AAAGGTT</i>	...CCGAATTAATTCGGCGTTAATTCAGTACATTA
CE90-3	...CGCAATGTGTTATTAAGTTGTCTAAGCGTCAA	<i>TTTGTGTT</i>	...CGCAATGTGTTATTAAGTTGTCTAAGCGTCAA
NE534	...TTACTAAGTTGTCTAAGCGTCAATTTGTTTACAC	<i>CACACCT</i>	...TTACTAAGTTGTCTAAGCGTCAATTTGTTTACAC
CE11-1	...ACAATATATCTGCCACGACGCAACAGC	<i>TC</i>	...ACAATATATCTGCCACGACGCAACAGC
CE87-2	...CCCAATGGTAAGAGGAAGTGTACCTGAATCCAC	<i>AAACAAAT</i>	...CCCAATGGTAAGAGGAAGTGTACCTGAATCCAC

图 9 T-DNA 左边界序列与水稻基因组 DNA 的小同源序列。

Figure 9. Microhomologous oligonucleotides at T-DNA left border/rice genomic DNA junction. Bold and italic letters representing microhomology.

2.5.3 T-DNA RB-LB 串联情况的分析

T-DNA 在整合到植物基因组时,不但会以独立的单个 T-DNA 插入到基因组 DNA 中,还会两个甚至两个以上的 T-DNA 以串联形式插入到植物基因中。在 182 条序列中发现有 14 例是 T-DNA 发生 RB-LB 的直接串联,中间没有任何的填充 DNA 的存在(图 10)。但在另外 2 例 RB-LB 的串联情况中,则发现有 filler-DNA 的存在(图 11)。其中 CE420 的 15bp 的 filler-DNA 前 14bp 虽然能在水稻基因组找到 100% 的同源性,且只有 7bp 的碱基能在载体中找到同源性,但也不能推测 filler-DNA 一定是来源于水稻的。CE302 的 32bp 的 filler-DNA 则无法从载体 DNA 和水稻基因组 DNA 找到完全同源的序列,也无法从农杆菌的基因组 DNA 中找到同源序列,推测可能是这些序列是由多个 DNA 片段连接而成的。

系号	T-DNA (R)		T-DNA (L)
NE398	CTATCAGTGTTTGA	0bp	ACTTAATAAC
NE527	CTATCAGTGTTTG	0bp	TGGTGTAACAAATTGACGCTTAGACA
NE666	CTATCAGTGTTTGA	0bp	CAGGATATATGTGGTGTAAACAAATTG
NE439	CTATCAGTGTTTGA	0bp	CAACTTAATAAC
NE315	CTATCAGTGTTTGA	0bp	AAATTGACGCTTAGACACTTAATAACA
NE702	CTATCAGTGTTTG	0bp	TAAACAAATTGACGCTTAGACACTTAA
NE694	CTATCAGTG	0bp	GTGTAACAAATGACGCTTAGACAACT
NE484	CTATCAGTGTTTGA	0bp	CAGGATATATGTGGTGTAAACAAATTG
NG109	CTATCAGTGTTT	0bp	ATATTGTGGTGTAAACAAATTGATGCT
NG110	CTATCAGTGTTT	0bp	ATATTGTGGTGTAAACAAATGACGCTT
NG127	CTATCAGTGTTTGA	0bp	AAATTAATCGGGGGATCTGGATTTTAG
NG137	AGTGTTTGA	0bp	CAGGATATATGTGGTGTAAACAAATT
NE658	AGTGTTTGA	0bp	ATCGCCCGCAGAAGCGGGCCGTCTGG
CE034	CTATCAGTG	0bp	CACATGAGCGNGGGCCTATAGGAACCC

图 10 右边界与左边界的直接串连,没有 filler DNA。

Figure 10. Direct tandem between right border and left border without any filler-DNA at junctions

CE420	CTATCAGTGTTTGA GCGACGTCTTTTG ATAACACATTGCGGACGTTTTTA
CE302	CTATC CAATTGGTATCAGTGTTCCTCCGCTTCAATTG ATTGTGTGGTGTAAACAAATGACGC

图 11 右边界与左边界的串联,中间间隔有 filler-DNA。

Figure 11. Direct tandem between right border and left border with filler-DNA at junctions.

2.5.4 Filler-DNA 分析

对 T-DNA 插入子的 DNA 序列进行分析,发现 T-DNA 边界序列与水稻基因组 DNA 的接合处存在 filler-DNA 是一种很普遍的现象。在总共 182 个 T-DNA 插入子中,只有 69 个序列(38%)中没有 filler-DNA 的存在,而在其余 T-DNA 插入子中或多或少都存在序列长短不一的 filler-DNA, filler-DNA 的长度大于 100bp 的有 58 个,占总数的 32% (表 4)。

表 4 DNA 边界结合处的不同长度的 filler-DNA

Table 4. Filler-DNA in different length at T-DNA/rice DNA junction

	filler-DNA				
	0	1-20	21-50	51-100	>100
个数	69	36	16	3	58

对大于 50bp 的 filler-DNA 的序列作分析, 根据本地比较的结果, 发现很多载体上的 filler-DNA 序列的主要来源为 T-DNA 内部序列, 结果总结于表 5。主要来源的序列有如 CaMV35S、*Ds3'*、*Ds5'*、GUS、NOS、ADH1, 总共有 49 个占总数的 80%; 总共有 8 个序列是来自于 T-DNA 的外部序列, 其中抗卡那霉素基因 *aadA* 的序列发现有 2 个, 还有另个 6 个序列来源于二元载体主干序列的其它部位。另有 4 个序列来源不明。成为 filler-DNA 最多的是 GUS 基因, 总共有 20 个。从结果来看, GUS 基因似乎是一个重组发生的热点。

表 5 Filler-DNA 的来源

Table 5 Origin of filler DNA

	filler-DNA									
	CaMV35S	<i>Ds3'</i>	GUS	NOS	ADH1	<i>Ds5'</i>	bone	<i>aadA</i>	unknown	total
个数	5	6	20	5	6	7	6	2	4	61

根据 TAIL-PCR 的结果, 将 filler-DNA 的表现形式分为以下几类:

1):

T-DNA	短的 filler-DNA	水稻 DNA
-------	---------------	--------

如 NG092、NE795、CE078 等, 在这些 T-DNA 插入植株中, 其 T-DNA 与水稻 DNA 的接合处存在只有几个碱基的 filler-DNA, 因为几个碱基的序列太短, 无法正确判断其来源 (图 12)。由于 filler-DNA 很短, 所以得到的水稻序列比较长, 对于确定 T-DNA 的插入位点没有什么影响。

	T-DNA	filler-DNA	水稻DNA
NG092----	GTGTTATTAAGTTGTCTAAGCGTCAATTTGTTTACACCACAAT	8bp	AAACATCCCTCCATTACTCGCAACCCCAACT
NE795----	ATCCCCGAATTAATTCGGCGTTAATTCAGTACATTAAAAACG	6bp	CTGAATTCGAATTCGTTAGCATCATCTCTCC
CE078---	AAACGTCGCCAATGTGTTATTAAAGTTGTCTAAGCGTCAATTT	8bp	GGCATGGATTCTGCTCCAGAGATTGTGTTA

图 12 T-DNA 与水稻基因组 DNA 接合处的短 filler-DNA。

Figure 12. Short filler-DNA at T-DNA/rice genomic DNA junction.

2):

T-DNA	长的 filler-DNA	水稻 DNA
-------	---------------	--------

如图 13 中的 NE612 和 NE653, 它们的 filler-DNA 都比较长。NE612 的 T-DNA

右边界紧接着的是一段长为 77bp 的来源于 CaMV35S 启动子的 filler-DNA, 其外紧接着的是水稻基因组 DNA 序列, 与水稻第 11 条染色体 BAC 克隆 OSJNBb0039K08 的同源性为 95%。T-DNA 边界、CaMV35S 启动子 filler-DNA 和水稻基因组 DNA 三者接合处的序列都有明确的界限, 并且它们之间不存在任何其它的碱基。

其中 NE653 的 filler-DNA 更是高达 537bp。其 T-DNA 左边界紧接着的是一段长达 523bp 的来源于 T-DNA 内部序列的 CaMV35S 启动子和 ADH1 内含子序列, 之外则是一段 14bp 难以确定来源的 DNA 序列, 这段 DNA 与水稻或载体的序列都没有完全的同源性, 其外部才是水稻基因组 DNA, 通过比较发现与水稻第二条染色体的 BAC 克隆 OSJNBb0088N06 的一段序列具有完全的同源性。

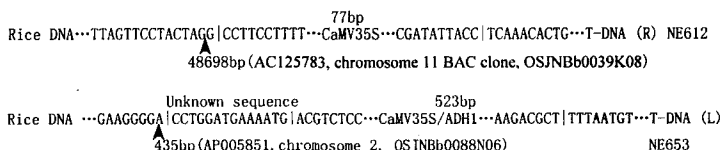


图 13 T-DNA 与水稻基因组 DNA 接合处的长的 filler-DNA

Figure 13. Long filler-DNA at T-DNA/rice genomic DNA junction.

3): T-DNA 长的 filler-DNA

图 14 中 CE315 其 T-DNA 右边界紧接着的 filler-DNA 为来源于 T-DNA 外的 265bp 载体主干序列(STA 区 DNA 序列), 由于 TAIL-PCR 扩增得到的序列长度比较短, 在其外没有扩增到水稻基因组 DNA。而 NE412 的 filler-DNA 的来源并不只是来源于某一段 DNA, 紧接着 T-DNA 边界的是一段长为 77bp 来源不明的 DNA 序列, 之外则是一段长度为 265bp 来源于 *Ds5'* 端的 DNA, 同样因为 TAIL-PCR 的局限性, 没有扩增到水稻 DNA 序列。

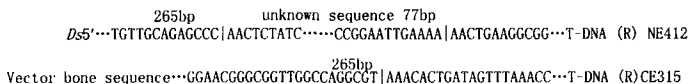


图 14 T-DNA 右边界处长的 filler-DNA

Figure 14. Long filler-DNA at T-DNA right border

3 讨论

3.1 T-DNA 插入拷贝数

在本研究中, 总共得到可育的独立转基因水稻植株 1188 株。根据 HYG^r 的抗性试验结果, T-DNA 在水稻基因组中的插入位点数为平均 1.3 个, 而通过 Southern 杂交的结果 T-DNA 在基因组中的插入拷贝数为 1.8 个。这个数值比 Jeon (2000) 报道的平均 1.4 个插入位点数和 2.2 个 T-DNA 插入拷贝数要小一些。这一数比采用基因枪法将 DNA 直接转入到植物细胞中(Kohli 等 1998)得到的拷贝数要少。插入位点数少, 插入的 T-DNA 拷贝数少对于插入位点旁侧的水稻基因组 DNA 的扩增会起到方便的作用, 而且在遗传分析方面也会大大的简化。

根据 TAIL-PCR 的结果, 大约 40% 的 T-DNA 插入到水稻的基因区, 因此可以推测在我们的研究中大约有 855 个 T-DNA 插入到了水稻基因区, 相当于可以获得 855 个插入突变体。这对于水稻功能基因的研究是非常有用的。根据我们的结果, 如果水稻基因以 40 000 个计, 则至少获得大约 56 000 个独立的转基因植株才有可能覆盖整个基因组。当然这种估计是远远偏小的, 因为随着 T-DNA 插入数量的增加, 插入到同一个基因中的 T-DNA 可能性就大大增加了, 尤其是对于那些大的基因。Krysan 等 (1999) 采用 $P=1-(1-X/125\ 000)^n$ 这一方程式来计算在拟南芥中找到一个大小为 Xkb 的基因概率为 P 条件下至少要 n 个 T-DNA 插入突变体。如果将这一方程式应用到水稻中, 即 $P=1-(1-X/430\ 000)^n$, 则如果要在水稻中找到一个基因大小为 3kb 的 T-DNA 插入突变体, 如果概率设定为 P=90%, 则大约需要 $n=33$ 万个 T-DNA 插入突变体。如果 P=99%, 则 $n=66$ 万, 相当于需要 51 万株独立的转基因水稻植株。

3.2 T-DNA 在水稻基因组的插入特点

T-DNA 在植物基因组中的整合位置通常发生在 AT 相对丰富的区域, AT 丰富区域可能是转基因发生重组的热点, 这些可能与染色体 DNA 在空间折叠结构有关 (Takano 等 1997)。T-DNA 整合到基因组 AT 丰富区在拟南芥(Mayerhofer 等 1991)、烟草(Gheysen 等 1991)中已有报道; 在水稻中 Takano 等(1997)在两个转基因水稻中序列也表明了 T-DNA 在转基因水稻中整合 AT 丰富的区域。在我们的研究中, 有 83% 的 T-DNA 插入子落在 AT 含量在 50-70% 的区域 (GC 含量在 30-50%), 这个区域也相当于水稻基因区或基因组 DNA 集中的区域。T-DNA 在水稻基因组中的整合偏向于水稻基因丰富区已有报道(Barakat 等 2000), 和 T-DNA 的这一特点对于水稻创建水稻突变体是相当有利的。表 3 中我们只是例出了其中一部分插入到水稻基因区中的 T-DNA 插入子, 从表中可以看出, T-DNA 插入分布在不同的染色体中, 在基因中的插入既

有在内含子的,也有在外显子的,还有一部分是插入在启动子区域的。从插入到酶的种类来看,也有各种各样的类型,如磷酸酯酶、蛋白激酶、脱水酶、异构酶、受体蛋白、锌指蛋白、氨基转移酶、磷酸转移酶等。这些插入在不同基因中的突变体,是进一步研究这些基因功能非常好的材料。因为已经知道了 T-DNA 在这些基因中的插入位置,就可以进行有目的的对可能产生的表型或生理变化安排试验,这也是所谓的反向遗传学的研究方法以及建立这个水稻插入突变体库的目的。对插入到基因区的插入子的分析正在进行之中。

3.3 T-DNA 在水稻基因组中的整合机理

T-DNA 在植物基因组中插入的具体机制相当复杂,至今没有一个清晰的认识。Tinland (1996) 认为由于致毒蛋白 VirD₂ 的保护,因此 T-DNA 右边界在整合到植物基因组的过程中,一直保持完整的 DNA 结构直到整合到植物基因组时,才在固定碱基处断裂,完成 T-DNA 在基因组 DNA 的精确插入。在对 158 例 T-DNA 右边界的研究中发现,将近有一半的序列保持了很完整的 TGA 标志序列,如果将保留 TG 和 T 的这两种情况加在一起则有大约 70% 的 T-DNA 右边界保持了精确的整合。这种整合方式与在拟南芥、烟草中的报道一致(Gheysn 等 1991, Mayerhofer 等 1991)。但是在其余的 46 例中 T-DNA 右边界都有不同程度的缺失,说明 T-DNA 右边界在基因组中的整合肯定还存在另外一种机制,这种整合方式是一种不精确的整合方式。

T-DNA 左边界在整合到水稻基因组 DNA 时,没有发现在某一个特定碱基处断裂的频率比较高,相对于右边界 T-DNA 的断裂点来讲,表现出很大的随机性。通过对 T-DNA 整合前的水稻基因组 DNA 与获得的 T-DNA 旁邻序列相比,在很多个 T-DNA 左边界序列与水稻基因组序列之间存在着小同源性序列(见图 9),这些同源性碱基在单链 T-DNA 与解链的基因组 DNA 进行互补配对,从而以单链修复模型完成 T-DNA 整合到基因组 DNA 中的过程(Mayerhofer 等 1991, Tinland 1996)。

图 10 和图 11 中表示了 T-DNA 的 RB-LB 直接串联分析。在串联的 T-DNA 中,绝大部分(14 个)在左边界与右边界之间不存在任何的 filler-DNA,而另外两个(CE302 和 CE420)串联中虽然有 filler-DNA,但经过比较其序列来自水稻基因组 DNA 的可能性不大。因此推测 RB-LB 的串联序列是在整合到水稻基因组 DNA 前就已经形成的。

在我们的研究中,没有扩增到 RB-RB 或者 LB-LB 这种倒转串联重复情况,而只出现了 RB-LB 的直接串联。如果在对 PCR 扩增原理做出分析以后,就不难理解为什么从 TAIL-PCR 的结果中无法得到 RB-RB 和 LB-LB 的串联,因为在这两种情况下,在 PCR 退火这一步中,由于 DNA 序列的完全倒转重复,引物可能就无法结合到 DNA 序列上,无法完成 PCR 的扩增。De Buck 等(1999)利用两个不同的 T-DNA 同时转化拟南芥,分析了 RB-RB 和 LB-LB 串联在基因组 DNA 中插入的情况。

对于在 T-DNA 左边界或者右边界出现的大量以 T-DNA 内部的一段 DNA 成为 filler-DNA 的现象。按照 Krizkova 和 Hroudá(1998)认为用 T-DNA 整合到植物基因组前形成 T-DNA 串联重复并且复制形成双链的模式解释不合理,因为在这种情况下,不太可能出现短的 T-DNA 序列作为 filler-DNA 存在。按照他们的模型,首先是 T-DNA 接近攻击染色体 DNA,使染色体 DNA 形成局部单链,T-DNA 的 3'端在染色体上找到小同源性序列配对形成双链,然后以此为引物修复植物基因组 DNA,T-DNA 的 5'端结合 VirD2 蛋白再与染色体 DNA 形成短暂的共价结合;然后第二条甚至第三条 T-DNA 链以同样的方式再攻击第一条 T-DNA 链,然后按照单链断裂修复模型整合到植物基因组 DNA 中。

3.4 Filler-DNA 对扩增插入位点水稻旁邻序列的影响

农杆菌介导的 T-DNA 转基因,通常只有 T-DNA 左右边界之间的序列才能转移到植株细胞中。但是 T-DNA 之外的序列有时也会随 T-DNA 一起转移到植物细胞中,在我们的研究中,也扩增得到了 7 个序列是来源于 T-DNA 之外的载体主干序列。这种现象也烟草和水稻中都有报道(方进等 2000, Maria 等 1997)。载体主干序列随 T-DNA 一起转入到水稻基因组中,对于突变体的分析增加的困难只是如其它 filler-DNA 片段一样,没有什么特别之处,但是在生物安全评价方面被认为是一个很大的隐患,在此不作展开性讨论。

Filler-DNA 在 T-DNA 边界中的大量出现,以及 T-DNA 串联重复序列的形成给 T-DNA 插入位点处水稻序列的获得造成了一定的困难。由于 TAIL-PCR 扩增产物的大小是很有限的,往往在几百个碱基左右,如果 filler-DNA 的长度也是几百个碱基,则就难扩增到水稻序列。在我们的研究中,在 182 个 T-DNA 边界序列中,存在有 filler-DNA 的序列有 113 个,占总数的 62%,虽然有些 filler-DNA 的片段很小,不会影响到 T-DNA 旁侧序列的扩增,但是超过 100bp 的 filler-DNA 达到了 58 个,占总数的 32%,这其中很大一部分会直接影响到 T-DNA 旁邻序列的获得。即使 TAIL-PCR 扩增的产物有足够长,应用 PCR 产物测序,按照目前一次测序反应能读到的碱基数平均也就在 500-600bp,而且 TAIL-PCR 的特点是只有一侧是特异性引物,只能采用用特异引物的一端测序,这样如果 filler-DNA 的长度比较长的话,也很难测到基因组序列。如果采用将 PCR 产物克隆到载体中再进行测序,那过程与时间就增加了很多,对于大量的 T-DNA 旁邻序列克隆与测序显然工作量是很大的。

第三章 *Ac/Ds* 转座子在水稻中跳跃规律的研究

1 材料与方法

1.1 材料

二元载体系统 *Ac*-T-DNA 和 *Ds*-T-DNA, 载体构建见上一章。结构示意图见图 1。转基因起始植株见上一章。

所用的工具酶、载体、酶也如上一章所述。

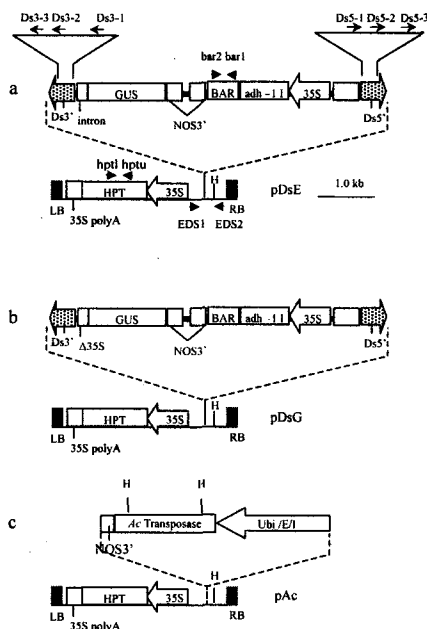


图 1 pDsE、pDsG、pAc 的结构示意图。

Figure 1. Schematic representation of pDsE, pDsG, pAc.

Ds3-1, *Ds3-2*, *Ds3-3* and *Ds5-1*, *Ds5-2*, *Ds5-3* are the specific primer of TAIL-PCR for cloning genomic sequence flanking *Ds* element. EDS1 and EDS2 are primer for amplification of empty donor sited left by excision of *Ds* element. All the other elements are described at the second chapter.

1.2 方法

1.2.1 F_2 群体的发展

含有单拷贝 *Ac* 的转基因水稻作父本, 与含单拷贝 *Ds* 作母本的转基因水稻杂交得 F_1 , F_1 自交得 F_2 。 *Ac* 元件已经缺失了 5' 和 3' 末端, 因此已经不会发生转座。不可发生转座的 *Ac* 元件带有用 UBI 启动的 *Ac* 转座酶基因, 并以此来激发 *Ds* 的割离。 *Ds* 元件含有报告基因 GUS 和选择标记基因 BAR。 GUS 报告基因的表达可以通过组织化学染色来进行检测, 增强子捕获器 *DsE* 元件在 GUS 报告基因前加了一个 CaMV35S 的弱启动子 (含有 TATA 盒), 当 *DsE* 插入到水稻基因组 DNA 的顺式作用元件中时 GUS 基因就会表达。 基因捕获器是在 GUS 报告基因前加了内含子剪接的受体位点和供体位点但是不加启动子成份, 当 *DsG* 插入到水稻基因组转录区时, 通过内含子的剪接作用使 GUS 与水稻基因形成融合基因而得到表达 (Sundaresan 等 1995)。 当 *Ds* 元件从原始供体位点割离后, HPT 基因将保留在空供体位点 (示意图 1)。

而对于含有 *Ds* 元件的苗, 可以通过喷洒除草剂对 BAR 基因进行选择。 我们采用将发芽一个星期后的水稻苗, 移栽到溶液培养中, 正常生长一个星期后, 喷洒 3mg/L PPT (Basta), 3-7 天后不含有 BAR 的植株将枯黄直至死亡。

1.2.2 小量 DNA 的提取

取 20 天苗龄的水稻叶片 5cm, 剪成小片段, 放入 1.5ml 离心管中, 加液氮磨碎, 加入 500 μ l DNA 提取液(100mM Tris-HCl pH8.0, 20mM EDTA, 500mM NaCl, 1.5% SDS), 短暂振荡使提取液与粉末混匀, 65 $^{\circ}$ C 水浴放置 20min, 加等体积的氯仿/乙醇/戊醇(20:4:1)振荡混匀 10 秒种, 12,000rpm 离心 5min, 取 500 μ l 上清液转移到新的 1.5ml 离心管中, 加 1ml -20 $^{\circ}$ C 的无水乙醇, 混匀, 室温下置 5min, 12,000rpm 离心 5min, 去上清, 加 1ml 70%乙醇洗涤, 倒扣置于吸水纸上自然晾干, 加入 100 μ l 含有 0.1mg/ml RNaseA 的 TE(pH8.0)溶解, 作为 PCR 反应的模板。

1.2.3 *Ds* 跳跃的 PCR 分析

检测 *Ds* 在 F_2 代中的跳跃。 根据载体的构建, 设计引物 EDS1 (5'-atgctccggctcgtatgtgtgt-3')和 EDS2 (5'-aagcgtaagggtatattggaagta-3')扩增 *Ds* 跳跃后的空供体位点 EDS (empty donor site), 引物 EDS1 和 *Ds*3-1 (5'-ggttcccggtccgatttgact-3')扩增完整的供体位点 FDS (full donor site)。 为减少 PCR 反应数量, 将三个引物同时置于一个反应管, 反应采用 30 μ l 体积(50ng DNA, 200 μ M dNTP, 1 μ M EDS1, 0.5 μ M EDS2 和 *Ds*3-1, 1.5mM $MgCl_2$, 50mM KCl, 0.1% Triton X-100, 10mM Tris-HCl, pH9.0, 1 unit

Taq DNA polymerase); 热循环条件为: 94℃ 4min; 94℃ 20sec, 56℃ 30sec, 72℃ 1min, 35 cycles; 72℃ 5min。根据设计, EDS 的扩增片段为 516bp, FDS 为 287bp。

检测 HPT 基因的引物设计为 hptu (5'-cggtcgcggaggctatggatg-3') 和 hptl (5'-gcttctcggggcgattgtgta-3'), 扩增产物为 599bp。反应采用 20μl 体积(50ng DNA, 100μM dNTP, 0.5μM hptu 和 hptl, 1.5mM MCl₂, 50mM KCl, 0.1% Triton X-100, 10mM Tris-HCl, PH9.0, 1 unit Taq DNA polymerase); 热循环条件为: 94℃ 4min; 94℃ 20sec, 62℃ 30sec, 72℃ 1min, 35 cycles; 72℃ 5min。

检测 BAR 基因的引物设计为 bar1 (5'-atgagcccagaacgacgccc-3') 和 bar2 (5'-tcagatctcgggtgacgggca-3'), 扩增产物大小为 354bp 反应采用 20μl 体积(50ng DNA, 100μM dNTP, 0.5μM baru 和 bar1, 1.5mM MCl₂, 50mM KCl, 0.1% Triton X-100, 10mM Tris-HCl, PH9.0, 1 unit Taq DNA polymerase); 热循环条件为: 94℃ 4min; 94℃ 20sec, 60℃ 30sec, 72℃ 1min, 35 cycles; 72℃ 5min。

1.2.4 Southern Blot 分析

水稻总 DNA 的大量提取法卢扬江(1992)的方法, 5μg 总 DNA 用 *Hind* III 酶切完全, 在 0.7% 的琼脂糖凝胶中电泳分离, 转移到带正电的尼龙膜(Hybond-N+, Amersham)上, 用 ³²P 放射性同位素标记的 HPT 基因(与亲本杂交)或 BAR 基因(与 F₂ 植株杂交)作探针。*Hind* III 在载体上的分布如图一所示, 在 HPT 基因和 *Ds* 上没有酶切位点, 杂交的条带数反映 T-DNA 或 *Ds* 的拷贝数。

1.2.5 *Ds* 旁邻序列的克隆

用 TAIL-PCR 的方法(Liu *et al.*, 1995a; Liu *et al.*, 1995b)扩增 *Ds* 插入位点两侧的水稻基因 DNA。*Ds* 3' 端的特异引物采用 *Ds*3-1: 5'-gggtcccggtccgattcgact-3', *Ds*3-2: 5'-cgattaccgtatttatcccgttc-3', *Ds*3-3: 5'-tcgtttccggtcccgcaagt-3'。*Ds* 5' 端的特异引物采用 *Ds*5-1: 5'-acggtcgggaaactagctctac-3', *Ds*5-2: 5'-tccgttccggtttcgttttttac-3', *Ds*5-3: 5'-cggtcgggtacgggattttcc-3'。随机引物用 AD1, NTCGA(G/C)T(A/T)T(G/C)G(A/T)GTT; AD2, (G/C)TTGNTA(G/C)TNCTNTGC; AD3, (A/T)GTGNAG(A/T)ANCANAGA。反应条件参照 Liu 等(1995)的方法进行。特异扩增得到的第三轮 PCR 产物经凝胶电泳回收后克隆到 pUC-T 载体中, 用 M13 测序引物及发玛西亚测序试剂盒(DYEnamic ET dye terminator kit)作反应在 MEGABASE1000 测序仪上测序。序列在 GenBank 中作 BLASTN 搜索确定 *Ds* 在水稻基因组中的插入位置, 序列本地分析在 DNASTAR (version 5.0)中进行。

1.2.6 *Ds* 旁邻序列的分析

序列在 GenBank(<http://www.ncbi.nih.nlm.gov/BLAST>)中作 BLASTN 搜索确定 *Ds* 在水稻基因组中的插入位置, 数据库 GenBank 中所有的水稻序列 nr (所有 GenBank+RefSeq Nucleotides+EMBL+DDBJ+PDB 序列, 不包含 EST、STS、GSS 和阶段 0、1、2 的 HTGS 序列)、gss (genome survey sequence)和 htgs (high throughput genomic sequence)。 *Ds* 插入元件在水稻物理图谱上的插入通过 BAC 号或 PAC 号在 <http://www.tigr.org/tdb/e2k1/osa1/BACmapping/description.shtml> 比较后确定。本地分析在 DNASTAR(version 5.0)中进行。核酸序列的同源分析按相似性>95% 或者 score>300 为准,对于有些片段比较短的序列, 按人工判断是否为水稻序列。

1.2.7 Gus 分析

将 1-2cm 的叶片、根或花放入装有 GUS 染色液 (GUS 染色液的配方: 100mM 磷酸缓冲液 pH7.0, Triton X-100 0.5%, 甲醇 20%, X-Gluc 0.5mg/ml) 的 1.5ml 离心管中, 真空抽气 15min 后盖上盖子, 放入 37℃ 恒温箱中反应 24hr。在 LEICA MZ95 体视镜下观察, 用 LEICA DC300 摄像头拍照。样品用戊二苯醛固定, 用 ERL-4206 树脂包包埋, 切片厚度为 4μm。

2 结果与分析

2.1 F_2 群体发展

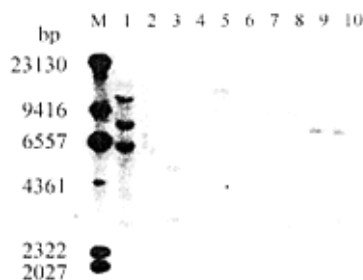


图2 亲本植株的Southern杂交分析。

Figure 2. Southern hybridization analysis of parent plants. M: λ /Hind III marker; 1:ce14; 2:ce37, 3:ce59, 4:ce70; 5:ce72; 6:ce78; 7:ce80; 8:cg38; 9:cac46, 10:cac48.

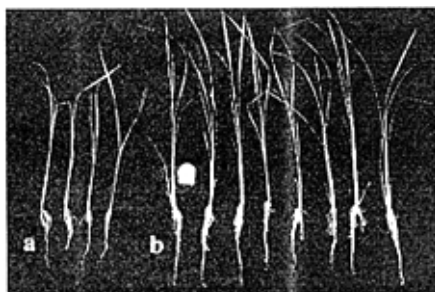


图3 PPT 抗性的 F_2 植株选择。

Figure 3. Screening PPT^r plants in F_2 population.

a: PPT^r rice plants.
b: PPT^s rice plants.

在本研究中, 我们采用二元转座子系统来发展 *Ds* 插入群体。将带有不可移动单拷贝 *Ac* 元件的水稻亲本植株作为父本与带有拷贝数不等的 *Ds* 元件(*DsE* 或 *DsG*)的水稻植株作母本进行杂交。在杂交前先对亲本植株 Southern 杂交进行拷贝数(图 2)和位点数(见上一章)的确定。*Ac* 全部为单拷贝单位点转基因植株, 而 *Ds* 作为亲本的 *Ds* 转基因水稻除了 CE14 外都表现出 3:1 (PPT^+ : PPT^-) 的分离, 说明它们的位点数都是 1 个, 但是 CE70, CE72, CE78 和 CE80 Southern 杂交的条带数则不至 1 条。杂交得到的 F_1 , F_1 自交得 F_2 , 分别进行单株收取种子以进行分析之用。

F_2 植株通过 BASTA 喷洒选择, 不含有 *Ds* 元件的 F_2 植株将枯萎死亡(图 3), 这样在对 F_2 植株的 PCR 筛选时, 死亡部分的植株就可以不进行 PCR 的检测。从而达到降低成本、减少工作量。

2.2 F_2 植株 *Ds* 元件从供体位点的割离

杂交之后, 总共得到的 108 个 F_1 植株, 在开花前提取水稻叶片的基因组 DNA 用作 PCR 的模板, 采用设计的三引物 EDS1、EDS2 和 *Ds3*-1 为引物来扩增分析确定 *Ds* 在 F_1 植株中的跳跃情况, 结果在所有 108 个样品中都没有扩增到 *Ds* 跳跃后留下的空位点 EDS。

同样采用 PCR 分析方法对 PPT^+ 的 F_2 植株 *Ds* 跳跃情况进行了分析, 结果表明 *Ds* 在 F_2 植株的跳跃频率大大增加了。按照设计在 F_2 植株的 *Ds* 跳跃检测中, 其三引物扩增的 PCR 产物应当出现四种条带模式, 分别为 EDS^+FDS^- 、 EDS^+FDS^+ 、 EDS^-FDS^+ 和 EDS^-FDS^- , 实验结果与预期的相一致, 在 F_2 代 PCR 检测中出现这四条带模式(见图 4)。 EDS^-FDS^- 条带模式(C19-01.002, C19-01.005 和 C19-01.010)代表只含有重新插入的 *Ds* 元件的(提供 PPT^+), 而不含有原来的供体位点; EDS^+FDS^- (C19-01.001, C19-01.009 和 C19-01.011)表明 F_2 植株含有一个空位点及提供 PPT 抗性的 *Ds* 位点; EDS^-FDS^+ 表明 F_2 植株只含有原先的位点而不含有 *Ds* 跳跃后的空位点 EDS; EDS^+FDS^+ (C19-01.004)表明 F_2 植株既含有跳跃后的空位点, 又含有原来的供体位点, 这种情况是由后代自由重组引起的。

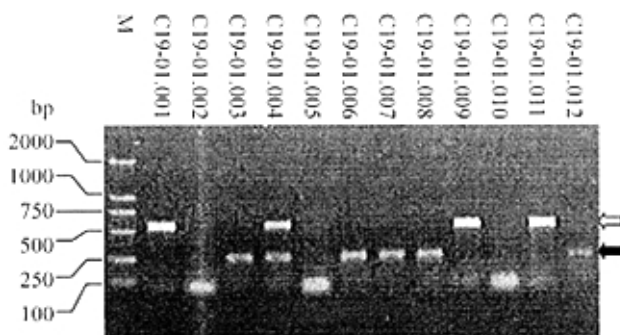


图 4 *Ds* 转座发生后留下的空供体位点 EDS 的 PCR 检测。

Figure 4. PCR amplification analysis of EDS left by excision of *Ds* element.

为证明用三引物扩增得到的 516bpEDS 片段确实代表 *Ds* 从供体位点跳跃后的情况, 我们对 3 个亲本的 10 个不同 EDS 片段进行了测序, 测序的结果表明 *Ds* 元件确实是从原来的供体位点上跳走了。*Ds* 跳跃后在供体位点发生了少量几个碱基的变化, 碱基的变化既有缺失如 C03-14.024 等, 也有替代如 C03-14.267 和 C03-14.272 等(见图 5)。*Ds* 跳跃后在空位点处发生碱基变化在玉米、拟南芥、酵母和水稻中都有发现(Izawa 等 1997, Rinehart 等 1997, Suzuki 等 2001, Weil and Kunze 2000)。

Ds element
▽

GGAATTCGAGCTC	GGTCACGCAACGC	donor site
GGAATTCGAGCTC	CGTCACGCAACGC	C03-14.272, C03-14.047, C03-14.003, C11-02.003
GGAATTCGAGCTG	AGTCACGCAACGC	C03-14.267, C15-43.153, C15-43.186
GGAATTCGAGCT-	-GTCACGCAACGC	C03-14.024, C03-14.058, C15-43.182, C19-01.001

图 5 EDS 位点的 DNA 序列碱基变化。

Figure 5. DNA sequence variety at EDS stie.

为了进一步确定用 PCR 方法扩增的 EDS 和 FDS 的正确性, 我们又采用了 HPT 基因作探针进行了 Southren 杂交分析(图 6)。与 PCR 分析的结果相一致, F_2 植株 C15-43.186、182 和 062 用 PCR 的方法只扩增出 EDS 片段, 在 Southern 杂交中也只出现了大约 2800bp 的杂交片段, 而 F_2 植株 C15-43.172 在扩增中既有 EDS 又有 FDS 片段, 在 Southern 杂交中则出现 8800bp 和 2800bp 大小的 2 个片段, 这两个片段分别与 FDS 和 EDS 相对应。

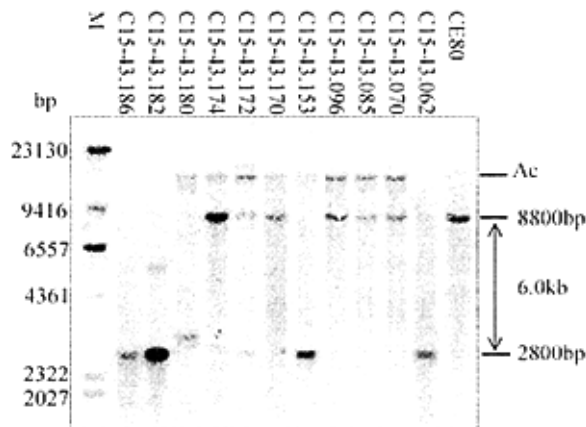


图 6 EDS 位点的 Southern 杂交分析。

Figure 6. Southern hybridization analysis of EDS site.

Ac: *Ac* transposase fragment, the 8800bp and 2800bp fragment are FDS and EDS respectively.

The band length is calculated with software Gel-Pro ANALYZER (version 4.0.00.001).

我们对来自 11 个 *Ds* 亲本总共 20 个 F_2 群体用 PCR 方法作了 *Ds* 割离(excision)情况包括体细胞割离(somatic excision)和生殖细胞割离(germinal excision)的分析。PCR 分析的结果统计列于表 1。*Ds* 在 F_2 不同群体的割离频率范围为 0-40%，来自同一个组合的不同 F_2 群体其 *Ds* 割离的频率变化不大。例如来自杂交组合 CE70×CAc46 的三个 F_2 群体它们的跳跃频率分别为 27%, 30% and 33%。杂交组合 CE37×CAc46 and CE76×CAc46 它们同样具有与 CE70×CAc46 的父本 CAc46，但显然它们的 *Ds* 割离频率要比 CE70 的相差很远，分别只有 0%与 6%。这说明 *Ds* 元件在水稻染色体中插入位点处的不同基因组 DNA 序列影响到 *Ds* 的割离频率。这种现象在拟南芥中已有报道(Smith 等 1996)，而亲本的 *Ds* 元件拷贝数则对 *Ds* 的跳跃频率影响不大，如 *Ds* 亲本 CE70 和 CE68 均为多拷贝，而 CE68 为单拷贝，但是它们的 *Ds* 割离频率都在 30% 以上，相差不大(表 1)。

根据 Southern 杂交的结果，转座事件有比较高的频率是发生在生殖细胞中。如果发生在生殖细胞的割离，其后代则通过 Southern 杂交的条带只留下 EDS 空位点条带而不带有 FDS 条带。图 6 显示在 6 个杂交的条带中，只有 C15-43.172 显示出既有 FDS 条带也有 EDS 条带，而这意味着 6 个发生割离的至少有 5 个是发生在生殖细胞中的。

2.3 *Ds* 元件在 F_2 代中的插入

为确定从供体位点割离后的 *Ds* 在 F_2 植株中重新插入的情况，用放射性同位素 ^{32}P 标记的 BAR 作探针对 8 个 F_2 群体的基因组 DNA 作了 Southern 杂交分析，图 7 显示出同一群体中不同个体的 F_2 植株表现出不一样的杂交带型，说明 *Ds* 在基因组 DNA 中的插入位置是不同的。根据 Southern 杂交结果，*Ds* 在 F_2 群体中的独立插入频率在 70%以上(表 2)。这一结果与在拟南芥中的报道(Sundaresan 等 1995)不一样，在他们的报道中，绝大部分的 *Ds* 插入子表现出相同的杂交带。

从 FDS 割离后的 *Ds* 在 F_2 植株中并不一定重新插入植物基因组 DNA 中的。根据 8 个 F_2 群体的 Southern 杂交结果统计分析，在我们的系统中 *Ds* 重新插入频率保持在高于 70%的水平(表 2)。大多数重新插入的 *Ds* 植株含有 *Ac* 元件，*Ac* 元件的存在可能会导致 *Ds* 在下一代中的重新跳跃。稳定插入(有 *Ds* 插入但是没有 *Ac* 元件的 F_2 植株)频率变化范围在 14-33%之间，这个频率范围符合单基因分离的孟德尔规律。

表 1. *Ds* 元件在 F_2 群体中的剥离与重新插入的频率

Table 1. Frequency of *Ds* excision and reinsertion in F_2 plants derived from individual F_1 plants

F_1 plant	<i>Ds</i> x <i>Ac</i>	PPT ^r F_2 plant	EDS				<i>Ac</i> - excision (%)	Reinsertion (%)	Independent reinsertin (%)	Stable reinsertion
			F_2 plants showing	Reinsertion	Independent					
C03-02	CE70× <i>CAc</i> 46	60	16	12	10	4	27	75	83	33
C03-13	CE70× <i>CAc</i> 46	60	18	13	10	4	30	72	77	31
C03-14	CE70× <i>CAc</i> 46	100	34	28	21	7	33	82	75	25
C05-05	CE76× <i>CAc</i> 46	35	2	-	-	-	6	-	-	-
C05-06	CE76× <i>CAc</i> 46	35	0	-	-	-	0	-	0	-
C06-14	CE37× <i>CAc</i> 46	35	2	-	-	-	6	-	-	-
C06-22	CE37× <i>CAc</i> 46	35	0	-	-	-	0	-	-	-
C11-02	CE14× <i>CAc</i> 46	35	5	-	-	-	14	-	-	-
C11-09	CE14× <i>CAc</i> 46	35	6	-	-	-	19	-	-	-
C19-01	CE68× <i>CAc</i> 46	39	13	10	8	2	33	77	80	20
C13-04	CE78× <i>CAc</i> 48	35	10	-	-	-	29	-	-	-
C13-09	CE78× <i>CAc</i> 48	35	9	-	-	-	25	-	-	-
C15-01	CE80× <i>CAc</i> 48	35	10	7	6	1	29	70	86	14
C15-43	CE80× <i>CAc</i> 48	100	41	30	22	8	41	73	73	27
C14-03	CE42× <i>CAc</i> 52	35	8	-	-	-	23	-	-	-
C18-05	CE72× <i>CAc</i> 52	35	0	-	-	-	0	-	-	-
C18-06	CE72× <i>CAc</i> 52	35	1	-	-	-	3	-	-	-
C20-03	CG38× <i>CAc</i> 52	84	18	15	12	4	22	83	80	27
C20-05	CG38× <i>CAc</i> 52	60	14	12	10	3	23	86	83	25
C26-01	CG32× <i>CAc</i> 24	39	9	-	-	-	23	-	-	-

Ac-: reinsertion plants lacking *Ac* element;
-: not determined;
Excision (%): No. of F_2 plants showing EDS/ No. of F_2 PPT^r plants analysed;
Reinsertion (%): No. of F_2 plants showing reinsertion/ No. of F_2 plants showing EDS;
Independent reinsertion (%): No. of F_2 plants showing independent reinsertion/No. of F_2 plants showing reinsertion;
Stable reinsertion (%): No. of reinsertion plants showing *Ac*-/ No. of F_2 plants showing reinsertion.

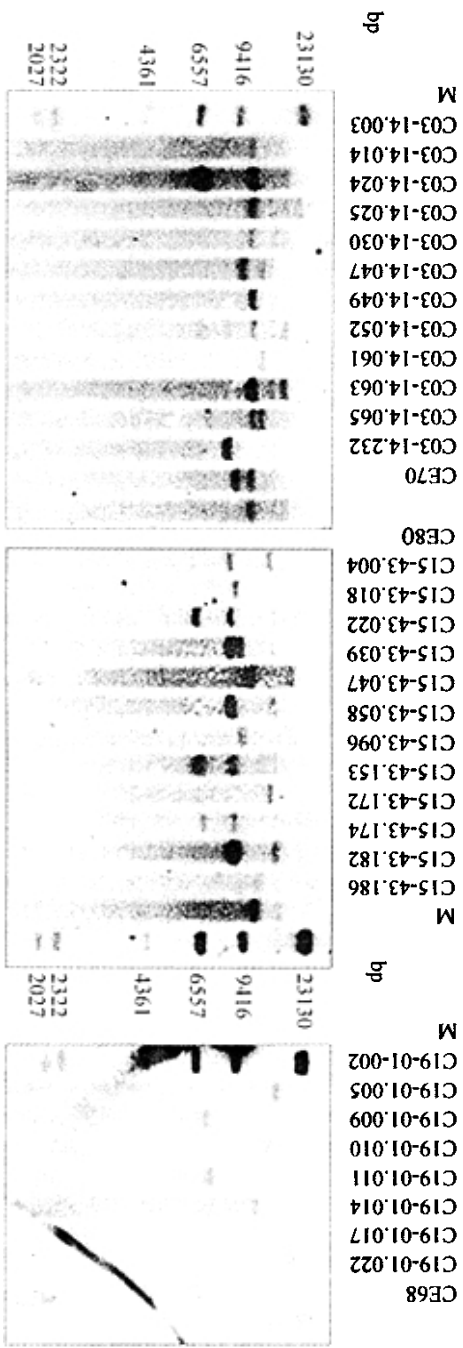


图 7 F₂ 植株的 Southern 杂交分析。

Figure 7. Southern hybridization analysis of F₂ plants.

CE70, CE80 and CE68 are D_s parent of C03-14, C15-43, C19-01 respectively. Genomic DNA was cut with *Hind* III. Probed with ³²P labeled BAR gene fragment.

2.4 *Ds* 元件在水稻基因中的插入位点分析

为确定转座后的 *Ds* 元件是否与供体点位相连锁, 我们采用了 TAIL-PCR 的方法克隆了 *Ds* 元件旁侧的水稻 DNA 序列, 序列扩增的特异性由第二轮与第三轮 PCR 产物的大小来确定, 通过电泳第三轮产物比第二轮产物要小一些来判断 TAIL-PCR 扩增的特异性 (图 8)。

在设计上, 第三轮特异扩增的 TAIL-PCR 保留了 *Ds* 序列 3' 末端的 39bp, *Ds* 5' 末端的 82bp, 克隆到 pUC-T 载体上测序列能保证这些标记序列的完整保留。通过对这些标记序列的判读, 确定得到的 29 个 *Ds* 元件旁邻的水稻 DNA 序列是特异扩增的。

这些序列在水稻中位置的确定通过与 GenBank 中水稻序列作 BLASTN 比较, 以同源性大于 95% 或 Score 值大于 300 为准。序列比较结果表明来自四个供体位点的 29 个序列分布于 7 条不同的染色体上 (表 2)。旁邻序列的结果表明在 Southern 杂交中的不同的条带对应于 *Ds* 在染色体上不同的位点。与供体位点相比表明, *Ds* 的插入位点与原始位点不相串联。

在 29 个 *Ds* 插入位点中, 有 17 个 (59%) 插入到了基因区, 其中有 6 个在启动子区域 (定义为开放阅读框前 1-1000bp), 7 个在外显子区域, 有 4 个在内含子区域。另外还有 10 个序列位于非基因区域 (定义为开放阅读框前 1000bp 以上和开放阅读框 200bp 以后)。 *Ds* 可以插入到不同的基因中, 如谷胱甘肽转移酶 (C03-14.011)、离子通道蛋白 (C20-5.090)、查尔酮合酶 (C20-3.169)、ECA1 相关蛋白 (C15-43.153) 等。 *Ds* 在水稻中以 59% 的高比例插入到水稻基因区表明 *Ds* 转座子适合于作为水稻基因的插入序列标签。

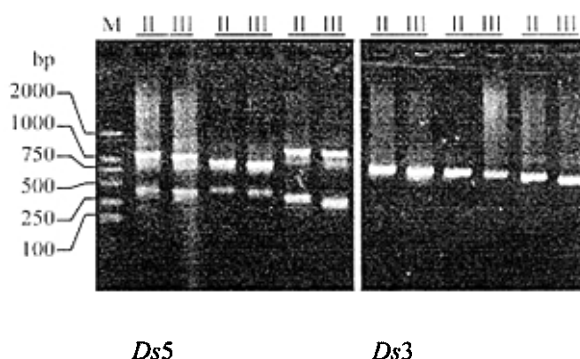


图 8 TAIL-PCR 扩增 *Ds* 旁侧序列的电泳图。

Figure 8. Electrophoresis of TAIL-PCR fragments amplification from *Ds* insertion site.

II represented the second round TAIL-PCR products, III represented the third round TAIL-PCR products., *Ds*5 indicated 5' end of *Ds* element and *Ds*3 3' end.

表 2 Ds 在水稻基因组中的插入位置。

Table 2 Ds reinserion positions in rice genome.

Ds lines	genomic Clone	Accession number position	insertion position of the clone	Insertion within Chr (cm)	Chr	relationship of insertion site toORF	Annotation insertion position
CE70 ^a	OJ1113_A10	AP004643	29302	92.2	8	upstream	Glucose-6-phosphate isomerase
C03-14.061	P0017H11	AP005859	11162	118.1	2	upstream	Putative protein
C03-14.052	OSJNBa0030C11	AC137930	2930	10.9	3	intergenic	
C03-14.030	OSJNBb0027B12	AC137075	67383	44.4	3	exon	leucoanthocyanidin dioxygenase-like protein
C03-14.070	OSJNBa0006D11	AC134231	136078	59	3	intergenic	
C03-14.024	OSJNBa0087G11	AC091787	49183	73.5	3	upstream	retroelement
C03-14.037	OSJNBb0081B07	AC093018	128626	154.8	3	upstream	OSEXp7
C03-14.049	OSJNBa0034E08	AC135597	145581	-	3	upstream	EST S055H03
C03-14.011	P0425F05	AP003569	58565	35.8	6	upstream	Glutathione transferase
C03-14.232	OSJNBa0090K04	AP003216	94113	-	6	exon	Predicted gene
C03-14.267	OSJNBa0064I23	AP005912	116741	20.7	9	exon	LTR retrotransposon Osr39-1
CE78-7 ^b	P0497A05	AP003380	93300	157.6	1	intron	Unknown protein
C15-01.077	OSJNBa0035A24	AP005514	119600	54.9	2	intergenic	
C15-43.186	OJ1148_D05	AP004118	65800	118.1	2	upstream	RING-H2 zinc finger protein-like
C15-43.153	OSJNBb0036D03	AC118670	5305	46.6	3	upstream	ECA1 protein-related
C15-43.047	OSJNBb0078C05	AC137634	31000	46.6	3	intergenic	
C15-43.166	OSJNBb0011H13	AC120983	12754	130.7	3	intergenic	
C15-43.004	P0642B07	AP003623	110004	82.9	6	exon	reverse transcriptase
CE68-3 ^c	OSJNBb0077A02	BX000503	15355	7.4	12	intergenic	
C19-01.017	OSJNBa0005N02	AL662954	41698	-	4	intergenic	
C19-01.015	OSJNBb0005G07	AP005149	72787	83.3	7	intron	Hypothetical protein
C19-01.002	OJ1339_F05	AP004009	23394	95.8	7	intergenic	
C19-01.005	OJ1136_D11	AP003749	107794	99.3	7	intron	Hypothetical protein

Table 2. continued

CG38 ^d	OSJNBa0032E21	AF377947	16584	139.8	3	exon	Unknown protein
C20-3.016	OSJNBb0011H13	AC120983	12474	130.7	3	intergenic	ligand-gated ion channel protein
C20-5.090	OSJNBa0090D06	AP004738	119180	18	6	exon	
C20-3.168	P0493C06	AP005193	107389	115.6	7	intergenic	
C20-3.003	P0465E03	AP005707	69713	49.3	9	intergenic	
C20-3.035	OSJNBa0015J03	AC131375	123989	10.9	10	intergenic	aldehyde 5-hydroxylase
C20-5.141	OSJNBa0055O03	AC073867	67178	57.5	10	exon	
C20-3.169	OSJNBa0026L12	AC068924	85693	57.5-58.6	10	intergenic	chalcone synthase
C20-3.027	OSJNBb0038A07	AC113948	55739	-	10	exon	glutathione S-transferase
C20-5.209	OSJNBa0026L12	AC068924	111561	57.5-58.6	10	intron	NOD26-like membrane integral protein

a, b, c, d: *Ds* donor site

-: insertion position (cm) in rice chromosome was not determin

为了确定 *Ds* 在水稻中是否存在重新插入的时候是否形成 8bp 的重复序列, 对 4 个 *Ds* 插入位点的两侧的水稻序列分别进行了克隆与序列分析, 结果表明在插入位点 *Ds* 的两侧形成了水稻基因组 DNA 的 8bp 碱基的重复, 结果见我们的结果与 Izawa(1997)的报道一致。

```
GCATCTGAAGGACTGAAG [Ds] GACTGAAGCACTTCGTTT C19-01.002
ACAATTGGTCTGAGCAGT [Ds] TGAGCAGTTGGGGTTCCC C19-01.005
AAGCAGCCAGATCAATAC [Ds] ATCAATACACTATACACT C15-43.166
CTCACTTTGGGTGTTCTT [Ds] GTGTTCTTAGTTCACGTC C20-05.209
```

图 9 *Ds* 元件在插入位点处 8bp 的正向重复序列。

Figure 9. Direct 8bp duplication at *Ds* insertion sites. Bold letters indicated duplication.

2.5 GUS 染色分析

当基因捕获器或增强子捕获器插入到水稻基因组时, 就有可能使报告基因 GUS 得到表达。为了估计这一系统中基因捕获器或增强子捕获器的效率, 我们将叶、根、花或不成熟种子作 GUS 活性分析。我们总共分析了来自 6 个 F_1 含有不同 *DsE* 插入的 100 个 F_2 植株, 总共有 28 个株植株表现出 GUS 活性, 增强子捕获器在这个系统中的 GUS 检出率为 28%; 对自 2 个 F_1 家系(C20-03 and C20-05)的 27 个 *DsG* 独立插入植株进行 GUS 染色, 发现其中有 6 株水稻表现 GUS 染色, GUS 检出率为 22%。部分 *DsG* 和 *DsE* 插入植株的 GUS 染色的结果如图 10 所示。GUS 表达的部分既有在根中的(如 C20-05.027, C20-05.090)、叶中(如 C20-5.194, C13-4.002), 也有在花中(如 C20-5.311)、种子中(C03-14.272)。结果说明捕获器可以捕获不同时期特异表达的基因。

2.6 C03-14.037 植株的分析

C03-14.037 是一个 *DsE* 插入子, GUS 染色结果在根尖(包括主根、不定根和侧根根尖)和花药处有强烈的 GUS 活性, 而在叶片和种子则不表现出任何的 GUS 活性。根据 GUS 染色和切片结果, 主要是在根尖分生细胞和花药绒毡经外的细胞有强烈的 GUS 活性。其根尖 GUS 基因始终有强烈的表达。TAIL-PCR 结果表明, *DsE* 元件插入到水稻第 3 条染色体分子标记 R273 附近的膨胀素基因 *OsEXP7* 启动子区域, 位于起始密码子 ATG 前 366bp。C03-14.037 F_2 植株在初期生长阶段并无异常, 但是在田间生长过程中发现其后期比其它植株要矮小, 并且结实率很底。在对 F_3 植株的观察当中, 在生长到 1 个月左右时生长与野生型水

稻没有差别, 并且不表现出性状的分离。但是在 GUS 分析过程中发现, 在总共 10 株苗中其中有 2 株是没有表现出 GUS 活性的, 说明 *Ds* 插入在后代中是分离的。由于 *Ds* 插入引起的突变性状可能产生在植株生长发育的后期阶段, 需要进行进一步的观察与确定。膨胀素基因是一类对细胞扩大生长起着很重要的作用, 因此这一突变体对于研究膨胀素在细胞生长扩大中的具体作用机制可能会起到很大的帮助。

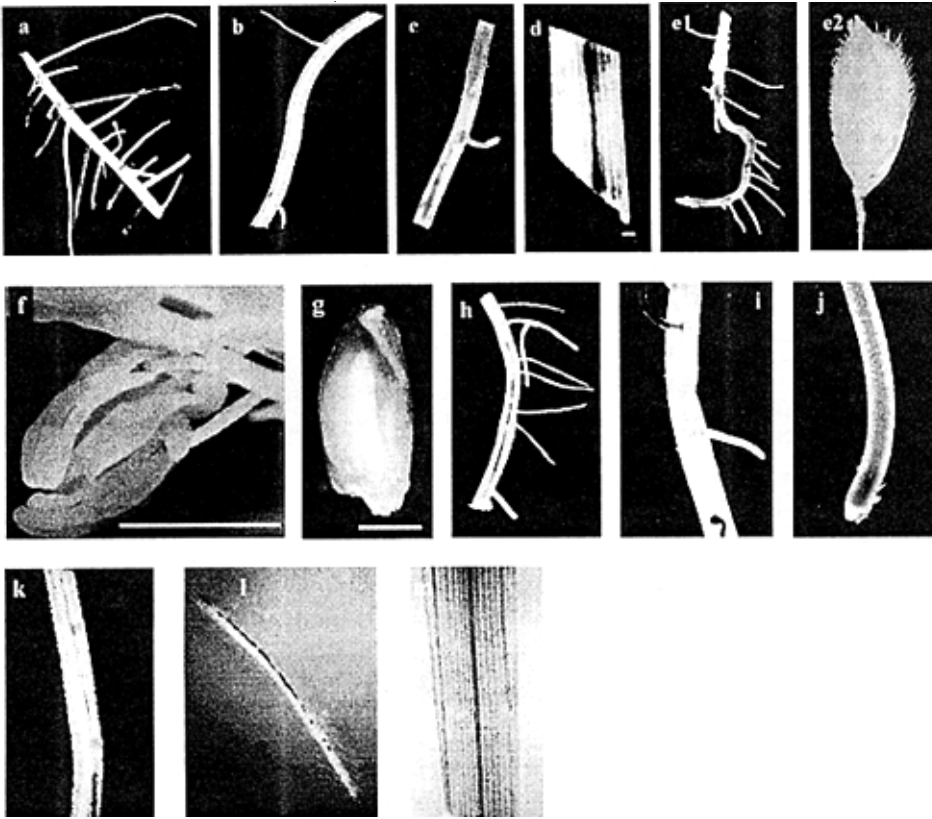


图 10 *DsE* and *DsG* 植株的 GUS 染色分析。

Figure 10. GUS staining analysis of *DsE* and *DsG* insertion plants

a-f: GUS staining photos of *DsG* insertion plants.

(a) Line C20-03.027 exhibits strongly GUS activity in lateral roots. (b) Line C20-05.090 exhibits the strong GUS activity at the base region of lateral roots. (c) Line C20-05.112 exhibits strong GUS activity in whole roots. (d) Line C20-05.194 exhibits strong GUS activity in main vein of leaf. (e1) and (e2) Line C20-05.209 exhibits GUS activity in root and glume. (f) Line C20-05.311 exhibits GUS activity at the junction of filament and anther.

g-m GUS staining photos of *DsE* insertion plants. (g) Line C03-14.272 exhibits GUS activity in seed capsule. (h) Line C14-03.113 exhibits strong GUS activity in stele. (i) Line C03-13.174 exhibits strong GUS activity in lateral root. (j) Line C11-02.030 exhibits strong GUS activity in whole root. (k) Line C03-13.191 exhibits GUS activity in hemi root where lateral root happened. (l) Line C03-02.024 exhibits strong GUS activity in hemi root. (m) Line C13-04.002 exhibits strong GUS activity in lateral vein of leaf.

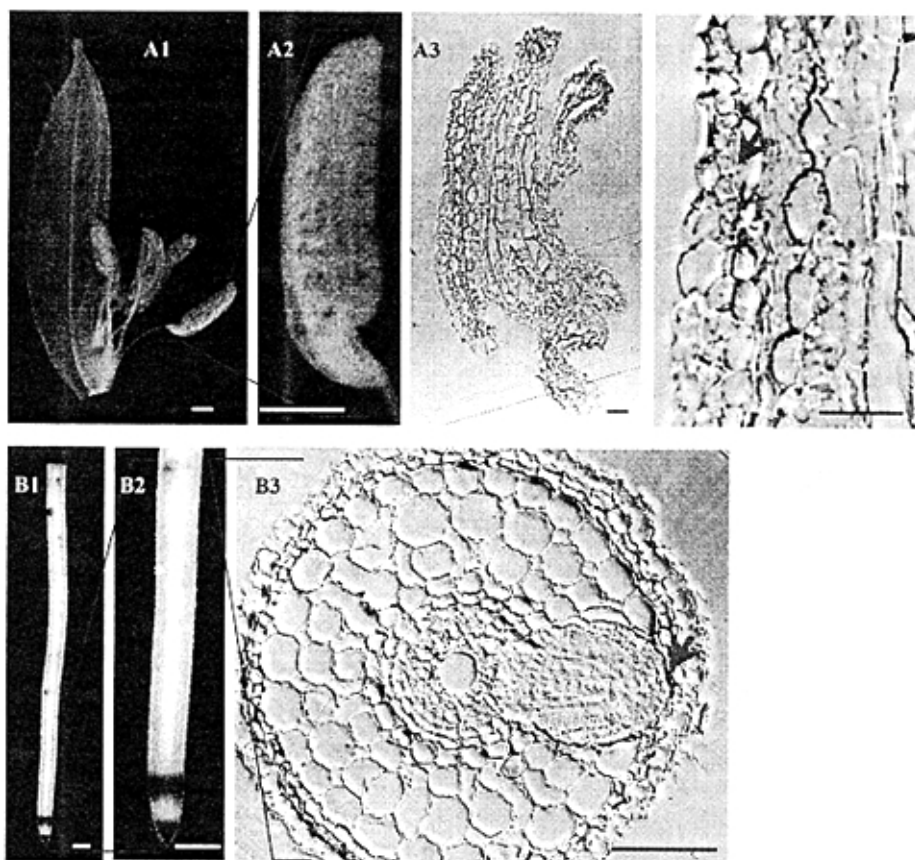


图 11 *DsE* 插入植株 C03-14.37 的 GUS 活性分析。

Figure 11. Analysis of GUS activity in *DsE* insertion rice plant C03-14.037.

A1-A4 indicated strong GUS activity at anther. B1-B3 indicated strong GUS activity at root tips. White bar: 200μm, black bar: 50μm.

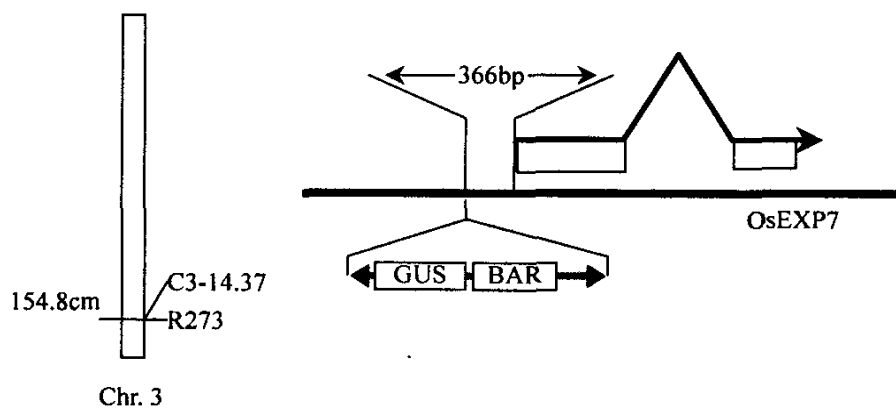


图 12 *DsE* 插入植株 C03-14.037 在水稻基因组中的插入位置分析。

Figure 12 Insertion site of *DsE* element at rice genome DNA in plant C03-14.037.

3 讨论

3.1 影响 *Ds* 转座的因素

在本研究中,我们采用含有增强子捕获器和基因捕获器的 *Ac/Ds* 转座子系统用作创建水稻突变体和基因标签。*Ds* 的转座频率在不同的系统与生物中变化很大,在已经报道的 *Ac/Ds* 转座子系统在拟南芥中转座频率为 50-60%(Bancroft 和 Dean1993, Ito 等 1999),在烟草中为 0-90%(Suzuki 等 2001),在水稻中为 0-51%(Nakagawa 等 2000),大麦中为 0-47%(Koprek 等 2000),在我们的研究中,转座的频率为 0%到 41%。

在不同的 F_2 群体中,即使它们具有同样来源的 *Ac* 转座酶的亲本,表现出很大转座频率很大的差异性,这说明 *Ds* 元件周围的基因组 DNA 序列肯定会影响转座的发生和发生的频率。

Ds 元件从供体位点剥离后重新插入到植物基因组的时候,在不同的系统中,有的时候在同一 F_2 群体中,*Ds* 的插入模式是一样的,而在有的系统中则发现在同一个 F_2 群体中 *Ds* 的插入可以插入到基因组 DNA 的不同位点,这种情况拟南芥、水稻和大麦中均有发现(Bancroft 和 Dean1993b, Koprek 等 2000, Nakagawa, 2000)。在来自同一 F_1 的 F_2 群体中,高比例的插入到同一位点被认为是由于 *Ac* 转座酶基因置于强启动子(如启动子 CaMV35 和 UBI)下,*Ac* 转座酶的高表达使得转座发生在发育的早期,因此在后代中会出现大量的同一转座位点的不同 F_2 植株(Balcells 和 Coupland 1994, Keller 等 1993, Long 等 1993)。而在比较弱的启动子下的 *Ac* 转座酶会在 F_2 代中得到不同插入位点的 F_2 植株。在我们的实验中,采用了强启动子 UBI 来启动 *Ac* 转座酶,但是得到的结果却与上面的结论不一样,通过对 8 个 F_2 群体的 Southern 杂交(图 7)发现独立的插入频率都在 70%以上(表 1)。在对 F_1 植株进行 PCR 检测 *Ds* 跳跃时,以成熟的水稻叶片 DNA 为模板都扩增不到 EDS 特征片段,所以认为转座发生是在 F_1 植株生长发育的后期,这可能是导致在我们这一系统中出现比率很高的独立转座频率的原因。同时也说明这一系统对于构建 *Ds* 插入突变体库是非常有效的。

3.2 *Ds* 元件在水稻中插入位置的特点

Ds 有一种倾向就是跳跃到核仁的组织区,这就是所谓的热点(Parinov 等 1999, 等 2002)。但是 *Ds* 还有一种倾向是跳跃到染色体的基因区,在拟南芥,有 48%的 *Ds* 插入子位于开放阅读框区(Raina 等 2002)。在我们的研究中,在总共扩增得到的 29 个 *Ds* 插入子中,有 7 个是位于外显子区域的,占有 24%,这

与 Enoki 等 (1999) 在 *Ac* 自主转座子系统中 29% 的结果比较一致。看起来似乎在水稻中 *Ds* 插入到基因的频率中只有在拟南芥中的一半, 我们分析认为在水稻中基因密度大约为每 10kb 为 1 个基因(Yu 等 2002), 而在拟南芥中相同长度的基因组序列则有 2 个基因。水稻和拟南芥中外显子的分布很相似, 所以我们认为 *Ds* 在水稻中的作为外显子的插入标签的能力可以与在拟南芥中的相比。

3.3 HPT 基因作为反向选择标记的讨论

为在 F_2 群体中获得稳定的 *Ds* 插入植株, 必需要得到与 *Ac* 转座酶相分离的 *Ds* 插入植株。为方便选择这样 *Ds* 插入植株, 应用反向选择标记就显得很有必要了, 在拟南芥中, *IAAH* 基因已成功的应用作为选择与供体位点及 *Ac* 转座酶不相连锁的 *Ds* 插入植株 (Sundaresan 等 1995)。水稻中 Chin(1999)应用细胞色素 P450 基因作为反向选择标记, 但是他没有作详细的描述, 效果如何不得而知。HPT 基因, 在我们的 *Ac*-T-DNA 及 *Ds*-T-DNA 亲本的转基因实验中是作为选择标记基因的, 在选择与供体位点及 *Ac* 转座酶不相连锁的 *Ds* 插入植株时可以作为反向选择标记。不含有 HPT 基因的水稻叶片在含有 100mg/L 潮霉素的 MS 培养基上培养, 一个星期后会坏死, 而含有 HPT 基因的水稻叶片则会保持绿色。(Ming-Bo Wang 和 Peter M. Waterhouse 1997)。在 39 个 F_2 植株中, 我们用这种方法筛选得到两株 PPT⁺HYG^s (C19-01.002 and C19-01.005) 植株。结果表明 HPT 基因可以用作反向选择标记以选择不含有供体位点及 *Ac* 转座酶的 *Ds* 插入植株, 这在大规模进行筛选独立远距离插入突变体是非常有效的。

3.4 基因捕获器的作用

基因捕获器和增强子捕获器是确定基因功能的一种很好的工具, 特别对于那些生长发育多效性基因及功能冗余的基因尤其有效(Springer 等 1995, Springer 1999)。表现出 GUS 活性的 *Ds* 插入植株对于基因功能的确定和表达模式的研究很有帮助。在我们的研究系统中, 基因捕获器的 GUS 活性表为 22% 增强子捕获 GUS 活性表为 28%。基因捕获器和增强子捕获器可以捕获不同时期特异表达的基因, 这些材料对于研究基因在特定时期的功能是非常有用的。在我们的研究中, 有一些 GUS 染色的植株非常有意思, 例如 C03-14.037 是一个增强子捕获器在根尖和花药表现出强烈的 GUS 活性, 但在叶中并不表现出 GUS 活性, TAIL-PCR 的结果表明, 它插入到的是水稻膨胀素基因 OsExp7 编码区 326bp 的位置, 也即是启动子部位, 其 F_2 植株已经表现出一定的生长状态的改变, 这对于 OsExp7 在水稻细胞伸长机理可能是非常好的一个材料。

3.5 大规模筛选的策略

基于本研究所获得的信息,我们认为可以采用本系统来获得大量的 *Ds* 插入子。例如 1 个 F_1 种子能产生 1000 个 F_2 , 则理论上能获得的独立稳定的 *Ds* 插入子数为 $1000 \times 1/4$ (无原始供体位点的 F_2 植株) $\times 30\%$ (转座子的平均跳跃频率) $\times 70\%$ (独立插入频率) $\times 1/4$ (无 *Ac* 元件的 F_2 植株频率) = 12。在这样的选择策略下,可以利用 HPT 基因作为反向选择标记来选择 *Ds* 与原来供体位点及 *Ac* 元件相分离的插入子。为了获得更多的 *Ds* 插入的 F_2 植株,而不考虑 *Ds* 与供体位点及 *Ac* 分离与否,则可以用本文中所采用的 PPT^r 抗性筛选含有 *Ds* 插入子、PCR 方法来筛选 EDS⁺FDS⁻ 或 EDS⁺FDS⁺ F_2 植株。根据我们前面的结果,这些 *Ds* 插入植株 70% 以上是独立插入子,这对于筛选突变体是非常有用的。此外,对于这些筛选到的 *Ds* 插入子还可通过 GUS 活性分析,以筛选 GUS 报告基因组织特异或时空特异性表达的突变体,以进一步研究特定基因的功能。

参考文献

1. 奥斯伯 F, 布伦特 R, 金斯顿 RE, 穆尔 DD, 塞德曼 JG, 史密斯 JA, 斯特拉尔 K. (1998) 精编分子生物学实验指南 科学出版社 北京 朱立煌 李素文 (2001) 转基因水稻 T-DNA 侧翼序列的扩增与分析 遗传学报 28: 345-351 卢扬江 郑康乐 提取水稻 DNA 的一种简易方法 中国水稻科学 1992 6: 47-48.
4. 萨母布鲁克 J, 弗里奇 EF, 曼尼阿蒂斯 T. (1995) 分子克隆实验指南 科学出版社 1995
5. 王关林 方宏筠 植物基因工程原理与技术 科学出版社 北京 1998
6. 曾广方 蒋德安 植物生理学 成都科技大学出版社 1998 p35
7. 朱冰 黄大年 杨炜 等1996 利用基因枪法获得可遗传的抗除草剂转基因水稻植株 中国农业科学 29(6): 15-20
8. 朱正歌 肖晗 傅亚萍 胡国成 于永红 斯华敏 张景六 孙宗修 2001 水稻转座子突变体库的构建及突变类型的遗传分析 生物工程学报 17(3): 288-292.
9. Aarts MGM, Dirkse WG, Stiekema WJ and Pereira A. (1993) Transposon tagging of a male sterility gene in *Arabidopsis*. *Nature* 363: 715-717.
10. Aldemita RR and Hodges TK. (1996) *Agrobacterium tumefaciens*-mediated transformation of Japonica and Indica rice varieties. *Planta* 199: 612-617
11. Allen ND, Cran DG, Barton SC, Hettle S, Reik W, and Surani MA. (1988) Transgenes as probes for active chromosomal domains in mouse development. *Nature* 333: 852-855
12. André D, Colau D, Schell J, Van Montagu M, and Hernalsteens JP (1986) Gene tagging in plants by a T-DNA insertion mutagen that generates APH(39)II-plant gene fusions. *Mol. Gen. Genet.* 204: 512-518.
13. Balcells L and Coupland G. (1994) The presence of enhancers adjacent to the *Ac* promoter increases the abundance of transposase mRNA and alters the timing of *Ds* excision in *Arabidopsis*. *Plant Mol.Biol.* 24: 789-798.
14. Bancroft I and Dean C. (1993a) Factors affecting the excision frequency of the maize transposable element *Ds* in *Arabidopsis thaliana*. *Mol.Gen.Genet.* 240: 65-72.
15. Bancroft I and Dean C. (1993b) Transposition pattern of the maize element *Ds* in *Arabidopsis thaliana*. *Genetics* 134: 1221-1229.
16. Barakat A, Gallois P, Raynal M, Mestre-Ortega D, Sallaud C, Guiderdoni E, Delseny M, Bernardi G. (2000) The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice. *FEBS lett.* 471: 161-164.
17. Bier E, Vaessin H, Shepherd S, Lee K, McCall K, Barbel S, Ackerman L, Carretto R, Uemura T, Grell E, Jan LY and Jan YN. (1989). Searching for pattern and mutation in the *Drosophila* genome with a P-*lacZ* vector. *Genes Dev.* 3: 1273-1287.
18. Brennan J, and Skarnes WC. (1999) Gene trapping in mouse embryonic stem cells. *Methods Mol. Biol.* 97: 123-138.

19. Brunaud V, Balzergue S, Dubreucq B, Aubourg S, Samson F, Chauvin S, Bechtold N, Cruaud C, DeRose R, Pelletier G, Lepiniec L, Caboche M and Lecharny A. (2002) T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites. *EMBO Rep.* 3: 1152-1157.
20. Chan MT, Chang HH, Ho SL, Tong WF and Yu SM (1993) Agrobacterium-mediated production of transgenic rice plants expressing a chimeric α -amylase promoter/ β -glucuronidase gene. *Plant Mol. Biol.* 22: 491-506.
21. Chan MT, Lee TM and Chang HH (1992) Transformation of indica rice (*Oryza sativa* L.) mediated by *Agrobacterium tumefaciens*. *Plant Cell Physiol.* 33: 577-583.
22. Chang YC, Pfitzner AJP, Pfitzner UM, Chang-KF, Chen CM, Tu J and Kuo TT (2000) Construction of an inducible transposon, INAc, to develop a gene tagging system in higher plants *Molecular Breeding.* 6: 353-367.
23. Chin HG, Choe MS, Lee SH, Park SH, Koo JC, Kim NY, Lee JJ, Oh BG, Yi GH, Kim SC, Choi HC, Cho MJ, and Han CD. (1999) Molecular analysis of rice plants harboring an *Ac/Ds* transposable element-mediated gene trapping system. *Plant J.* 19:615-623.
24. Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H and Sheen J. (1996) Engineered GFP as a vital reporter in plants. *Curr.Biol.* 6: 325-330.
25. Christou P, Ford TM and Kofron M. (1991) production of transgenic rice (*Oryza sativa* L.) plants from agronomically important Indica and Japonica varieties via electric discharge particle acceleration of exogenous DNA into immature zygotic embryos. *Bio/Technology* 9: 957-962.
26. Chuck G, Robbins T, Nijjar C, Ralston E, Chourtney-Gutterson N and Dooner HK. (1993) Tagging and cloning of a petunia flower color gene with the maize transposable element Activator. *Plant Cell* 5:371-378.
27. De Buck S, Jacobs A, Van Montagu M and Depicker A. (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. *Plant J.* 20: 295-304.
28. Dong JJ, Teng WM, Buchholz WG and Hall TC (1996) Agrobacterium-mediated transformation of Javanica rice. *Molecular Breeding* 2: 267-276.
29. Dooner HK and Belachew A. (1989) Transposition Pattern of the Maize Element *Ac* from the *bz-m2(Ac)* Allele. *Genetics* 122: 447-457.
30. Dooner HK, Keller J, Harper E, and Ralston E. (1991) Variable Patterns of Transposition of the Maize Element Activator in Tobacco. *Plant Cell* 3: 473-482.
31. Enoki H, Izawa T, Kawahara M, Komatsu M, Koh S, Kyojuka J and Shimamoto K. (1999) *Ac* as a tool for the functional genomics of rice. *Plant J.* 19: 605-613.
32. Fedoroff NV and Smith DL. (1993) A versatile system for detecting transposition in Arabidopsis. *Plant J.* 3: 273-289.
33. Feldmann KA. (1991) T-DNA insertion mutagenesis in Arabidopsis: mutational spectrum.

- Plant J. 1: 71-82.
34. Fortsthoefel NR, Wu Y, Schulz B, Bennett MJ and Feldmann KA. (1992) T-DNA insertion mutagenesis in Arabidopsis: prospects and perspectives. Aust. J. Plant Physiol. 19: 353-366
35. Gheysen G, Villaroel R and Van Mantagu M. (1991) Illegitimate recombination in plants: a model for T-DNA integration. Genes Dev. 5: 287-297.
36. Goldsbrough AP, Tong Y and Yoder JI (1996) *Lc* as a non-destructive visual reporter and transposition excision marker gene for tomato. Plant J. 9: 927-933.
37. Gordon-Kamm WJ, Spencer TM, Mangano ML, Adams TR, Daines RJ, start WG, O'Brien JV, Chambers SA, Adams WR, Jr Willetts NG, Rice TB, Mackey CJ, Krueger RW, Kausch AP and Lemaux PG (1990) Transformation of maize cells and regeneration of fertile transgenic plants. Plant Cell 2: 603-618.
38. Greco R, Ouwerkerk BF, Sallaud C, Kohli A, Colombo L, Puigdomenech P, Guiderdoni E, Christou P, Hoge JHC and Pereira A (2001) Transposon insertional Mutagenesis in rice. Plant physiol. 125: 1175-1177
39. Grossniklaus U, Bellen HJ, Wilson C and Gehring WJ. (1989) P-element-mediated enhancer detection applied to the study of oogenesis in *Drosophila*. Development 107: 189-200.
40. Hanin M, Volrath S, Bogucki A, Briker M, Ward E, Paszkowski J. (2001) Gene targeting in Arabidopsis. Plant J. 28: 671-677.
41. Healy J, Corr C, Deyoung J and Baker B (1993) Linked and unlinked transposition of a genetically marked Dissociation element in transgenic tomato. Genetics 134: 571-584.
42. Hiei Y, Ohta S, Komari T and Kumashiro T. (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J. 6: 271-282.
43. Ito T, Seki M, Hayashida N, Shibata D and Shinozaki K. (1999) Regional insertional mutagenesis of genes on Arabidopsis thaliana chromosome V using the *Ac/Ds* transposon in combination with a cDNA scanning method. Plant J. 17: 433-444.
44. Izawa T, Miyazaki C, Yamamoto M, Terada R, Iida S and Shimamoto K. (1991) Introduction and transposition of the maize transposable element *Ac* in rice (*Oryza sativa* L.). Mol. Gen. Genet. 227: 391-96.
45. Izawa T, Ohnishi T, Nakano T, Ishida N, Enoki H, Hashimoto H, Itoh K, Terada R, Wu C, Miyazaki C, Endo T, Iida S, and Shimamoto K. (1997) Transposon tagging in rice. Plant Mol. Biol. 35: 219-229.
46. Jeon JS, Lee S, Jung KH, Jun SH, Jeong DH, Lee J., Kim C., Jang S , Yang,K., Nam,J., An,K., Han,M.J., Sung,R.J., Choi,H.S., Yu,J.H., Choi,J.H., Cho,S.Y., Cha,S.S., Kim,S.I., and An,G. (2000) T-DNA insertional mutagenesis for functional genomics in rice. Plant J. 22:561-570.
47. Jeong DH, An S, Kang HG, Moon S, Han JJ, Park S, Lee HS, An K and An G. (2003) T-DNA Insertional Mutagenesis for Activation Tagging in Rice. Plant Physiol. 130: 1636-1644.

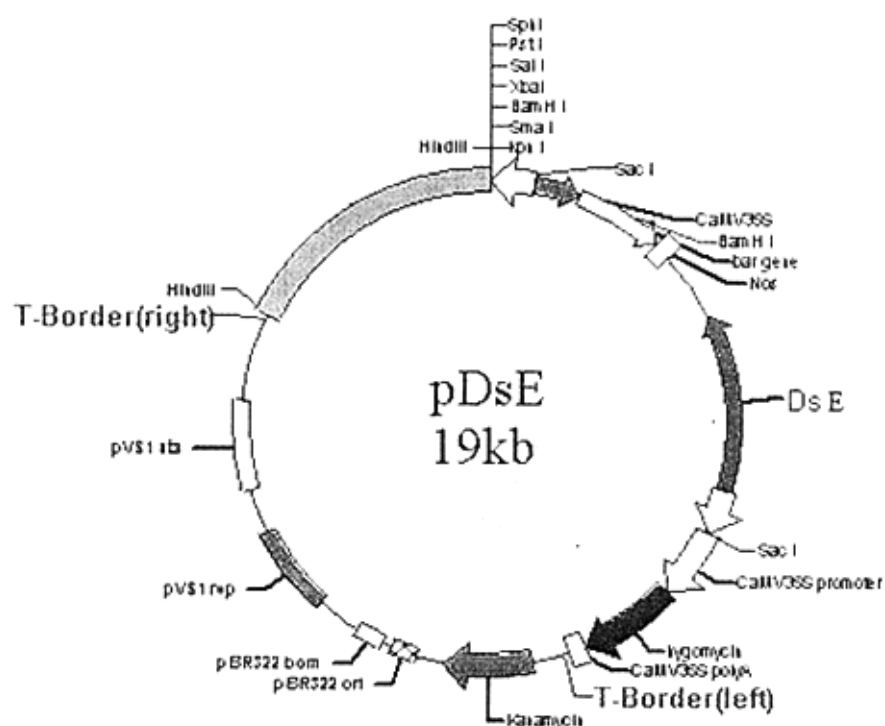
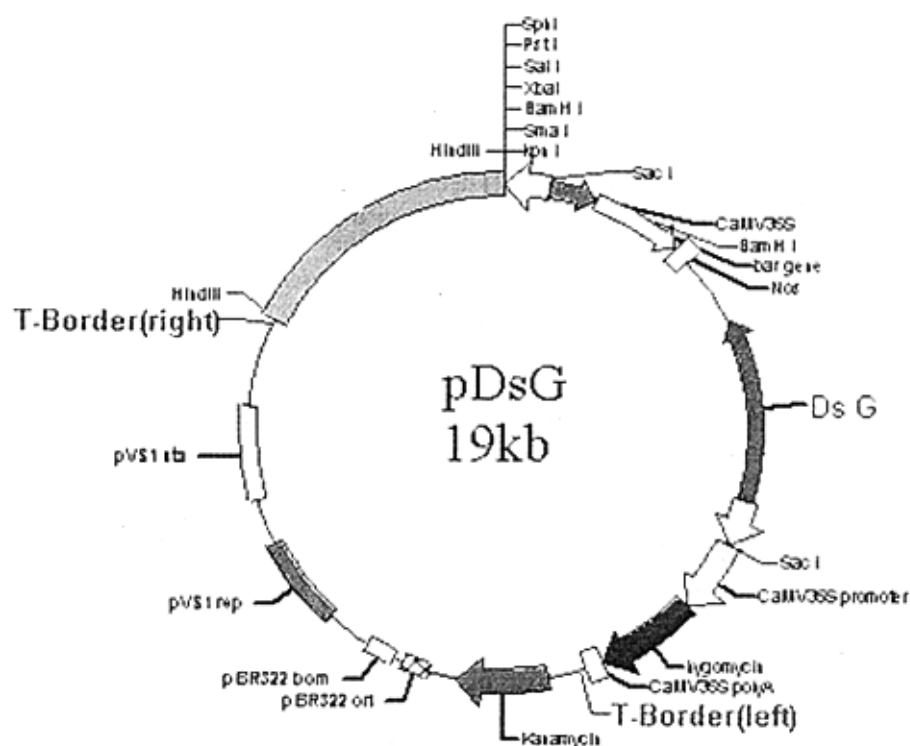
48. Jones DA, Thomas CM, Hammond Kosack KE, Balintkurti PJ and Jones JDG. (1994) Isolation of the tomato Cr-9 gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* 266: 789-793.
49. Jones JDG, Carland FM Maliga P and Dooner HK (1989) Visual detection of transposition of the maize element Activator (Ac) in tobacco seedlings. *Science* 244: 204-207.
50. Keller J, Jones DJ, Harper E, Lim E, Carland F, Ralston EJ and Dooner HK (1993) Effects of gene dosage and sequence modification on the frequency and timing of transposition of the maize element *Activator* (Ac) in tobacco. *Plant Mol. Biol.* 21: 157-170.
51. Kempin SA, Liljegren SJ, Block LM, Rounsley SD, Yanofsky MF and Lam E. (1997) Targeted disruption in *Arabidopsis*. *Nature* 389(6653): 802-803.
52. Kertbundit, S., De Greve, H., Deboeck, F., Van Montagu, M., and Hernalsteens, J.-P. (1991). *In-vivo* random b-glucuronidase gene fusions in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 88: 5212-5216.
53. Klee HJ, Horsch RB, Hinchee MA., Hein MB and Hoffmann NL. (1987). The effects of overproduction of two *Agrobacterium tumefaciens* T-DNA auxin biosynthetic gene products in transgenic petunia plants. *Genes Dev.* 1: 86-96.
54. Kohli A, Leech M, Vain P, Laurie DA and Christou P (1998) Transgene organization in rice engineered through direct DNA transfer supports a two-phase integration mechanism mediated by the establishment of integration hotspots. *Proc. Natl. Acad. Sci. USA* 95: 7203-7208
55. Koncz C, Nemeth K, Redei GP and Schell J. (1992) T-DNA insertion mutagenesis in *Arabidopsis*. *Plant Mol. Biol.* 20: 963-967
56. Koprek T, McElroy D, Louwerse J, Williams-Carrier R and Lemaux PG. (2000) An efficient method for dispersing *Ds* elements in the barley genome as a tool for determining gene function. *Plant J.* 24: 253-263.
57. Kothary R, Clapoff S., Brown A, Campbell R, Peterson A, and Rossant J. (1988) A transgene containing *lacZ* inserted into the *dystonia* locus is expressed in neural tube. *Nature* 335: 435-437
58. Krizkova L and Hroudá M. (1998) Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. *Plant J.* 16: 673-680.
59. Krysan PJ, Young JC and Sussman MR. (1999) T-DNA as an insertional mutagen in *Arabidopsis*. *Plant cell* 11: 2283-2290.
60. Liu YG, Mitsukawa N, Oosumi T and Whittier RF. (1995) Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* 8: 457-463.
61. Liu YG and Whittier RF. (1995) Thermal asymmetric interlaced PCR: Automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 25: 674-681.

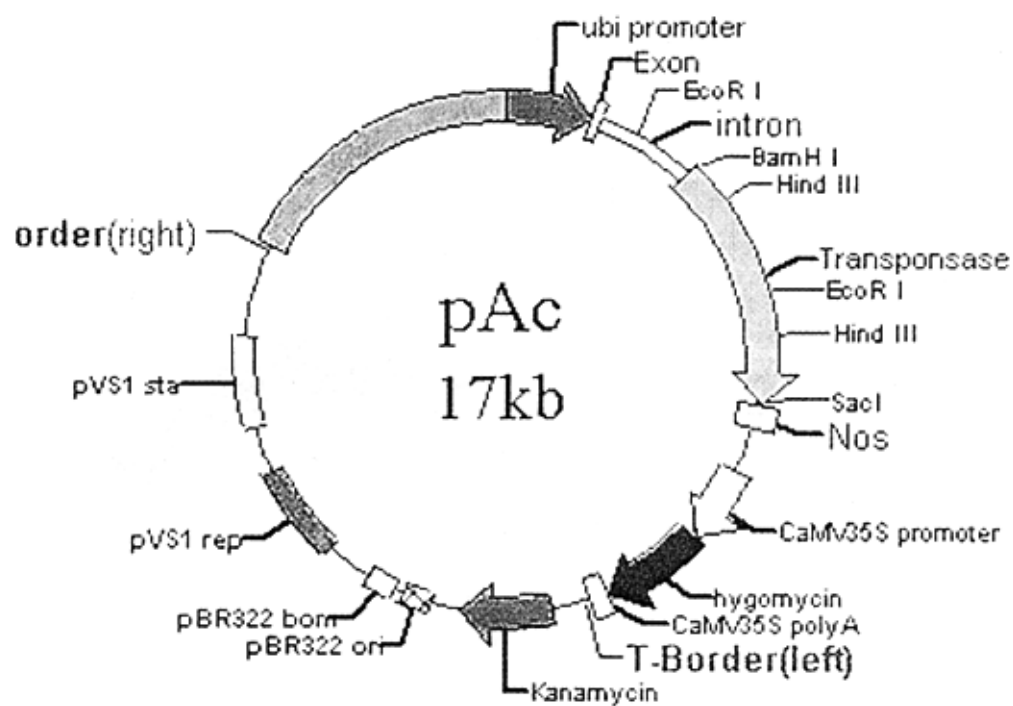
62. Lloyd AM, Walbot V and Davis RW. (1992) *Arabidopsis* and *Nicotiana* anthocyanin production activated by maize regulators *R* and *C1*. *Science* 258: 1773-1775.
63. Long D, Swinburne J, Martin M, Wilson K, Sundberg E, Lee K and Coupland G. (1993b) Analysis of the frequency of inheritance of transposed *Ds* elements in *Arabidopsis* after activation by a CaMV 35S promoter fusion to the *Ac* transposase gene. *Mol.Gen.Genet.* 241: 627-636.
64. Maria E, Burgund B, Stanton B. (1997) Intergration of T-DNA binary vector 'backbone' sequence into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J.* 11(5): 945-957.
65. Mayerhofer R, Koncz-Kalman Z, Nawrath C, Bakkeren G, Cramer A., Angelis K, Redei GP, Schell J, Hohn B and Koncz C. (1991) T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J.* 10: 697-704.
66. McClintock B. (1954) Mutations in maize and chromosomal aberrations in *Neurospora*. *Carnegie Inst. Wash. Year book* 54: 254-260.
67. Ming-Bo Wang and Peter M. Waterhouse (1997) A Rapid and Simple Method of Assaying Plants Transformed with Hygromycin or PPT Resistance Genes. *Plant Mol. Bio. Rep.* 15: 209-215.
68. Nakagawa Y, Machida C, Machida Y and Toriyama K. (2000) Frequency and pattern of transposition of the maize transposable element *Ds* in transgenic rice plants. *Plant Cell Physiol.* 41: 733-742.
69. O'Kane CJ and Gehring WJ. (1987). Detection *in situ* of genomic regulatory elements in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 84: 9123-9127.
70. O'Keeffe DP, Tepperman JM, Dean C, Leto KJ, Erbes DL and Odell JT. (1994). Plant expression of a bacterial cytochrome P450 that catalyzes activation of a sulfonylurea pro-herbicide. *Plant Physiol.* 105: 473-482.
71. Parinov S, Sevugan M, De Y, Yang WC, Kumaran M and Sundaresan V. (1999) Analysis of flanking sequences from dissociation insertion lines: a database for reverse genetics in *Arabidopsis*. *Plant Cell* 11: 2263-2270.
72. Pereira A, Schwarz-Sommer Zs, Gierl A, Bertram I, Peterson PA and Saedler H. (1985) Genetic and molecular analysis of the Enhancer (En) transposable element system of *Zea mays*. *EMBO J.* 4: 17-25.
73. Peterson PA. (1953) A mutable pale green locus in maize. *Genetics* 38: 682-683.
74. Raina S, Mahalingam R, Chen F and Fedoroff N. (2002) A collection of sequenced and mapped *Ds* transposon insertion sites in *Arabidopsis thaliana*. *Plant Mol.Biol.* 50: 93-110.
75. Rashid H, Yokoi S, Toriyama K and Hinata K (1996) Transgenic plant production mediated by *Agrobacterium* in Indica rice. *Plant Cell Reports*, 15: 727-730.
76. Richmond T and Somerville S. (2000) Chasing the dream: Plant EST microarrays. *Curr. Opin. Plant Biol.* 3: 108-116.

77. Ross-Macdonald P, Coelho P, Roemer T, Agarwal S, Kumar A, Jansen R, Cheung K, Sheehan A, Symoniat D, Ulmanský L, Heidtman M, Nelson F K, Iwaskai H, Hager K, Gerstein M, Miller P, Roeder GS and Snyder M. (1999). Large-scale Analysis of the Yeast Genome by Transposon Tagging and Gene Disruption. *Nature* 402: 413-418.
78. Sheen J, Hwang S, Niwa Y, Kobayashi H and Galbraith DW. (1995) Green-fluorescent protein as a new vital marker in plant cells. *Plant J.* 8: 777-784.
79. Skarnes WC. (1993). The identification of new genes: Gene trapping in transgenic mice. *Curr. Opin. Biotechnol.* 4: 684-689.
80. Somers DA, Rines HW, Gu W, kaeppler HF and Bushnell WR (1992) Fertile, transgenic oat plants. *Bio/Technology* 10: 1589-1594.
81. Springer PS, McCombie WR, Sundaresan V and Martienssen RA. (1995) Gene trap tagging of *PROLIFERA*, an essential MCM2-3-5-like gene in Arabidopsis. *Science* 268: 877-880.
82. Springer PS. (2000) Gene trap: tools for plant development and genomics. *Plant cell* 12(7): 1007-1020.
83. Stougaard J. (1993) Substrate-dependent negative selection in plants using a bacterial cytosine deaminase gene. *Plant J.* 3: 755-761.
84. Sundaresan V, Springer P, Volpe T, Haward S, Jones JD, Dean C, Ma H and Martienssen R. (1995) Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* 9: 1797-1810.
85. Suzuki Y, Uemura S, Saito Y, Murofushi N, Schmitz G, Theres K and Yamaguchi I. (2001) A novel transposon tagging element for obtaining gain-of-function mutants based on a self-stabilizing *Ac* derivative. *Plant Mol. Biol.* 45:123-131.
86. Szabados L, Kovacs I, Oberschall A, Abraham E, Kerekcs I, Zsigmond L, Nagy R, Alvarado M, Krasovskaja I, Gal M, Berente A, Redei GP, Haim AB and Koncz C. (2002) Distribution of 1000 sequenced T-DNA tags in the Arabidopsis genome. *Plant J.* 32: 233-242.
87. Takano M, Egawa H, Ikeda JE, Wakasa K (1997) Thr structures of integration sites in transgenic rice. *Plant J.* 11: 353-361.
88. Tarada R, Hrawa H, Inagaki Y, Tsugane K and Iida S. (2002) Efficient gene targeting by homologous recombination in rice. *Nat Biotechnol.* 20(10): 1030-1034.
89. Teeri, T.H., Herrera-Estrella, L., Depicker, A., Van Montagu, M., and Palva, E.T. (1986). Identification of plant promoters *in situ* by T-DNA-mediated transcriptional fusions to the *npt-II* gene. *EMBO J.* 5: 1755-1760.
90. Tinland B. (1996) The integration of T-DNA into plant genomes. *Trends Plant Sci.* 1: 178-184.
91. Topping JF, Wei W and Lindsey K. (1991). Functional tagging of regulatory elements in the plant genome. *Development* 112: 1009-1019.
92. van Enckevort LJG, Lasschuit J, Stiekema WJ, Jacobsen E and Pereira A. (2001) Development of *Ac* and *Ds* transposon tagging lines for gene isolation in diploid potato.

- Molecular breeding 7: 117-129.
93. Vasil V, Castillo M, Fromm ME and Vasil IK (1992) Herbicide resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus. *Bio/Technology*, 10:667-678.
94. Wang YC, Klein TM, Fromm M, Cao J, Sanford JC and Wu R (1988) Transient expression of foreign genes in rice, wheat and soybean cells following particle bombardment. *Plant Mol Bio*, 11: 33-439.
95. Weil CF and Kunze R. (2000) Transposition of maize *Ac/Ds* transposable elements in the yeast *Saccharomyces cerevisiae*. *Nat.Genet.* 26: 187-190.
96. Whitham S, Dineshkumar SP, Choi D, Hehl R, Corr C and Baker B. (1994) The product of the tobacco mosaic virus resistance gene N, similarity to *TOLL* and interleukin-1 receptor. *Cell* 78: 1101-1115.
97. Wilson C, Pearson RK, Bellen HJ, O'Kane CJ, Grossniklaus U and Gehring WJ. (1989). P-element-mediated enhancer detection: An efficient method for isolating and characterizing developmentally regulated genes in *Drosophila*. *Genes Dev.* 3: 1301-1313.
98. Yu J., Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L and Yang H. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79-92.
99. Zhang W and Wu R (1998) Efficient regeneration of transgenic plants from rice protoplasts and correctly regulated expression of the foreign gene in the plants. *Theor Appl Genet.* 76: 835.
100. Zupan J, Muth TR, draper O and Zambryski P. 2000 The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. *Plant J.* 23(1): 11-28.

附录 1





攻读博士学位期间发表的科研论文

1. **Weizheng Jin**, Fan Zhang, Yunrong Wu, Saomin Wang, Min Xu, Ruijun Duan, Shuangyan Chen, Feiyan Liu, Ping Wu* (2003) *Ac/Ds* system harboring enhancer trap and gene trap for gene tagging in rice. Submitted
2. 金维正 段瑞君 吴运荣 吴平* 2003 利用 *Ac/Ds* 转座子系统在水稻中获得无选择标记转基因植株的方法 生物工程学报 (接收)
3. 金维正 张帆 陈双燕 侯兴亮 刘洪家 吴运荣 周洁 刘非燕 吴平 T-DNA 插入标签在水稻中的应用 (投稿)
4. Shuangyan Chen, **Weizheng Jin**, Mingyi Wang, Fan Zhang, Jie Zhou, Qiaojun Jia, Yunrong, Wu, Feiyan Liu, Ping Wu*. Distribution and characterization of over 1000 T-DNA tags in rice genome. (submitted).
5. 张帆 金维正 陈双燕 吴运荣 刘非燕 吴平* 2003 质粒营救法和 TAIL-PCR 方法获得水稻 T-DNA 旁邻序列的效率比较 农业生物技术学报 (接收)