根癌农杆菌介导的大豆遗传转化

林树柱[1] 曹越平[1*] 卫志明[2]

[1](上海交通大学农业与生物学院,上海 201101)

[2](中国科学院上海生命科学研究院植物生理生态研究所,上海 200032)

摘 要 农杆菌介导法是大豆遗传转化的重要方法之一,许多实验室应用该方法得到了转基因大豆,但目前使用该方法进行转化的效率还比较低,尚需深入研究。农杆菌菌株、大豆基因型、组织培养条件、T-DNA 的转移效率和转化后的筛选模式都会影响大豆转化的效率。概述了近年来根癌农杆菌介导的大豆遗传转化的一些重要成果,以及转化过程中大豆的易感性与农杆菌的转化能力、乙酰丁香酮促进 *vir* 基因活化、转化的受体系统和巯基混合物减轻受体材料的褐化、提高 T-DNA 的转移效率等几个重要因素的研究进展,并介绍了转化中常用的几个筛选标记基因(*npt*Ⅱ、*hpt*、*bar* 基因和突变的 *ahas* 基因)及通过共转化法去除标记基因的方法,同时对今后研究的重点进行了讨论。

关键词 根癌农杆菌,大豆,遗传转化

中图分类号 Q75 **文献标识码** A **文章编号** 1000-3061(2004)06-0817-04

农杆菌介导的遗传转化是通过感染将所带的经过或未经改造的 T-DNA 导入植物细胞,引起相应的植物细胞可遗传的变异[1]。自 1983 年采用根癌农杆菌 Ti 质粒作为载体通过遗传转化产生第一例转基因烟草以后,20 余年来农杆菌介导的转化系统在双子叶植物中得到了广泛的应用。1988 年 Hinchee 等[2]首次用农杆菌介导法获得大豆转基因植株。到目前为止农杆菌介导法因其外源基因整合拷贝数较低且完整,操作简便,并可转移大于 30kb 的外源基因,一直是大豆转化的重要方法。

1 农杆菌介导的大豆遗传转化的一些重要成果

1988 年 Hinchee 等[2]首次用农杆菌介导法获得大豆转基因植株。他们从 100 个栽培品种中筛选到对农杆菌敏感的品种 Peking,用去毒的农杆菌转化其子叶外植体,再以卡那霉素进行筛选,获得再生植株。经检测再生植株中仅有 6% 为转基因植株,后代分析表明外源基因是以单拷贝整合进大豆基因组的。Townsend 和 Thomas[3]用与 Hinchee 等[2]同样的方法获得了转基因大豆栽培种"Poineer 9341"。2000 年 Clemente 等[4]也使用了与 Hinchee 等[2]相似的方法,以草甘膦作为筛选剂,获得了 196 个可育的转基因植株。

1996 年 Di 等[5]用农杆菌/子叶节系统以卡那霉素为筛选剂将 BPMV(bean pod mottle virus)外壳蛋白基因转入大豆。1997 年徐香玲等[6]以 Ti 质粒为载体将 B.T.k-δ 内毒素蛋白基因导入大豆,经卡那霉素筛选,得到 7 株转基因大豆,并获得后代。1999 年 Zhang 等[7]以 *bar* 基因作选择标记基因,用除草剂进行筛选获得了转化的植株,但是转化频率仍比较低。Donaldson 等[8]用农杆菌转化 12 个短季大豆品种子叶节,感染率可达 92%,但仅从品种 Accolibri 得到了遗传稳定的转基因植株,且转化率极低。2001 年 P.M.Olhoft 等[9]发现在固体共培养培养基中加入 L-半胱氨酸(L-Cys)能大大提高农杆菌对子叶节区的感染率,同时能降低转化后组织培养时外植体的褐化现象,从而使转化效率提高。之后 P.M.Olhoft 等[10]进一步将这一系统的转化效率提高到了 16.4%。

1989 年 Parrott 等[11]用 14 个大豆品种的未成熟子叶为外植体,接种土壤农杆菌 EHA101 或 LBA4404,两株菌都含有质粒 pH5PZ3D,该质粒带有 15kD 的玉米醇溶蛋白基因(*zein*)和新霉素磷酸转移酶基因(*npt*Ⅱ),通过体细胞胚发生得到 18 株再生植株,其中有三株表现为转基因植株,但都为嵌合体。2000 年 B.Yan 等[12]利用根癌农杆菌感染大豆未成熟子叶,通过体细胞胚发生过程得到三株转基因的再生植株,转化效率为 0.03%。这一系统的转化效率虽然还很低,但是由于体细胞胚发生再生系统自身的优点,这一系统可以有效的解决转化后产生嵌合体的问题,所以成为当前大豆转化研究的一个热点。

最近一些新的、潜在高效的农杆菌转化方法,如超声辅助农杆菌转化法(SAAT)[13,14]、真空抽滤辅助农杆菌转化完整植株法、微粒轰击与农杆菌综合法得到了发展。SAAT 法是将与农杆菌接触的植物组织先进行短期的超声处理,使外

收稿日期:2004-03-26,修回日期:2004-06-18。

基金项目:国家自然科学基金资助(No.30370892,30370790);上海市科委基础研究重点项目(No.02DJ14047)。

* 通讯作者。 Tel:86-21-64788330; Fax:86-21-64788330; E-mail: yuepingcao@sjtu.edu.cn

植体表面形成数千条穿过组织的小隧道,有利于农杆菌的感染进而转化组织。这一方法的应用必将提高农杆菌介导的转化效率,同时使农杆菌介导的转化方法也广泛的应用于其他目标组织。真空抽滤可以产生与 SAAT 类似的效果。微粒轰击与农杆菌综合法是在农杆菌感染之前或之后,进行微粒轰击以提高转化效率[15]。

2 根癌农杆菌介导的大豆遗传转化的几个重要因素的研究

虽然有关农杆菌介导的大豆遗传转化的研究已经很多,但是转化效率依然比较低,转化还远远没有模式化,转化受到多种因素的影响。

2.1 大豆的易感性与农杆菌的转化能力

虽然双子叶植物是农杆菌的天然宿主,但是农杆菌介导的大豆转化仍然存在一定程度的种和组织的特异性[16]。Owens 和 Cress 通过对 24 个栽培大豆品种和 3 个野生大豆株系活体转化研究,筛选出 3 个栽培品种("Biloxi"、"Jupiter"和"Peking")及 1 个野生材料(PI393.693B)对农杆菌 A348(pTiA6)高度易感[1]。

另外不同农杆菌菌株对大豆的转化能力也有差异,Byrne以大豆品种"Peking"为材料对 11 个根癌农杆菌菌株的转化能力进行了研究。以致瘤频率和瘤的大小为指标,他们发现不同菌株的转化能力的差异主要是 Ti 质粒的差异所致,胭脂碱型 Ti 质粒的转化能力强于章鱼碱型 Ti 质粒。转化中选择对农杆菌敏感的大豆基因型和强毒农杆菌菌系有利于转化效率的提高[1]。

2.2 促进 *vir* 基因活化的研究

根癌农杆菌侵染植物细胞,首先要求 *vir* 基因的活化。农杆菌 *vir* 区的活化是通过 *virA*/*virG* 双组分系统调节的[17]。酚类物质如乙酰丁香酮可以诱导农杆菌 *vir* 基因的活化,首先是 *virA* 受到酚类诱导物的诱导后激活 *virG*,然后进一步激活 *vir* 区基因的表达。*vir* 基因的表达控制着 T-DNA 的转移。在大豆转化的共培养培养基中添加乙酰丁香酮可以部分克服农杆菌介导的大豆转化的种和组织的特异性[16],是农杆菌介导的大豆转化成功的重要因素[4]。转化中应用的乙酰丁香酮的浓度通常是 100 ~ 200μmol/L[10,18]。

2.3 农杆菌介导的大豆遗传转化的受体系统

高再生能力受体系统与转化方法的匹配是大豆遗传转化成功的关键,建立高效的大豆转化受体系统具有重要意义。

大豆不定芽器官发生再生系统是根癌农杆菌介导的大豆遗传转化中最常用的受体系统。这一系统是由外植体细胞直接分化出不定芽获得再生植株。目前大豆不定芽器官发生再生系统所用的外植体主要有:无菌苗子叶[2],子叶节[19,20],上胚轴[21]和茎尖[22]等,其中最常用的是子叶节再生系统。诱导不定芽所用的生长调节物质一般都是 6-BA,基本培养基一般为 MS 培养基或 B5 培养基。1996 年 Y. Kaneda 等[23]报道应用 TDZ(thidiazuron,噻重氮苯基脲)和低盐浓度的基本培养基可以提高大豆的再生频率。采用这一系统的优点是:获得再生植株的时间短,通常 1 ~ 3 个月就可以获得生根的再生植株;系统变异小遗传稳定性好;而且比较少地受材料和季节的限制。但是,由于不定芽常常起源于多细胞,所以采用该系统转化容易产生嵌合体。

大豆体细胞胚发生再生系统也是农杆菌介导的大豆转化的重要受体系统。这一系统主要采用的外植体是未成熟胚的子叶[24,25],胚轴[26]或完整幼胚[27,28]。1983 年 Christianson 等[26]首次对大豆体细胞胚再生进行了报道。1988 年 Finer 等[29,30]报道了体细胞胚发生悬浮培养系统,在这一系统中将未成熟的大豆子叶放入含高浓度(40mg/L)2,4-D 的 MSB(MS 基本盐 + B5 有机盐)培养基中诱导体细胞胚发生,在含有较低浓度 2,4-D(5mg/L)的液体培养基中进行胚性组织的增生。球形体细胞胚在不含 2,4-D 的固体培养基上进行再生。在这一系统中因为新的体细胞胚是在旧胚的表面或近表面增生的,且能在众多位点上增生,所以转化非常方便;而且由于液体培养使得培养基与胚性细胞团可以充分接触,使得筛选更加有效。但是体细胞胚发生再生系统传代时间长,后代容易产生变异与不育,而且取材受季节限制,这是此系统应用于遗传转化的主要缺点。

2.4 减轻受体材料褐化的研究

农杆菌介导大豆转化时,常常会引起外植体的褐化,这种外植体的褐化可能是植物对受伤和病原体侵害的一种防御反应,而这种因农杆菌的侵染和制造伤口引起的褐化和细胞死亡可能会阻碍 T-DNA 进入植物细胞。2001 年 P. M. Olhoft[9]报道了在共培养培养基中加入 L-Cys 可以有效的减少共培养后大豆子叶节外植体的褐化现象,同时将农杆菌的感染率从 37% 提高至 91%。后来他们又发现[10]加入巯基混合物 L-Cys、DTT(二硫苏糖醇)和硫代硫酸钠对转化的促进作用更加明显,他们推测巯基混合物提高 T-DNA 的转移效率的机理可能是抑制了植物对创伤和病原体侵染发生褐化反应的酶的活性,例如过氧化物酶(PODs)和多酚氧化酶(PPOs)。Okpuzor 等[31]报道 DTT 能够完全抑制 POD 的活性,但对 PPO 的活性只能抑制 21%;而 L-Cys 对 PPO 的活性能抑制 60%,但对 POD 的活性只能抑制 30%。可能就是由于这种作用的不同使得巯基混合物促进 T-DNA 向植物细胞中转移和减少外植体褐化的效果好于单一的巯基化合物。

不同的组织和物种对创伤和病原体侵染的反应机理是有区别的,因此在共培养时加入巯基混合物对不同组织和物种的农杆菌介导的转化效率的影响也会有所不同[10]。目前研究仅限于农杆菌/子叶节系统,巯基混合物对农杆菌介导的大豆其他组织和细胞转化影响的研究尚未见报道。

3 筛选标记基因及其消除方法的研究

在植物遗传转化中,一般只有一部分细胞被真正转化,因此建立一个好的筛选系统,在有效去除非转化体干扰的同时保证转化体正常生长,是转化成功的关键。

3.1 筛选标记基因

筛选标记基因的主要功能是其编码的产物可以使转化体对某些选择剂产生抗性,使其在特定的培养基中具有生长

优势,从而被选择出来[32]。目前在大豆的遗传转化中报道的筛选标记基因主要有抗生素抗性基因(*npt* Ⅱ、*hpt*)和除草剂抗性基因(*bar* 基因和突变的 *ahas* 基因)。

新霉素磷酸转移酶基因 *npt* Ⅱ的筛选剂是卡那霉素,早期的大豆转化大多数用的是这一筛选系统[5],但是大豆对卡那霉素的敏感性不是特别强,所以现在应用较少。潮霉素磷酸转移酶基因 *hpt* 的筛选剂是潮霉素,用基因枪法转化大豆体细胞胚后,应用潮霉素作为选择剂已经成功的获得了可育的转基因植株[33,34]。2003 年 P.M.Olhoft 等[10]首次在农杆菌介导的大豆子叶节转化系统中应用潮霉素作为选择剂,发现潮霉素可以快速杀死非转化的细胞,有效地防止了非转化植株的产生。

1999 年 Zhang 等[7]首次将 *bar* 基因作为选择标记基因应用于农杆菌介导的大豆子叶节的转化,获得转基因植株。*bar* 基因是目前大豆转化中应用较多的选择标记基因。突变的 *ahas* 基因是从拟南芥中分离得到的,在 653bp 附近有一突变,编码的产物能引起植物对咪唑酮类除草剂分子的抗性。Aragão 等[35]使用该系统通过基因枪法转化成熟种子浸泡 24h 后的茎尖,直接获得了再生的非嵌合转基因植株。

3.2 筛选标记基因的去除

在植物遗传转化过程中筛选标记基因往往是必不可少的,但是当得到转基因植株以后,仍然存在于植物体内的筛选基因就没有任何用处了,而且还会带来生物安全性问题,因此去除筛选标记基因已成为转化研究的新的热点之一。目前在大豆转化中主要是通过共转化法来删除标记基因。共转化法,即把分别携带目的基因和筛选标记基因的两个 T-DNA,通过共转化导入植物细胞内,转化植株经过后代的有性阶段即可使目的基因和筛选标记基因分离,经筛选得到只含目的基因而无筛选标记基因的转化体。这一方法要求共转化的频率要高,并且目的基因和标记基因要整合在染色体的不连锁位点上[36]。利用农杆菌介导的共转化去除标记基因的方法有:利用分别带有目的基因和筛选标记基因的两个农杆菌菌株进行共转化,利用一个农杆菌携带分别含有目的基因和筛选标记基因的两个双元载体进行共转化和利用一个农杆菌携带一个含两段 T-DNA 的双元载体进行共转化。Xing A.等利用第三类方法得到去除标记基因的转基因大豆,F1 代标记基因的去除率可达 40%[37]。

4 展望

植物遗传转化研究的最终目的是对植物进行品质和性状的改良及功能基因的研究,转化效率的提高是实现这一目标的关键前提。经过多年的研究,农杆菌介导的大豆转化已经取得了很大的进展,但尚不能满足大豆育种和功能基因研究的需要。因此,仍需对大豆转化的受体系统和转化过程中的影响因素做进一步研究。转化后基因的表达调控也是大豆转化研究的重要内容。另外,随着植物基因工程的发展,生物安全性问题受到广泛重视,转化后标记基因的去除也成为大豆转化研究的热点之一。在设法去除标记基因的同时,积极寻找更加安全有效的筛选系统也是非常重要的。

REFERENCES(参考文献)

[1] Xu ZH(许智宏). Plant Biotechnology, Shanghai: Shanghai Scientific and Technical Publishers(上海科学技术出版社), 1998, p. 321

[2] Hinchee MAW, Connor-Ward DV, Newell CA *et al*. Production of transgenic soybean plants using *Agrobacterium*-mediated DNA transfer. *Biotechnology*, 1988, **6**: 915 - 922

[3] Townsend JA, Thomas LA. An improved method of *Agrobacterium*-mediated transformation of cultured soybean cells. *Patent*, 1993, WO94/02620

[4] Clemente TE, LaValle BJ, Howe AR *et al*. Progeny analysis of glyphosate-selected transgenic soybeans derived from *Agrobacterium*-mediated transformation. *Crop Sci*, 2000, **40**: 797 - 803

[5] Di R, Purcell V, Collins GB *et al*. Production of transgenic soybean lines expressing the bean pod mottle virus coat protein precursor gene. *Plant Cell Rep*, 1996, **15**: 746 - 775

[6] Xu XL(徐香玲), Gao J(高晶), Liu WH(刘伟华) *et al*. Studies on transferring B.t.k-delta endotoxin gene into soybean with Ti-plasmide primaril. *Soybean Science*(大豆科学), 1997, **16**(1): 6 - 11

[7] Zhang ZY, Xing A, Paul S *et al*. The use of glufosinate as a selective agent in *Agrobacterium*-mediated transformation of soybean. *Plant Cell Tiss Org Cult*, 1999, **56**: 37 - 46

[8] Danaldson PA, Simmonds DH. Susceptibility to *Agrobacterium tumefaciens* and cotyledonary node transformation in short-season soybean. *Plant Cell Rep*, 2000, **19**: 478 - 484

[9] Olthoft P M, Somer D A. L-cysteine increases *Agrobacterium*-mediated T-DNA delivery into soybean cotyledonary-node cells. *Plant Cell Rep*, 2001, **20**: 706 - 711

[10] Paula MO, Lex EF, Christopher MD *et al*. Efficient soybean transformation using hygromycin B selection in the cotyledonary-node method. *Planta*, 2003, **216**(5): 723 - 735

[11] Parrott WA, Hoffman LM, Hildebrand DF *et al*. Recovery of primary transformation of soybean. *Plant Cell Rep*, 1989, **7**: 615 - 617

[12] Yan B, Reddy M SS, Collins GB *et al*. *Agrobacteriym tumefaciens*-mediated transformation of soybean [*Glycine max*(L.)Merrill.] using immature zygotic cotyledon explants. *Plant Cell Rep*, 2000, **19**: 1090 - 1097

[13] Trick HN, Finer JJ. Sonication-assiated *Agrobacterium*-mediated transformation of soybean [*Glycine max*(*L.*) *Merrill*] embryogenic suspension. *Plant Cell Rep*, 1998, **17**: 482 - 488

[14] Santarem ER, Trick HN, Essig JS *et al*. Sonication-assisted *Agrobacterium*-mediated transformation of soybean immature cotyledons: optimization of transient expression. *Plant Cell Rep*, 1998, **17**: 752 - 759

[15] Deng XY(邓向阳), Wei ZM(卫志明). Technique of transformation of soybean. *Plant Physiology Communication*(植物生理学通讯), 1998, **34**(5): 381 - 387

[16] Harold NT, Randy DD, Eliane RS *et al*. Recent advances in soybean transformation. *Plant Tissue Culture and Biotechnology*, 1997, **3**: 9 - 26

[17] Wang CT(王翠亭), Wei ZM(卫志明). Genetic transformation of important Cereal Crops. *Plant Physiology Communication*(植物生理学通讯), 2001, **38**(2): 105 - 110

[18] Zhou SJ(周思军), Li XC(李希臣), Liu ZJ(刘昭军) *et al*. Optimization of the *Agrobacterium*-mediated transformation systems of soybean. *Journal of Northeast Agricultural University*(东北农业大学学报), 2001, **32**(4): 313 - 319

[19] Cheng TV, Saka T, Voqui-Dinh TH. Plant regeneration from soybean cotyledonary node segments in culture. *Plant Sci Lett*, 1980, **19**: 91 - 99

[20] Wright MS, Koehler SM, Hichee MA *et al*. Plant regeneration by organogenesis in *Glycine max*. *Plant Cell Rep*, 1986, **5**: 150 - 154

[21] Wright MS, Williams MH, Pierson PE *et al*. Initiation and propagation of *Glycine max* L. M.: Plants from tissue-cultured epicotyls. *Plant Cell Tiss Org Cult*, 1987, **8**: 83 - 90

[22] Kartha KK, Pahl K, Leung NL *et al*. Plant regeneration from meristems of grain legumes: soybean, cowpean, peanut, chickpea and bean. *Can J Bot*, 1981, **59**: 1671 - 1679

[23] Kaneda Y, Takel Y, Nishimura S *et al*. Combination of thidiazuron and basal media with low salt concentrations increases the frequency of shoot organogenesis in soybeans [*Glycine max* (L.)Merr]. *Plant Cell Rep*, 1997, **17**: 8 - 12

[24] Lazzeri PA, Hildebrand DF, Collins GB. A procedure for plant regeneration from immature cotyledon tissue of soybean. *Plant Mol Bio Rep*, 1985, **3**: 160 - 167

[25] Hartweck LM, Lazzeri PA, Cui D *et al*. Auxin-orientation effects on somatic embryogenesis from immature soybean cotyledons. *In Viro-cellular and Developmental Biology*, 1988, **24**(8): 821 - 828

[26] Christianson ML, Warnick DA, Carlson PS. A morphpgenetically competent soybean suspension culture. *Science*, 1983, **222**: 632 - 634

[27] Ranch JP, Oglesby L, Zielinski AC. Plant regeneration from embryo-derived tissue cultures of soybean. *In Vitro Cell Dev Bio*, 1985, **21**: 653 - 658

[28] Ranch JP, Oglesby L, Zielinski AC. Plant regeneration from tissue culture of soybean by somatic embryogenesis. In Vasil IK(ed). Cell Culture and Somatic Cell Genetics of Plants(Vil3). New York: Academic Press, 1986, pp.97 - 110

[29] Finer JJ. Apical proliferation of embryogenic tissue of soybean (*Glycine max L. Merr.*). *Plant Cell Rep*, 1988, **7**: 238 - 241

[30] Finer JJ, Nagasawa A. Development of an embryogenic suspension culture of soybean (*Glycine max Merrill.*). *Plant Cell Tiss Org Cult*, 1988, **15**: 125 - 136

[31] Okpuzor J, Omidiji O. Peroxidase-polyphenol oxidase association in Dioscorea esculenta. *Z Naturforsch*, 1998, **53C**: 957 - 960

[32] Yan XF(闫新甫). Transgenic Plant, Beijing: Science Press(科学出版社), 2003, pp.118

[33] Stewart CN, Adang MJ, JNAII *et al*. Genetic transformation, recovery and characterization of fertile soybean transgenic for asynthetic bacillusthuringiensis cryiac gene. *Plant Physiol*, 1996, **112**: 121 - 129

[34] Maughan PJ, Philip R, Cho MJ *et al*. Biolistic transformation expression and inheriface of bovine β-casein in soybean (*Glycine max*). *In Vitro Cell Dev Biol*, 1999, **35**: 344 - 349

[35] Aragão FJL, Sarokin L, Vianna GR *et al*. Selection of transgenic meristematic cells utilizing a herbicidal molecule results in the recovery of fertile transgenic soybean [*Glycine max* (L.) Merrill] plants at a high frequency. *Thero Appl Genet*, 2000, **101**: 1 - 6

[36] Brasileiro ACM, Aragão FJL. Marker genes for *in vitro* selection of transgenic plants. *J Plant Biotechnology*, 2001, **3**(3): 113 - 121

[37] Xing A, Zhang Z, Stato S *et al*. The use of the T-DNA binary system to derive marker-free transgenic soybeans. *In Vitro Cell and Dev Biol Plant*, 2000, **36**(6): 456 - 463

Genetic Transformation of Soybean Mediated by *Agrobacterium tumefaciens*

LIN Shu-Zhu[1] CAO Yue-Ping[1]* WEI Zhi-Ming[2]

[1] (*Agriculture and Biology School, Shanghai Jiaotong University, Shanghai* 201101, *China*)

[2] (*Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai* 200032, *China*)

Abstract Genetic transformation mediated by *Agrobacterium tumefaciens* is one of the most important methods in soybean transformation. And transgenic soybean has been obtained in many laboratories by this method, and different *Agrobacterium tumefaciens* stains and soybean cultivars were used. Because of the low efficiency of this method, further researches are necessary. Many factors, such as *Agrobacterium tumefaciens* stains, soybean cultivars, the efficiency of T-DNA transfer, tissue culture conditions and selection protocols, can affect the efficiency of the transformation. In this paper, recent advances in soybean transformation mediated by *Agrobacterium tumefaciens* were introduced. And some affecting factors to soybean transformation, such as the interactions between soybean cultivars and *Agrobacterium tumefaciens* stains, regeneration systems in transformation, the chemical inducer, acetosyringone, inducing expression of the *vir* gene and the thiol-containing compounds decreasing tissue browning of explants and increasing T-DNA delivery were also introduced. In addition, the selectable markers, involving *npt* Ⅱ, *hpt*, *bar* and mutant *aras genes*, and marker-free method of co-transformation were discussed. At the same time the emphases were probed into the further researches.

Key words *Agrobacterium tumefaciens*, soybean, genetic transformation

Received: 03-26-2004

This work was supported by National Natural Science Foundation of China (No.30370892, 30370790), and Science and Technology Committee of Shanghai (No.02DJ14047).

* Corresponding author. Tel: 86-21-64788330; Fax: 86-21-64788330; E-mail: yuepingcao@sjtu.edu.cn

Selection of Optimal Primers for TAIL-PCR in Identifying Ds Flanking Sequences from Ac/Ds Insertion Rice Lines

SUN Bing-Yao[1,2*] PIAO Hai-Long[2] PARK Sung-Han[2] HAN Chang-Deok[2]

[1] (*College of Life Sciences, Suzhou University, Suzhou 215006, China*)

[2] (*Plant Molecular Biology & Biotechnology Research Center, Gyeongsang National University, Jinju 660-701, Korea*)

Abstract Thermal asymmetric interlaced-PCR, as a PCR-based technique in identifying DNA fragments flanking known sequences, has obtained wide application in different organisms thereby greatly promotes the efficiency in reverse genetics. Unfortunately, in spite of the fact that TAIL-PCR technique has been expanded vastly and adopted in transposon mutagenesis in rice, a reliable, highly reproducible TAIL-PCR procedure especially for rice genomic DNA is still not available, mainly due to the complexity of rice genome and the lack of optimal primers for TAIL-PCR in rice. Given the current situation, we designed 12 specific primers corresponding to 3′ end or complimentary to 5′ end of Ds insertion, which constitute 32 sets, each with 3 specific primers for three rounds of TAIL-PCR, for screening the optimal combinations of Ds-specific primers. Based on the massive results from pilot experiments, two optimal sets of specific primers (Ds3L1/Ds3L2/Ds3S3 at 3′end; Ds5L1/Ds5S2/Ds5S3 at 5′end) were chosen, and used together with six arbitrary degenerate (AD) primers, respectively, to comparatively investigate the effects of arbitrary degenerate primers on the specificity of TAIL-PCR. Among the tested six AD primers, AD4 (5′-NTCAGSTWTSCWGWT-3′), possessing 128 fold degeneracy, was proved to be the most efficient for TAIL-PCR with rice genomic DNA. Moreover, our results also implied that long specific primers (36～40mers) in the primary reaction favored the TAIL-PCR by increasing specificity, and different AD primers led to significant differences in PCR amplification, presumably due to great difference in degeneracy. Our data may provide helpful information for TAIL-PCR technique to improve the efficiency in identifying DNA fragments flanking Ds insertion in rice or other organisms.

Key words Arbitrary degenerate (AD) primers, Ds-specific primers, Ds flanking sequences, *Oryza sativa* L. (rice), TAIL-PCR

A variety of methods including partial library construction, inverse PCR (IPCR), and thermal asymmetric interlaced PCR (TAIL-PCR) can be used to clone unknown DNA sequences flanking known sequences, but the method of choice is TAIL-PCR[1]. Since it was developed in 1995[2], TAIL-PCR has become a powerful and efficient procedure to obtain the DNA fragments adjacent to known sequences, and has received wide application in isolating insert end segments from P1 and YAC clones[2], or cloning pathogenicity-related gene sequences[3], and in reverse genetics involving in higher plants[4,5,6]. The widespread use of TAIL-PCR greatly attributed to its distinct advantages over other traditional methods, such as its obvious simplicity (without manipulations of restriction digestion, ligation, etc. required in other methods), high specificity and efficiency. TAIL-PCR is a tri-reaction technique for *in vitro* DNA amplification, in which three sequential PCR reactions (primary, secondary, and tertiary reactions) using nested insert-specific primers along with a single degenerate primer are performed. The specific target products can usually be obtained after one round of super-PCR followed by two rounds of nested PCR. Ac/Ds system, a two- component transposon system based on the Ac element of maize, was proved to be an insertional mutagen in the heterologous host *Arabidopsis*[7]. In the case of screening rice mutants from Ds-tagging lines in our laboratory, this procedure was utilized as a conventional method to identify the transposon (maize Ac/Ds system) insertion sites in rice genome[8,10]. According to our experience in manipulating in TAIL-PCR with rice genomic DNA, this method is still exposed to failure due to unknown sources. When it was adopted directly to analyze the Ds-flanking fragment from *Oryza sativa* genomic DNA, a very high background

Received: 03-15-2004

This work was supported by a grant (No. CG1613) from Crop Functional Genomics Center of the 21st Century Frontier Research Program, Republic of Korea, and in part by a grant from the National Natural Science of China (No.30170556) and by the Natural Science Foundation (No.BK2001139) from Science and Technology Department of Jiangsu Province, China. Also, the work of Dr. Han's laboratory was supported by the Brain Korea 21 Project.

* Corresponding author. Tel: 86-512-65226781; E-mail: sunbingyao@suda.edu.cn

from smear products in agarose gel electrophoresis was detected, and sometimes it was difficult to identify the specific target products. So, a reliable, reproducible TAIL-PCR procedure especially for rice genomic DNA is not available. Apart from the complexity of rice genomic DNA, utilization of specific and arbitrary degenerate primers in TAIL-PCR may greatly owe to the TAIL-PCR performance[11]. Therefore, in this study we investigated the effects of primers on the TAIL-PCR performance with the objective to improve the reproducibility and efficiency of TAIL-PCR for identifying Ds-flanking sequences from rice genome.

1 Materials and Methods

1.1 Plant materials and templates for TAIL-PCR

Genomic DNA of *Oryza sativa* L. cv. Dongjin Bye, a japonica rice cultivar, was extracted from leaf tissues using a urea extraction procedure[8]. After extraction, the concentration and purification of genomic DNA were checked by measuring the absorbance of a sample at 260nm as well as 280nm on a spectrophotometer.

Rice genomic DNA containing both Ac and Ds was screened by PCR using a pair of specific primers, Ac1031 (5′-ATAAGATTGGCCAAGTTGATGTC-3′) and Ac1543a (5′-TTCTTGGTGAAATGCTGCCATAC-3′), and used as templates in TAIL-PCR.

1.2 Primers tested for TAIL-PCR in this study

Arbitrary degenerate (AD) primers Six arbitrary degenerate primers with different degeneracies (ranging from 32 to 256) were chosen for this study, which were summarized in Table 1.

Table 1 Arbitrary degenerate (AD) primers used in this study

AD primer	Sequence (5′-3′)	Degeneracy (X)
AD1GTG	ASNTGSWATGG	32
AD2	TTGIAGNACIANAGG	64
AD3	GWWGGTSCWASWCTG	64
AD4	NTCAGSTWTSGWGWT	128
AD5	GWSIDRAMSCTGCTC	192
AD6	AGWGNAGWANCANAGA	256

Notes: Mixed bases were present in IUB code name, D = G, A, T; M = A, C; N = A, G, C, T; R = A, G; S = G, C; W = A, T. I stands for Inosine. The sequence information on arbitrary degenerate primers was obtained from Tsugeki *et al*[9], and from personal communication with U. Grossniklaus (University of Zürich, Switzerland), and D. McCarty (University of Florida, USA).

Ds-specific primers: In order to evaluate the possible influences of specific primers and of oligonucleotide length of primers on the TAIL-PCR, twelve Ds-specific primers were designed, with 3 short primers (19 ~ 24mer) and 3 long primers (34 ~ 40mer) for each end of Ds insertion. All the twelve Ds-specific primers used in this study were marked in Figure 1. Because TAIL-PCR consists of three PCR reactions, which were carried out sequentially, these primers were combined to constitute totally 32 sets of specific primers (Table 2), and each set contained 3 specific primers for using in three rounds of TAIL-PCR, respectively. Totally, there were 12 and 20 sets of specific primers at 3′ end and 5′ end of Ds insertion, respectively.

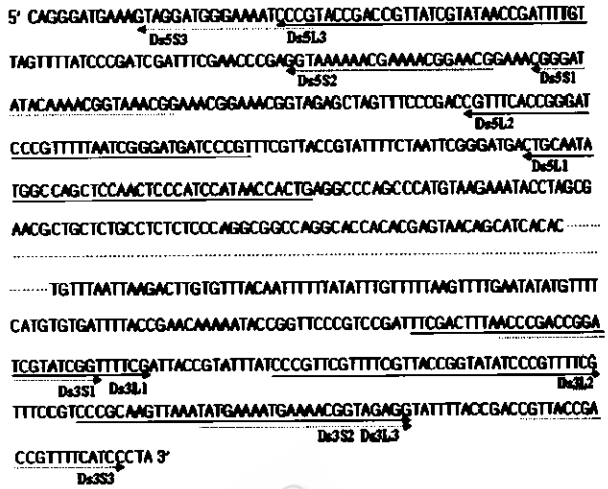

Fig.1 The corresponding positions of all the 12 specific primers tested for TAIL-PCR at both ends of Ds insertion

Specific primers corresponding to the 3′ end or complimentary to 5′ end of Ds insertion were marked by arrow underlines, with dashed lines for short primers and thick lines for long primers

Table 2 All the 32 sets of specific primers tested in the three successive rounds of TAIL-PCR

Primer set	1st/2nd/3rd reactions	Size difference of shift bands*
1	Ds3L1/Ds3L2/Ds3L3	44
2	Ds3L1/Ds3L2/ Ds3S2	57
3	Ds3S1/Ds3L2/Ds3L3	44
4	Ds3S1/ Ds3L2/ Ds3S2	57
5	Ds3L1/ Ds3L2/ Ds3S3	93
6	Ds3S1/ Ds3L2/ Ds3S3	93
7	Ds3L1/ Ds3L3/ Ds3S3	49
8	Ds3L1/ Ds3S2/Ds3S3	36
9	Ds3S1/ Ds3L3/ Ds3S3	49
10	Ds3S1/Ds3S2/Ds3S3	36
11	Ds3L2/Ds3L3/ Ds3S3	49
12	Ds3L2/ Ds3S2/Ds3S3	36
13	Ds5L1/Ds5L2/Ds5S1	70
14	Ds5L1/Ds5L2/Ds5S2	88
15	Ds5L1/Ds5L2/Ds5L3	150
16	Ds5L1/Ds5L2/Ds5S3	180
17	Ds5L1/Ds5S1/Ds5S2	18
18	Ds5L1/Ds5S1/Ds5L3	80

Primer set	$1^{st}/2^{nd}/3^{rd}$ reactions	Size difference of shift bands*
19	Ds5L1/Ds5L2/Ds5L3	110
20	Ds5L1/Ds5L2/Ds5L3	62
21	Ds5L1/Ds5L2/Ds5L3	92
22	Ds5L1/Ds5L2/Ds5L3	30
23	Ds5L2/Ds5S1/Ds5S2	18
24	Ds5L2/Ds5S1/Ds5L3	80
25	Ds5L2/Ds5S1/Ds5S3	110
26	Ds5L2/Ds5S2/Ds5L3	62
27	Ds5L2/Ds5S2/Ds5S3	92
28	Ds5L2/Ds5L3/Ds5S3	30
29	Ds5S1/Ds5S2/Ds5L3	62
30	Ds5S1/Ds5S2/Ds5S3	92
31	Ds5S1/Ds5L3/Ds5S3	30
32	Ds5S2/Ds5L3/Ds5S3	30

Notes: Sets 1 ~ 12 are specific primers corresponding to 3′ end of Ds insertion, whereas sets 13 ~ 32 are complimentary to 5′ end of Ds insertion. *: expected size difference between specific products from the 2nd and 3rd reactions calculated according to the distance from 5′ end of the primer to the corresponding end of Ds.

1.3 Cycling conditions of TAIL-PCR

Thermal asymmetric interlaced PCR (TAIL-PCR) was performed in the PTC-200 Peltier Thermal Cycler (MJ Research, Watertown, MA, USA) basically according to the original protocol by Liu & Huang[12] with the minor modification that running the low stringency cycle at 35℃ instead of 25℃ in the primary reaction (20μL), 15 supercycles in the secondary reaction (20μL) and 35 reduced - stringency cycles in the tertiary reaction (30μL) were performed. 20ng rice genomic DNA was used as a template in the 1st reaction. rTaq (Takara Shuzo, Japan) or NeoTherm (GeneCraft, Germany) was used for Taq polymerase in all the three consecutive rounds of TAIL-PCR.

1.4 Agarose gel analysis of TAIL-PCR products

The amplified products were analyzed by 2% agarose gel electrophoresis in 1 × TAE buffer (40mol/L Tris-acetate, 2mmol/L EDTA [pH 8.3]). The Mupid-21 apparatus (COSMO BIO CO, LTD) was set at 50 V for approximately 40min. After electrophoresis, the gels were stained with 0.5 μg of ethidium bromide per mL for 30 min and destained with water for 20 min. To record the amplification pattern, the gels were placed on a UV light box and the pictures of the fluorescent ethidium bromide-stained DNA separation were taken with a Polaroid camera. Since specific products in the 1st reaction may not reach a visible level in agarose gels[12], only amplified products from the 2nd and 3rd reactions were analyzed in this study.

2 Results

Based on the results from two rounds of pilot experiments (data not shown), which were carried out to determine the appropriate specific primers, we investigated the possible effects of different arbitrary degenerate (AD) primers on the performance of TAIL-PCR. As shown in Figure 2, when combined with specific primers at 3′end (Figure 2A) or 5′ end of Ds insertion (Figure 2B), different AD primers produced quite different amplification patterns of TAIL-PCR. Comparatively, it is easy to find that AD4 provided the highest efficiency in TAIL-PCR as evidenced by the strong bands amplified

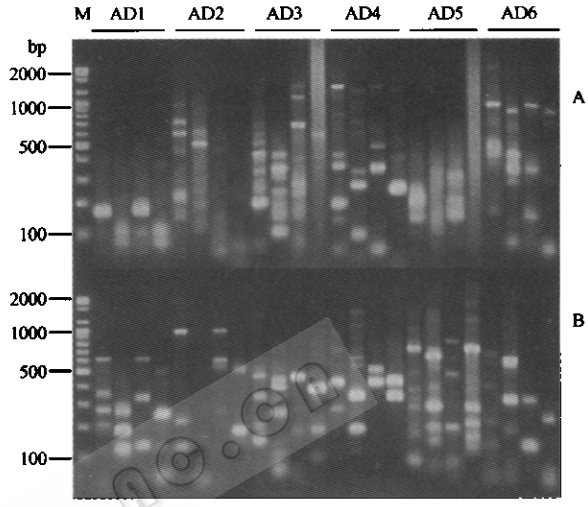

Fig.2 Agarose gel analysis of amplified products by TAIL-PCR using six different arbitrary degenerate (AD) primers

Two samples (genomic DNA) were used as templates for each AD primer, and each pair of lanes contains amplified products from the 2nd and 3rd reactions. DNA size markers (lane M) are indicated on the left. A) AD primers together with 3′ end specific primers (set 5 in Table 2); B) AD primers together with 5′ end specific primers (set 21 in Table 2)

from most samples (here, data from two samples are presented in Figure 2). Alternatively, AD3 and AD6 can be chosen for TAIL-PCR with rice genomic DNA. Both AD1 and AD5 failed to yield any specific products when they were used together with specific primers at 3′ end of Ds insertion, but they can be used in combination with 5′ end specific primers. Due to the considerable complexity of AD primers, the close relationship between degeneracy of AD primers and its effects on TAIL-PCR can not be established. Presumably, lower degeneracy is highly related to lower efficiency of TAIL-PCR. According to the data in this study, AD primers with lower degeneracy didn't provide the results as good as those from AD primer with higher degeneracy, though AD5 (with 192 fold degeneracy) yielded smear products when it was used together with 3′ end specific primers. One possible explanation for AD5's failure in combination with 3′ end specific primers is that the amplification performance may be affected not only by its degeneracy but also by its oligonucleotide sequence. Among the six AD primers tested, AD5 possesses the highest ratio of degenerate base to particular base, with 50% bases degenerated, which may closely related to the efficiency of TAIL-PCR.

AD4, the most optimal arbitrary degenerate (AD) primer as

indicated above, was employed to perform TAIL-PCR together with specific primers with an objective to select the optimal Ds-specific primers. Together with specific primers at 3′ end or 5′ end of Ds insertion, respectively, AD4 was exploited to perform the TAIL-PCR, and the resulting amplification patterns are shown in Figure 3 and Figure 4.

As far as specific primers at 3′ end are concerned, in comparison with other sets of specific primer, sets 1, 2, 5, 7 and 8 have clearly displayed strong specific bands as confirmed by the shift distance between the 2nd and its corresponding 3rd products (Figure 3), which unexceptionally

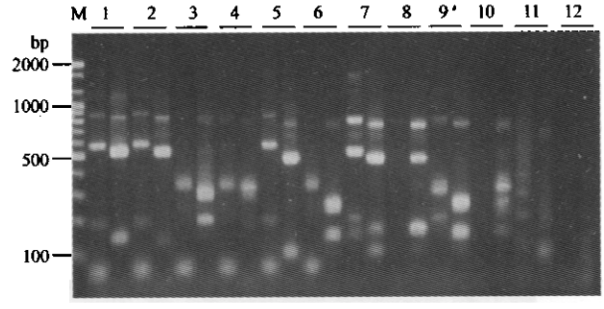

Fig. 3 Agarose gel analysis of amplified products by TAIL-PCR using AD4 (AD primer) together with specific primers at 3′ end of Ds insertion

For testing, 12 sets (a pair of lanes for each set) of specific primers were examined. The combinations of specific primers are designated in Table 2. The 2nd and 3rd products were loaded side by side to check the possible shift bands

generated from Ds3L1 in the 1st reaction. For sets 11 and 12 using long primer Ds3L2 in the 1st round, no obvious specific products were observed, which was even worse than those from sets 3, 4, 6, 9 and 10, in all of which short primer Ds3S1 was used in the 1st reaction. To disclose the reasons for the above results, we carried out a database search for short, nearly exact matches using the sequence of specific primer as a query against NCBI-BLAST [http://www.ncbi.nlm.nih.gov/BLAST/]. These results can be well explained by the fact that Ds3L1 matches only 2 hits when NCBI search was done against short, nearly exact sequences among rice genome, while Ds3L2 holds 10 matches, even more than 7 hits of Ds3S1. Collectively, results present here showed that the nucleotide sequence of specific primer can be regarded as the key factor influencing PCR specificity, and using long primers in the 1st reaction significantly favored TAIL-PCR. As for the 2nd and 3rd reactions, no significant difference exhibited in using short or long primers.

Compared with the amplification patterns from specific primers at 3′ end of Ds (as shown in Figure 3), each set of specific primers at 5′ end of Ds insertion yielded stronger specific products (Figure 4). As expected, the distance between the 2nd and its corresponding 3rd product from each set of 5′ end primers is consistent with the calculated size difference (Table 2), which indicated that all the strong bands were specific targeted products.

TAIL-PCR products from set 13 to set 22 were generated by specific primer Ds5L1 in the 1st reaction (shown in the upper panel of Figure 4). It can be found that most of the products are 300 ~ 500 bp in size, while other products with bigger size are very weak in gel. Sets 23 ~ 28 using Ds5L2 in the 1st reaction (shown in the lower panel of Figure 4) also yielded highly specific products, with only two strong bands present, while sets 29 ~ 32, in which short primers Ds5S1 or Ds5S2 were adopted, recovered more than three bands, which implied the lower specificity.

Interestingly, by comparing the set 9 (Figure 3) with set 31 (Figure 4), in both of which short primers were used in the 1st reaction, and long primers in the 2nd reaction, respectively, it was found that these two sets gave rise to quite different amplification patterns, though both primers (Ds3S1 in set 9 and Ds5S1 in set 31) have the same hit number of nearly exact matches with rice genomic DNA, and similar distance from their corresponding end of Ds. Nevertheless, if sequencing of shift bands was taken into consideration, set 31 is appreciated relatively, because it recovered PCR products appropriate in size for sequencing.

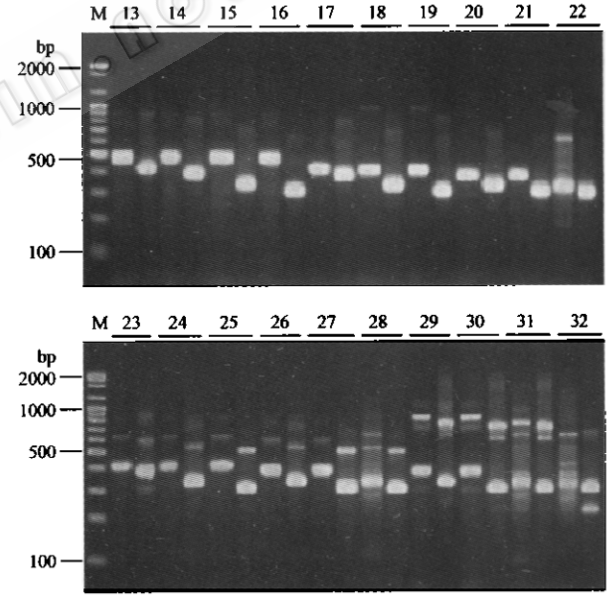

Fig. 4 Agarose gel analysis of amplified products by TAIL-PCR using AD4 (AD primer) together with specific primers at 5′ end of Ds insertion

For testing, 20 sets (13 ~ 32) of specific primers were examined. The combinations are signed in Table 2. The 2nd and 3rd products were loaded side by side to check the possible shift bands

3 Discussion

As a powerful technique to identify the DNA fragments flanking a known sequence, TAIL-PCR has obtained a wide range of applications in different higher plants including rice[6,8]. Originally, TAIL-PCR was developed and used successfully for chromosome walking in YAC clones[2,12]. To increase its efficiency in plant

reverse genetics, different researchers have given some modifications to TAIL-PCR. Sessions *et al*[4] developed a two-reaction mTAIL-PCR for Arabidopsis reverse genetics from the original version of TAIL-PCR[2]. With the objective to efficiently identify the DNA fragments flanking Ds insertion from rice genome on a large scale, we think it is necessary to optimize TAIL-PCR reaction conditions. Though a number of parameters can be varied to optimize reaction conditions for PCR including TAIL-PCR, the most crucial factor influencing TAIL-PCR is the utilization of primers. In this study, based on many pilot experiments, we fixed most parameters at the standard levels[2] that are suitable for successful application in other organisms, and concentrated on the effects of primers on the performance of TAIL-PCR amplification, thereby screened the optimal specific primers and AD primers.

Successful TAIL-PCR depends tremendously on the utilization of optimal primers, though effects of other cycling conditions can not be ruled out. Among the three consecutive reactions in TAIL-PCR, the primary reaction is the most critical because relative amplification efficiencies of specific and non-specific products is thermally controlled in this step, in which the specific product can be preferentially amplified over nonspecific ones by interspersion of two high-stringency PCR cycles with one reduced-stringency PCR cycle[12]. Two factors related to the specific primer used in 1st reaction may have played a crucial role in the performance of TAIL-PCR. One is the oligonucleotide length, the other is the base composition. As shown in Figure 4, when long sequence primer (40mer) was used, TAIL-PCR can produce highly specific products, whereas the specificity may decrease obviously when short primer (22-24mer) was used. As for the sequence of primer, we found that the background of primer sequence against the whole rice genome was another major factor for the success of TAIL-PCR. To support this, we run a NCBI BLAST search for short, nearly exact matches using primer sequence as a query. The search results indicated that primers Ds5L1 and Ds5L2 have 3 and 21 hits against rice sequence, respectively, which can explain the results in Figure 4. Most often, more hits mean more sequences similar to the primer, and imply noisy DNA, which lead to the decrease in specificity of amplification.

Though primer selection is often empirical, some of the criteria should be considered to select optimal specific primer sets for TAIL-PCR, which include (a) higher specificity with less non-specific products, (b) stronger of specific bands in gel, (c) size of PCR products around 500 bp. These rules will apparently favor the subsequent manipulations of eluting and sequencing. Only those specific primer sets that fulfill the above criteria can be regarded as optimal for TAIL-PCR. In this study, we screened the two most optimal specific primer sets, set 5 and set 21, for TAIL-PCR, which are corresponding to 3′ end or complimentary to 5′ end of Ds insertion, respectively. Obviously, one common feature for both sets is long primer in the 1st reaction. Moreover, though there was no direct evidence of positioning effect of specific primer, the results suggested specific primers at 5′ end of Ds generated better amplification than those at 3′ end did.

Our data turn out that AD4 (5′-NTCAGSTWTSGWGWT-3′) is the most optimal AD primer for TAIL-PCR using Ds-tagged rice genomic DNA as a template, because it can cover most samples at each manipulation. Understandably, even the optimal AD primers can not always work for every sample, so, when no specific products are recovered, using another arbitrary degenerate primer (such as AD 3) in place of AD 4 is recommended. It should be pointed out that two inosine-containing AD primers tested in this study, AD2 and AD5 didn't show better results, though inosine was usually synthesized in AD primers in other research work[13].

Undoubtedly, the optimal primers screened in this study will increase specificity of TAIL-PCR and thereby facilitate reliable application of TAIL-PCR in rice reverse genetics.

REFERENCES

[1] Campisi L, Yang Y, Yi Y *et al*. Generation of enhancer trap lines in *Arabidopsis* and characterization of expression patterns in the inflorescence. *Plant J*, 1999, **17**(6):699 - 707

[2] Liu YG, Whittier RF. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, 1995, **25**(3):674 - 681

[3] Ying G(应 革), Wu W(武 威), He CZ(何朝族). Cloning of *Xanthomonas campestris* cv. *campestris* pathogenicity-related gene sequences by TAIL-PCR. *Chinese Journal of Biotechnology*(生物工程学报), 2002, **18**(2):182 - 186

[4] Singer T, Burke E. High-throughput TAIL-PCR as a tool to identify DNA flanking insertions. *Methods Mol Biol*, 2003, **236**:241 - 272

[5] Sessions A, Burke E, Presting G *et al*. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell*, 2002, **14**(12): 2985 - 2994

[6] Sha Y, Li S, Pei Z *et al*. Generation and flanking sequence analysis of a rice T-DNA tagged population. *Theor Appl Genet*, 2004, **108** (2):306 - 314

[7] Long D, Martin M, Sundberg E *et al*. The maize transposable element system Ac/Ds as a mutagen in *Arabidopsis*: identification of an albino mutation induced by Ds insertion. *Proc Natl Acad Sci U S A*. 1993, **90**(21):10370 - 10374

[8] Chin HG, Choe MS, Lee SH *et al*. Molecular analysis of rice plants harboring an Ac/Ds transposable element-mediated gene trapping system. *Plant J*, 1999, **19**(5): 615 - 624

[9] Tsugeki R, Kochieva EZ, Fedoroff NV. A transposon insertion in the *Arabidopsis SSR*16 gene causes an embryo-defective lethal mutation. *Plant J*, 1996, **10**(3):479 - 489

[10] Kim CM, Je BI, Piao HL *et al*. Reprogramming of the activity of Ac/Ds transposon family during plant regeneration in rice. *Mol Cell*, 2002, **14**(2):231 - 237

[11] Tong L, Nakashima S, Shibasaka M *et al*. A novel histidine-rich CPx-ATPase from the filamentous cyanobacterium *Oscillatoria brevis* related to multiple-heavy-metal cotolerance. *J Bacteriol*. 2002, **184** (18): 5027 - 5035

[12] Liu YG, Huang N. Efficient amplification of insert end sequences from bacterial artificial chromosome clones by thermal asymmetric interlaced PCR. *Plant Mol Biol Rept*, 1998, 16: 175 - 181

[13] Fordham-Skelton AP, Yarwood A, Croy RR. Synthesis of saporin gene probes from partial protein sequence data: use of inosine-oligonucleotides, genomic DNA and the polymerase chain reaction. *Mol Gen Genet*, 1990, 221(1): 134 - 138