

# 大豆球蛋白 G1 启动子的克隆 及植物表达载体构建

王永芹, 陈德富, 王绘砖, 陈喜文\*

(南开大学 生命科学学院, 天津 300071)

**摘要:** 从大豆冀 nf37 和冀豆 15 中克隆了大豆球蛋白 G1 基因的启动子片段。序列分析表明, 两种启动子片段均为 688 bp, 与 GenBank 现有的 3 种启动子序列(四川大豆(DQ250808)、南农 87-c38(AY649096)和 Dare(X15121))间的同源性在 96.4%~99.6%之间。其中来自冀 nf37 的启动子片段除 Legumin 盒上有一个碱基差异外, 其它元件与 DQ250808 完全相同, 据此推测该启动子片段具有种子特异性启动子活性。将其与已有  $\gamma$ -生育酚甲基转移酶基因连接, 构建了种子特异性表达载体 pBG1TMT, 为通过代谢工程手段调控油料作物种子维生素 E 组成、提高其营养品质奠定了基础。

**关键词:** 大豆; 球蛋白 G1; 启动子;  $\gamma$ -生育酚甲基转移酶( $\gamma$ -TMT); 种子特异性表达载体

**中图分类号:** Q943 **文献标识码:** A **文章编号:** 1000-3142(2009)02-0254-06

## Cloning of soybean glycinin G1 promoter and construction of plant expression vector

WANG Yong-Qin, CHEN De-Fu, WANG Hui-Zhuan, CHEN Xi-Wen\*

(College of Life Sciences, Nankai University, Tianjin 300071, China)

**Abstract:** A fragment of glycinin G1 promoter was amplified from soybean cultivars Ji nf37 and Jidou 15, respectively. Sequence analysis indicated that both the fragments are 688 bp and had 96.4%–99.6% homologies compared with the registered sequences in GenBank (Sichuan(DQ250808), Nannong87-c38(AY649096) and Dare(X15121)). The promoter fragment from Ji nf37 contains the completely identical elements for seed-specific expression with DQ250808 except only one base difference existing in the Legumin box, which was supposed to have the activity of seed-specific promoter. It was recombined with the  $\gamma$ -TMT gene and plant expression vector pBG1TMT was constructed. The work laid foundation for metabolic engineering to increase  $\alpha$ -tocopherol level in oilseed crop seeds.

**Key words:** soybean; glycinin G1; promoter;  $\gamma$ -tocopherol methyltransferase( $\gamma$ -TMT); seed-specific expression vector

维生素 E 是 8 种生育酚类化合物的总称, 仅由植物和某些光合细菌所合成, 为维持人体正常生理功能所必需(Fryer, 1992)。临床上可用来治疗与预防高血压、冠心病、心肌梗塞、动脉硬化、血栓、不孕症、癌症等疾病(Rimm 等, 1993; Stampfer 等, 1993; Buring & Hennekens, 1997)。8 种维生素 E 中,  $\alpha$ -

生育酚活性最高。 $\gamma$ -生育酚甲基转移酶( $\gamma$ -TMT)催化  $\gamma$ -生育酚向  $\alpha$ -生育酚的转化(Eckardt, 2003; Karin, 2003), 但它在种子中的活性很低, 从而导致油料种子(如大豆)中  $\alpha$ -生育酚仅占生育酚总量的 7%~10%, 而其前体  $\gamma$ -生育酚却高达 60%~70%(Grusak & Dellapenna, 1999)。因此, 提高油料种

收稿日期: 2007-08-13 修回日期: 2007-10-29

基金项目: 天津市自然科学基金重点基金(07JCZDJC03800)[Supported by Key Project of Municipal Natural Science Foundation of Tianjin(07JCZDJC03800)]

作者简介: 王永芹(1981-), 女, 山东临沂人, 硕士研究生, 研究方向为生物化学与分子生物学, (E-mail)yongqin129@tom.com.

\* 通讯作者(Author for correspondence, E-mail: xiwenchen@nankai.edu.cn)

子中  $\alpha$ -生育酚含量已成为维生素 E 代谢工程的重要方向。为避免非特异性启动子驱动外源基因持续表达对植物正常生理活动产生干扰, 种子特异性启动子的使用已成为改变植物种子成分的代谢工程研究的重要环节。

大豆球蛋白(glycinin)是大豆种子中主要贮藏蛋白之一, 占种子干重的 20% 左右, 它只在种子形成的胚乳期表达(Baumlein 等, 1992; Lelievre 等, 1992)。目前已知的 6 种大豆球蛋白 G1、G2、G3、G4、G5 和 G7 中, G1 基因的启动子能在种子中启动外源基因表达, 且启动子活性相对较高(Ding 等, 2006)。在从甘蓝型油菜中克隆出  $\gamma$ -TMT 全长 cDNA 序列并验证其功能的基础上(钱文成等, 2007), 本文从高油大豆冀 nf37 克隆了大豆球蛋白 G1 基因启动子片段, 并构建了 G1 启动子驱动  $\gamma$ -TMT 的表达载体, 为通过代谢工程手段调控作物种子中维生素 E 的组成奠定基础。

## 1 材料与方法

### 1.1 材料

1.1.1 植物材料 大豆(*Glycine max*)高油品种“冀 nf 37”和高蛋白品种“冀豆 15”由河北省农林科学院粮油作物研究所王文秀先生惠赠。种子播种于  $25 \pm 3$  °C、15 h 光照/9 h 黑暗、光照强度  $54 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$  的人工气候室中。取 4~5 叶龄期幼苗叶片提取总基因组 DNA。

1.1.2 菌株和质粒 大肠杆菌 DH5 $\alpha$ -FT、植物表达载体 pBin438(含 35S 双启动子)、甘蓝型油菜全长 cDNAs(Chen 等, 2004)为我室保存。pMD18-T 载体购自大连宝生物工程有限公司。

### 1.2 方法

1.2.1  $\gamma$ -TMT 基因的扩增 在获得甘蓝型油菜  $\gamma$ -TMT 基因序列(GenBank 注册号为 DQ508019)的基础上, 设计引入限制性内切酶位点的引物 TMT-P ( $5'$  aaactgcagatgaaagcgactctgcacc- $3'$ , 下划线为 *Pst* I 酶切位点)和 TMT-E4 ( $5'$ -acgcgtcgacttagagag-gtttctggcaagtgatg- $3'$ , 下划线为 *Sal* I 酶切位点), 以甘蓝型油菜全长 cDNA 为模板扩增  $\gamma$ -TMT 全长基因。PCR 体系为 15  $\mu\text{L}$ , 包含 3 nmol dNTPs、1.9 pmol 两种引物、0.375 U Blend *Taq* DNA 聚合酶和 2.0 ng 模板 DNA。PCR 条件为 94 °C 5 min; 94 °C 40 s; 58 °C 32 s; 72 °C 100 s; 40 个循环; 72 °C 7

min。扩增产物经琼脂糖凝胶电泳回收纯化后, 16 °C 下与 pMD18-T 过夜连接, 再转化进 DH5 $\alpha$ -FT, 然后涂布在含 IPTG、X-gal 和氨苄青霉素 (50  $\mu\text{g}/\text{mL}$ ) 的 LB 筛选平板上, 过夜培养。挑选白色转化子, 碱法小量提取质粒进行质粒 PCR 筛选, 阳性转化子再经限制性内切酶分析, 以检查插入片段大小及方向, 正向插入的克隆命名为 pMD18-T-TMT。

1.2.2 大豆种子特异性启动子片段的克隆 根据 GenBank 现有大豆球蛋白 G1 基因的启动子序列, 设计和合成引物 G1-F ( $5'$ -ccccaaagcttagcctaagtaag-tactcaaaatgcc- $3'$ , 下划线为 *Hind* III 酶切位点)和 G1-R ( $5'$ -aaaactgcaggggtgatgactgatgtgtaagg- $3'$ , 下划线为 *Pst* I 酶切位点)。用 CTAB 法(Saghai-Maroof 等, 1984)提取大豆叶片基因组 DNA, 以此为模板扩增 G1 基因启动子片段。PCR 体系为 15  $\mu\text{L}$ , 包含 3 nmol dNTPs、1.9 pmol 两种引物、0.375 U Blend *Taq* DNA 聚合酶和 50 ng 基因组 DNA。PCR 条件同上。扩增产物同样经琼脂糖凝胶电泳回收纯化后, 用 *Hind* III/*Pst* I 双酶切, 再与同样经 *Hind* III/*Pst* I 双酶切的 pMD18-T-TMT 重组质粒连接, 转化 DH5 $\alpha$ -FT 感受态细胞。阳性转化子的筛选方法同上, 并命名为 pMD18-T-G1-TMT, 送北京三博远志公司测序。利用软件 Lasergene Suite MegAlign(ver 7.0)的 Clustal V 方法进行序列比对。

1.2.3 种子特异性表达载体 pBG1TMT 的构建 用 *Hind* III 与 *Sal* I 双酶切重组质粒 pMD18-T-G1-TMT, 回收约 1.7 kb 片段(G1-TMT)。用同样的限制性内切酶双酶切植物表达载体 pBin438, 回收大片段, 以切除其上的  $\Omega$  增强子和 CaMV35S 双启动子。将 G1-TMT 片段与 pBin438 大片段连接, 转化进 DH5 $\alpha$ -FT, 涂布到含 20 mg/L 卡那霉素的 LB 筛选平板上。PCR 及酶切鉴定的阳性重组子命名为 pBG1TMT。

## 2 结果与分析

### 2.1 pMD18-T-TMT 载体的构建

以油菜全长 cDNA 为模板, TMT-P 和 TMT-E4 为引物, 扩增出一长约 1.0 kb 的片段。将该片段连接到 pMD18-T 上, 转化进 DH5 $\alpha$ -FT。菌落 PCR 法初步筛选出阳性转化子, 提取质粒 DNA, 进行限制性内切酶酶切分析。用 *Sal* I 酶切发现, 重组子 1、2、3、4、6、7 产生 1.0 kb 和 2.7 kb 两片段,

为正向插入;重组子 5 产生一个 3.7 kb 片段,为反向插入(图 1: A)。筛选出正向插入的重组子,用 *Hind* III/*Pst* I 双酶切确证,均产生一个 3.7 kb 片段,确证为正向插入的重组子(图 1: B)。筛选出正

向插入重组子,命名为 pMD18-T-TMT。

### 2.2 种子特异性启动子 G1 片段的克隆

提取大豆基因组 DNA,电泳分析显示 DNA 呈单一条带,分子量大于 21 kb;*Eco*RI 酶切呈连续片

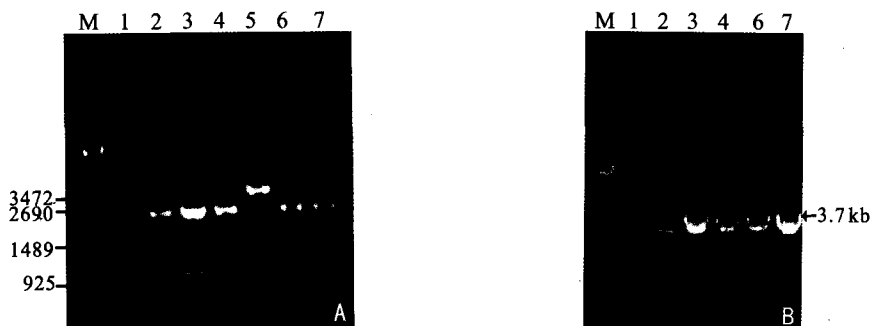


图 1  $\gamma$ -TMT 基因插入重组子的限制性内切酶分析

Fig. 1 Restriction analysis of recombinant plasmid pMD18-T-TMT

M: DNA 分子量标准(bp); A. *Sal* I 酶切; 1-7: 重组子编号; B. *Hind* III/*Pst* I 双酶切, 重组子编号同 A。

M:  $\lambda$ /*Eco*T14 I digest marker(bp); A. digested by *Sal* I; 1-4 and 6-7:  $\gamma$ -TMT inserted in forward orientation into pMD18-T; 5:  $\gamma$ -TMT inserted in reverse orientation into pMD18-T; B. digested by *Hind* III/*Pst* I. The number of recombinant lines is the same as in Fig. 1 A.

段,说明提取的 DNA 是完整的、纯度高,可用作启动子扩增模板。然后以 G1-F 和 G1-R 为引物,扩增 G1 启动子片段,电泳结果见图 2,扩增片段长约 700 bp,与预期大小相符。

0.7 kb 和 3.7 kb 两片段,证明从冀 nf37 和冀豆 15 中克隆的 G1 启动子均插入到 pMD18-T-TMT 中。阳性重组子命名为 pMD18-T-G1-TMT。

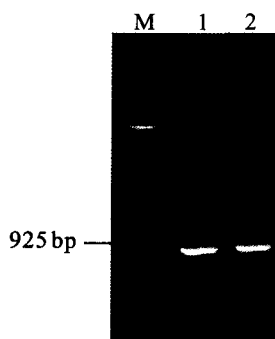


图 2 大豆 G1 启动子的 PCR 扩增

Fig. 2 PCR amplification of soybean glycinin G1 promoter

M: DNA 分子量标准(bp); 1: 冀 nf37; 2: 冀豆 15。

M:  $\lambda$ /*Eco*T14 I digest Molecular Marker (bp); 1: Ji nf37; 2: Jidou15

将该启动子片段用 *Hind* III/*Pst* I 双酶切后,与同样经 *Hind* III/*Pst* I 双酶切的 pMD18-T-TMT 重组质粒连接,转化进 DH5 $\alpha$ -FT, PCR 法和酶切法筛选阳性克隆。图 3 显示,经 PCR 初步筛选出的阳性重组子用 *Pst* I/*Sal* I 双酶切,产生 1.0 kb 和 3.4 kb 两片段;同时用 *Hind* III/*Pst* I 双酶切时,产生

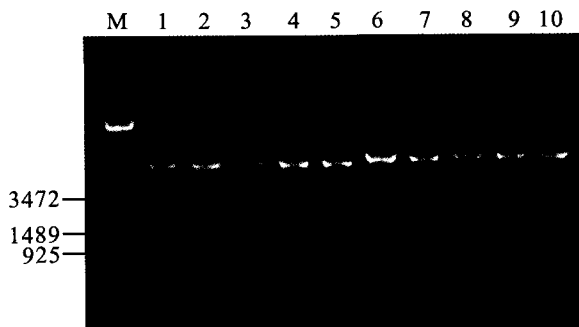


图 3 重组子 pMD18-T-G1-TMT 的酶切分析

Fig. 3 Restriction analysis of recombinant plasmid pMD18-T-G1-TMT

M: DNA 分子量标准 (bp); 1-5: 重组子 1-5 的 *Pst* I/*Sal* I 酶切; 6-10: 重组子 1-5 的 *Hind* III/*Pst* I 酶切

M:  $\lambda$ /*Eco*T14 I digest DNA Marker (bp); 1-5: recombinant lines 1-5 digested by *Pst* I/*Sal* I; 6-10: recombinant lines 1-5 digested by *Hind* III/*Pst* I

### 2.3 序列测定与分析

将阳性重组子测序后发现,两种启动子片段均为 688 bp,图 4 是其与 GenBank 现有 3 种该启动子序列(四川大豆(DQ250808)、南农 87-c38 (AY649096)和 Dare(X15121))比对结果。克隆的两种启动子片段与 GenBank 现有 3 种启动子序列

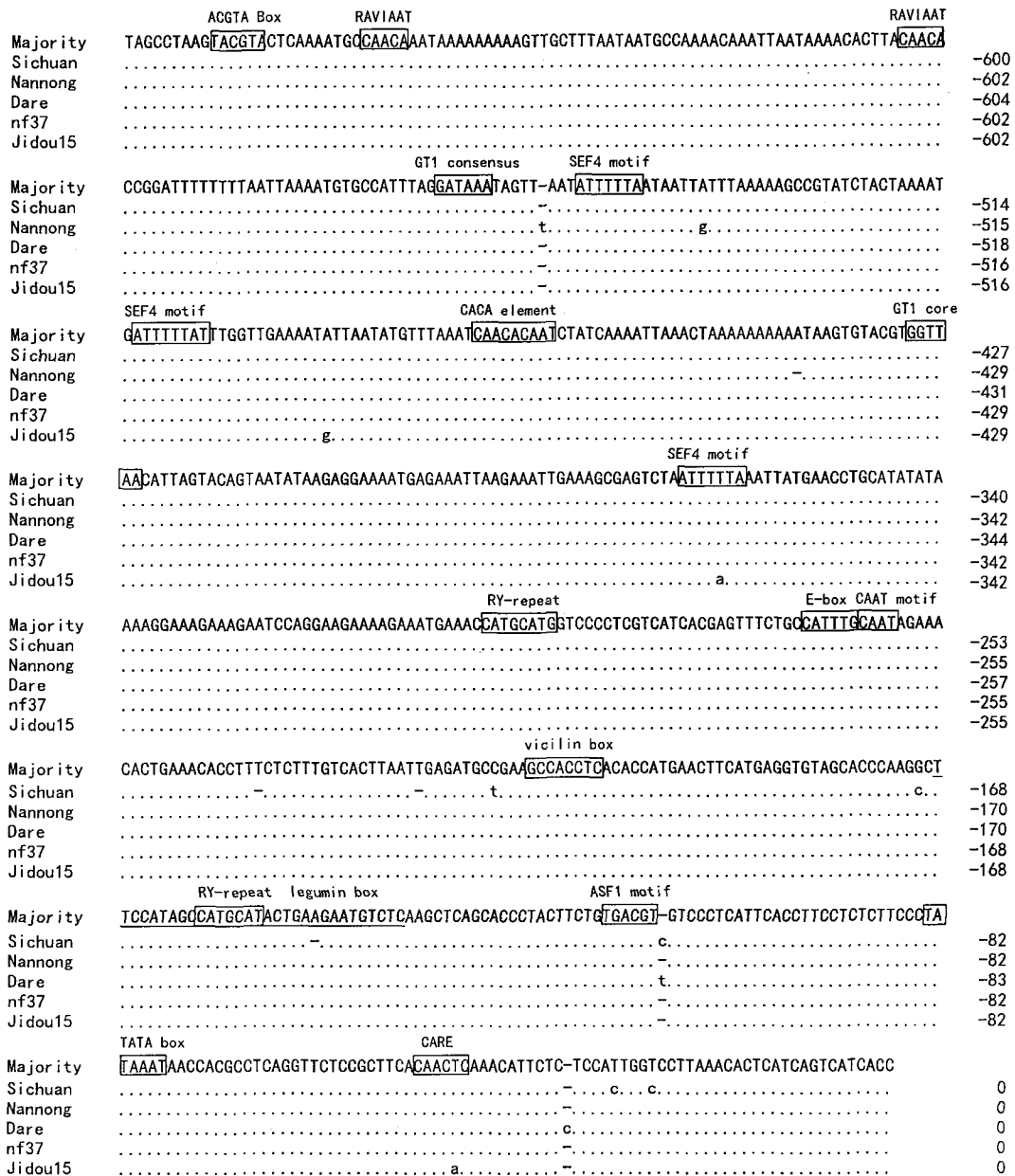


图 4 大豆球蛋白 G1 启动子序列及元件的比对分析

Fig. 4 Sequence and element alignment among glycinin G1 promoters from soybean. 序列比对所用的其他启动子序列分别来自于四川大豆(DQ250808)、南农 87-c38(AY649096)和 Dare (X15121)。启动子顺式调控元件用方框标示,数字表示距离起始密码子的位置

G1 promoter sequences used for MegAlign are from Sichuan(DQ250808), Nannong87-c38(AY649096) and Dare(X15121). Cis-regulatory elements are shown in boxes. Numbers represent the positions relative to the ATG translation start codon

间的同源性介于 96.4%~99.6%之间,其中冀 nf37 与 Dare 间的同源性最高,冀豆 15 与四川大豆间的同源性最低。启动子元件分析发现,克隆的两种启动子片段均含有 TATA 盒、CAAT 盒等启动子特征元件,同时含有 RY 重复序列元件、Legumin 盒、SEF4 基元、E 盒、CACA 元件等种子特异表达所需的顺式作用元件(Beilinson 等,2002)。由于 3 种

G1 启动子中,只有来源于四川大豆的启动子功能得到证实(Ding 等,2006),因此以它为基准进行启动子元件比对。结果发现,克隆于冀豆 15 的启动子片段在 SEF4 基元、CARE 和 Legumin 盒处与之均有差异;而克隆于冀 nf37 的启动子片段仅在 Legumin 盒上与之有一个碱基的差异,其它元件与之完全相同,由于 Legumin 盒上的该碱基在其它 4 种来源的启

动子序列(除四川大豆外)中完全一致,推测它很可能不会影响启动子的功能,因此选择克隆于冀 nf37 的 G1 启动子用于种子特异性表达载体的构建。

#### 2.4 植物表达载体的构建

用 *Hind* III 和 *Sal* I 双酶切重组质粒 pMD18-T-G1-TMT,回收 1.7 kb 片段(G1-TMT),与同样经 *Hind* III 和 *Sal* I 双酶切的双元植物表达载体 pBin438 大片段连接,转化进 DH5 $\alpha$ -FT,提取质粒,用 *Hind* III 和 *Sal* I 酶切,产生 1.7 kb 和 14 kb 两片段,说明载体构建正确(图 5)。筛选出的阳性克隆,命名为 pBG1TMT。图 6 即为所构建的种子特异性表达载体示意图,即 G1 启动子片断驱动  $\gamma$ -TMT 目的基因,同时含有 UTT 和 Nos-T 转录终止序列,以及新霉素磷酸转移酶基因(NPT II)抗性筛选标记。

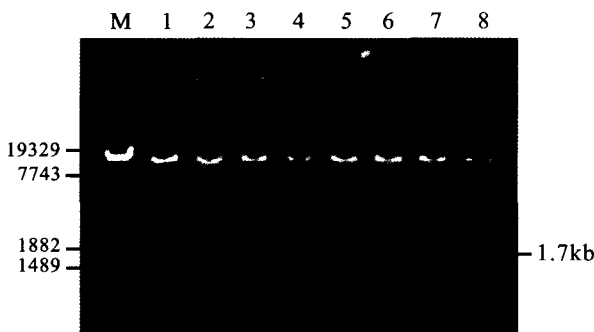


图 5 pBG1TMT 重组质粒的酶切鉴定  
Fig. 5 Restriction analysis of recombinant plasmid pBG1TMT

1-8: 重组 pBG1TMT *Hind* III/*Sal* I 双酶切  
1-8: *Hind* III/*Sal* I digestion of recombinant pBG1TMT

### 3 讨论

$\gamma$ -TMT 是维生素 E 生物合成途径中的限速酶,催化  $\gamma$ -生育酚向  $\alpha$ -生育酚的转化,因此它的活性直接影响着维生素 E 的总体活性。研究表明,种子中高含量  $\gamma$ -生育酚并非是由于  $\gamma$ -TMT 活性低,很可能是由于  $\gamma$ -TMT 基因的表达在种子中受到了抑制。欧阳青等(2003)曾使用半定量 RT-PCR 方法证实,在结球甘蓝叶片中  $\gamma$ -TMT 的表达量很高,而在种子却仅有少量表达。Tavva 等(2007)将种子特异性启动子 vicilin 与  $\gamma$ -TMT 基因连接,转化大豆,结果大豆种子中  $\gamma$ -生育酚几乎完全转变成了  $\alpha$ -生育酚,因此推测植物天然种子中  $\gamma$ -TMT 的表达受到抑制,可能是野生型启动子受到了某种调节作用,这种作用机理虽然还需相关分子生物学实验证实,但为我们利用代谢工程提高油料种子  $\alpha$ -生育酚含量提供了重要的资料。

影响外源基因在转基因植物中表达的因素很多,其中启动子的调控最关键,因此选择合适启动子是增强外源基因表达首先考虑的要素。种子特异性启动子,仅在种子成熟的中期至晚期启动下游基因特异性地表达,将这类启动子的转录调控模式应用于植物基因工程,不仅可以避免外源基因在转基因植物中非特异表达所造成的能量负荷、物质浪费以及对其它组织正常生理代谢活动的影响;更重要的是,可以使外源基因在宿主中的表达具有器官或组织特异性,可以实现外源基因特异、高效地定位表

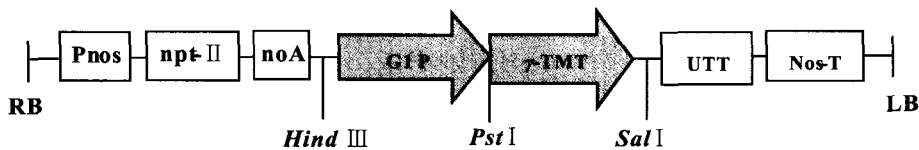


图 6 种子特异性表达载体 pBG1TMT 的示意图

Fig. 6 Schematic representation of the seed-specific expression vector pBG1TMT

G1 P: G1 启动子序列;  $\gamma$ -TMT:  $\gamma$ -生育酚甲基转移酶基因; UTT 和 Nos-T: 转录终止序列; NPT II: 新霉素磷酸转移酶基因。

G1 P: G1 promoter;  $\gamma$ -TMT:  $\gamma$ -tocopherol methyltransferase; UTT and Nos-T: termination sequence of transcription; NPT II: Neomycin phosphatransferase II.

达。因此,种子特异性启动子的使用对调控油料作物种子的成分意义非常重大。

大豆球蛋白是大豆种子中的一种主要贮藏蛋白,其表达具有高度的组织特异性和发育特异性,即在大豆种子形成的胚乳期大量表达。目前已发现 6

种大豆球蛋白 G1、G2、G3、G4、G5 和 G7。Sims & Goldberg(1989)报道了大豆 Dare 的 G1 基因全序列(X15121),但并未见其功能验证的报道。Ding 等(2006)从四川大豆中克隆了 G1 启动子,并构建了种子特异性表达人碱性成纤维细胞生长因子

(bFGF) 基因的表达载体, 转化大豆后, 表达出的 bFGF 蛋白仅存在于种子中, 含量可达种子可溶蛋白总量的 2.3%, 证明 G1 启动子具有种子特异性表达活性。本文克隆的两种启动子片段与之相比具有很高的同源性, 其中冀 nf37 的 G1 启动子片段上除 Legumin 盒上有一个碱基的差异外, 其它元件与之完全相同, 而 Legumin 盒上该差异碱基与除四川大豆外的其它四种来源的启动子完全一致, 因此猜测它的改变不会影响启动子功能, 所以选择其用于进一步构建种子特异性表达载体。冀豆 15 启动子片段在特征元件 SEF4 基元、CARE 和 Legumin 盒处与之均有差异, 对启动子功能的发挥是否有影响, 有待进一步研究。

自  $\gamma$ -TMT 基因从拟南芥和集胞藻中分离出来 (Shinatani & Dellapenna, 1998), 有关  $\gamma$ -TMT 基因的遗传转化研究仅局限在拟南芥、烟草等模式植物上, 而对农作物特别是大豆等油料作物的实用性研究却很少。大豆是我国重要的油料作物, 在人类膳食中占有重要位置, 是目前人类维生素 E 的主要来源。因此, 提高种子中  $\alpha$ -生育酚含量, 以提高食用油的维生素 E 活性, 将具有重要的营养学价值和应用前景。本文克隆了大豆球蛋白 G1 基因启动子, 并构建了  $\gamma$ -TMT 基因的植物载体, 对于进一步利用种子特异性启动子提高油料作物种子中  $\alpha$ -生育酚的含量, 从而提高作物营养品质等方面的研究具有重要的应用前景。

## 参考文献:

- 欧阳青, 樊春涛, 孙卉, 等. 2003. 结球甘蓝  $\gamma$ -生育酚甲基转移酶 cDNA 的克隆、分析及其异源表达酶蛋白的功能研究[J]. 自然科学进展 13:709-715
- Baumlein H, Nagy I, Villarroel R, et al. 1992. Cis-analysis of a seed protein gene promoter: the conservative RY repeat CATG-CATG within the legumin box is essential for tissue-specific expression of a legumin gene[J]. *Plant J*, 2:233-239
- Beilinson V, Chen Z, Shoemaker C, et al. 2002. Genomic organization of glycinin genes in soybean[J]. *Theor Appl Genet*, 104:1 132-1 140
- Buring JE, Hennekens CH. 1997. Antioxidant vitamins and cardiovascular disease[J]. *Nutr Rev*, 55:S53-60
- Chen DF, Chen XW. 2004. Cloning and activity analysis of *in vitro* expression of plant NAD-IDH genes[J]. *Chin Sci Bull*, 49:328-336
- Ding SH, Huang LY, Wang YD, et al. 2006. High-level expression of basic fibroblast growth factor in transgenic soybean seeds and characterization of its biological activity[J]. *Biotechnol Lett*, 28:869-875
- Eckardt NA. 2003. Vitamin E-defective mutants of *Arabidopsis* tell tales of convergent evolution[J]. *Plant Cell*, 15:2 233-2 235
- Fryer MJ. 1992. The antioxidant effects of thylakoid vitamin E ( $\alpha$ -tocopherol)[J]. *Plant Cell Environ*, 15:381-392
- Grusak MA, Dellapenna D. 1999. Improving the nutrient composition of plants to enhance human nutrition and health[J]. *Annu Rev Plant Physiol Plant Mol Biol*, 50:133-161
- Karin H. 2003. Vitamin production in transgenic plants[J]. *Plant Physiol*, 160:821-829
- Lelievre JM, Oliveira LO, Nielsen NC. 1992. 5'-CATGCAT-3' elements modulate the expression of glycinin genes: comparison of the DNA sequence[J]. *Plant Physiol*, 98:387-391
- Qian WC(钱文成), Chen DF(陈德富), Wang HZ(王绘砖), et al. 2007. Gene cloning and function confirmation of  $\gamma$ -tocopherol methyltransferase from *Brassica napus* (甘蓝型油菜  $\gamma$ -生育酚甲基转移酶基因的克隆和功能验证)[J]. *Plant Physiol Commun*(植物生理学通讯), 43:683-688
- Rimm EB, Stampfer MJ, Ascherio A, et al. 1993. Vitamin E consumption and the risk of coronary heart disease in men[J]. *New Engl J Med*, 328:1 450-1 456
- Saghai-Marouf MA, Soliman KM, Jorgensen, et al. 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics[J]. *Proc Natl Acad Sci USA*, 81:8 014-8 018
- Shinatani D, DellaPenna D. 1998. Elevating the vitamin E content of plants through metabolic engineering[J]. *Science*, 282:2 098-2 100
- Sims TL, Goldberg RB. 1989. The glycinin Gyl gene from soybean[J]. *Nucleic Acids Res*, 17:4 386
- Stampfer MJ, Hennekens CH, Manson JE, et al. 1993. Vitamin E consumption and the risk of coronary disease in women[J]. *New Engl J Med*, 328:1 444-1 449
- Tavva VS, Kim YH, Kagan IA, et al. 2007. Increased alpha-tocopherol content in soybean seed overexpressing the *Perilla frutescens* gamma-tocopherol methyltransferase gene [J]. *Plant Cell Rep*, 26:61-70