

# 大豆 *P39-1* 基因及其启动子表达方式的初步研究

赵 艳<sup>1</sup>, 赵 丹<sup>2</sup>, 刘 博<sup>1</sup>, 国春晖<sup>1</sup>

(1. 齐齐哈尔大学 生命科学与农林学院, 黑龙江 齐齐哈尔 161006; 2. 吉林市第七中学, 吉林 吉林 132011)

**摘要:** 利用实时定量 PCR 方法, 检测大豆 *P39-1* 基因在大豆各组织中的表达方式, 结果显示该基因在大豆根、茎、叶中的表达活性低, 而在花和种子中的相对表达活性极高。利用 PCR 方法, 克隆大豆 *P39-1* 基因 5' 端上游 2 000 bp 序列, 命名为 *P39-1-p*。在线启动子预测分析表明 *P39-1-p* 序列中含有多种典型的种子特异表达元件和花特异表达的元件, 如 *SEF1* motif、*SEF3* motif、*SEF4* motif、*E-box*、*G-box*、*AACACA*、*AACA*、*ACGT*、*CCAA*、*52-box*、*ntp303-box*、*GTGA*、*TACPyATbox*, 推测大豆 *P39-1-p* 启动子具有调控下游基因大量表达在花和种子中的特性。

**关键词:** 大豆; *P39-1* 基因; 启动子; 克隆

中图分类号: S565.1

文献标识码: A

文章编号: 1000-9841(2013)03-0425-04

## Preliminary Studies on the Expression Pattern of Soybean *P39-1* Gene and Promoter

ZHAO Yan<sup>1</sup>, ZHAO Dan<sup>2</sup>, LIU Bo<sup>1</sup>, GUO Chun-hui<sup>1</sup>

(1. College of Life Science and Agroforestry, Qiqihar University, Qiqihar 161006, China; 2. Jilin City No. 7 Middle School, Jilin 132011, China)

**Abstract:** The expression pattern of soybean *P39-1* gene in soybean tissues was detected by real-time quantitative PCR (RTQ-PCR), and the result showed that it had lower activity in roots, stems and leaves, while higher activity in flowers and seeds. The 5'-flanking upstream sequence of soybean *P39-1* gene, named *P39-1-p*, was isolated from soybean genomic DNA by PCR method, and the length was 2 000 bp. Sequence analysis revealed that this fragment contained a series of motifs related to seed-specific promoters and some pollen-expressed elements, such as *SEF1* motif, *SEF3* motif, *SEF4* motif, *E-box*, *G-box*, *AACACA*, *AACA*, *ACGT*, *CCAA*, *52-box*, *ntp303-box*, *GTGA* and *TACPyATbox*. It is inferred that *P39-1-p* promoter possesses the function driven downstream gene expression abundantly in soybean flowers and seeds.

**Key words:** Soybean; *P39-1* gene; Promoter; Cloning

在转基因大豆中增加或降低某些物质含量, 实现外源目的基因在转基因大豆中正常、有效表达, 首要考虑的问题是选择适合的植物启动子。组织特异性启动子能够调控下游基因的表达具有组织特异性, 避免造成植物营养不必要的浪费, 因而成为当前研究热点。

*P39* 蛋白存在于多种高等植物中, 如玉米<sup>[1]</sup>、水稻、大麦、拟南芥和落花生等<sup>[2]</sup>, 是蛋白凝胶电泳分离卵磷脂过程中获得的分子量为 39 kDa 的一种蛋白质, 该蛋白的 N 端具有保守氨基酸序列 SDKEDS-VFKGIN<sup>[3]</sup>。大豆中的 *P39* 编码一个多基因家族, 包括 *P39-1*、*P39-2*、*P39-3*、*P39-4*、*P39-5* 和 *P39-6*。2008 年, Xiang 等克隆了大豆的 *P39-1* 和 *P39-2* 两个同源基因, 同源率为 92%。Northern 杂交实验表明 *P39-1* 基因在大豆种子中具有较高的活性<sup>[2]</sup>。推测 *P39-1* 基因上游的启动子很可能具有调控下游基因在种子中高表达的特性。为明确该基因在大豆各组织中的表达量, 利用实时荧光定量 PCR

(qRT-PCR) 检测 *P39-1* 基因在大豆各组织中的表达量; 并克隆 *P39-1* 基因 5' 端上游 2 000 bp 的序列, 利用生物信息学方法对 *P39-1* 基因启动子的表达方式进行分析。

## 1 材料与方法

### 1.1 材料

大豆品种(吉豆 2 号)、大肠杆菌 DH 5 $\alpha$  均为本实验室保存; RNAiso Reagent 试剂盒、逆转录试剂盒、pMD18-T 克隆载体、限制性内切酶 *Hind* III、*Eco*R I、*T*<sub>4</sub> 连接酶、*ExTaq* 均购自 Takara 公司; 引物由上海生工合成; DNA 凝胶回收试剂盒购自爱思进公司, 其他试剂均为进口或国产分析纯。

### 1.2 总 RNA 提取及 cDNA 合成

利用 RNA 提取试剂盒, 提取大豆根、茎、叶、花和种子的总 RNA, 以各自提取的总 RNA 为模板, 按照逆转录试剂盒说明分别合成大豆根、茎、叶、花和

收稿日期: 2013-03-01

基金项目: 黑龙江省教育厅科学技术研究项目(12521611); 作物生物学国家重点实验室开放基金(2012KF13); 齐齐哈尔大学青年教师科研启动支持计划项目(2010K-M09)。

第一作者简介: 赵艳(1981-), 女, 博士, 讲师, 主要从事植物分子生物学研究。E-mail: zhaoyan3053877@yahoo.com.cn。

种子的 cDNA 第一条链。

### 1.3 qRT-PCR

分别以大豆各组织的 cDNA 为模板,进行 qRT-PCR 反应。以大豆管家基因  $\beta$ -tubulin (GMU12286) 为内参, B<sub>1</sub> (5'-GGAAGGCTTTCTTGCATTGGTA-3') 和 B<sub>2</sub> (5'-AGTGGCATCCTGGTACTGC-3') 为引物。根据大豆 P39-1 基因的核苷酸序列 (42722638), 设计检测大豆 P39-1 基因在大豆各组织中表达量的引物, 分别为 PS (5'-GCTTTCCTAGTCTTTTCTAATC-CAG-3') 和 PA (5'-GCAACCTCATCAGCCAAACA-3')。

### 1.4 大豆 P39-1 基因启动子的克隆及预测分析

以大豆叶片为材料,采用 CTAB 法提取大豆的基因组 DNA<sup>[4]</sup>。根据大豆基因组序列 (<http://www.phytozome.net/soybean>), 以基因组 DNA 为模板,设计引物 (P<sub>1</sub>: 5'-AAATAGTAGTCAGATAATGAT-TATGG-3'; P<sub>2</sub>: 5'-GGTGAGATTGAATCTCAAAGG-3'), 扩增大豆 P39-1 基因 ATG 上游 2 000 bp 的序列。PCR 反应条件为预变性 95℃ 5 min; 94℃ 30 s, 50℃ 40 s, 72℃ 1 min, 共 30 个循环; 72℃ 后延伸 7 min。将 PCR 扩增片段回收,并连接到 pMD18-T 克隆载体上,获得重组质粒并进行测序。利用启动子预测软件 PlantCARE (<http://www.dna.affrc.go.jp/PLACE/signalscan.html>) 分析启动子序列中的顺式作用元件。

## 2 结果与分析

### 2.1 大豆 P39-1 基因的组织表达特性

以大豆管家基因  $\beta$ -tubulin 作为内参基因,采用 qRT-PCR 方法,对大豆根、茎、叶、花和种子中 P39-1 基因的表达量进行分析。结果显示,大豆 P39-1 基因在根、茎、叶中表达活性低,而在花和种子中的相对表达活性极高,种子中相对表达量最高(图 1)。

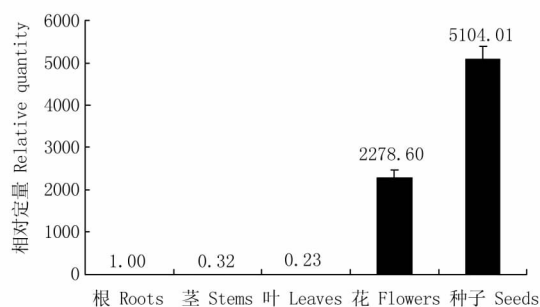
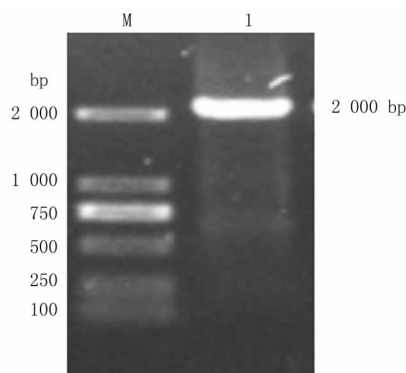


图 1 P39-1 基因在大豆各组织中的相对表达

Fig. 1 The relative expression of soybean P39-1 gene in different soybean tissues

### 2.2 大豆 P39-1 基因启动子的克隆

以大豆基因组 DNA 为模板,PCR 扩增大豆 P39-1 基因上游序列,获得长度为 2 000 bp 的片段,命名为 P39-1-p(图 2)。将该片段连接到 pMD18-T 载体上,获得重组质粒 pMD18-T-P39-1-p。将其转入大肠杆菌中,提取质粒进行鉴定,用 Hind III 和 EcoR I 双酶切鉴定结果得到与预期相符的片段(图 3)。

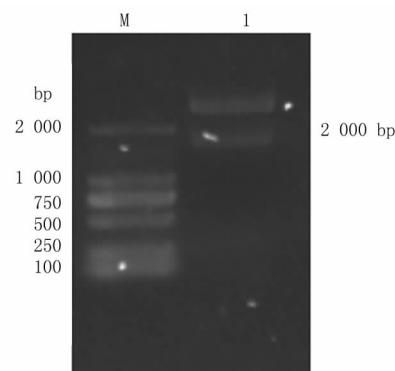


M. 2 000 bp DNA marker; 1. PCR 产物

M. 2 000 bp DNA marker; 1. PCR amplification of P39-1-p

图 2 P39-1-p 的扩增产物

Fig. 2 PCR amplification of P39-1-p



M. 2 000 bp DNA Marker; 1 ~ 2. Hind III 和 EcoR I 双酶切片段

M. 2 000 bp DNA Marker; 1 ~ 2. Hind III and EcoR I digestion of pMD18-T-P39-1-p

图 3 pMD18-T-P39-1-p 双酶切鉴定

Fig. 3 Restriction enzyme digestion identification of pMD18-T-P39-1-p

### 2.3 大豆 P39-1-p 启动子的序列分析

生物信息学分析表明,P39-1-p 序列中含有多种典型的种子特异表达元件和花特异表达元件(图 4), 如:SEF1 motif<sup>[5]</sup>、SEF3 motif<sup>[6]</sup>、SEF4 motif(胚特异蛋白 SEF4 的结合位点)<sup>[7]</sup>、E-box(常出现在参与三酰甘油合成和植物种子特异表达基因的启动子中)<sup>[8-9]</sup>、G-box(广泛存在于各种种子特异性启动子中)<sup>[10-11]</sup>和多个种子特异表达元件 AACACA<sup>[12]</sup>、AACACA<sup>[13]</sup>、ACGT<sup>[14]</sup>、CCAA<sup>[6]</sup>等;花特异表达元件如:52-box<sup>[15]</sup>、ntp303-box<sup>[16]</sup>、GTGA<sup>[17]</sup>、TACPyATbox<sup>[18]</sup>等。

1 AAATAGTAGTCAGATAATGATTATGGTTCTTAAGTAATTTCTAGCCACTGTAATTTGTGGCAATTGAAAC  
 71 TTATGTGATGGTTTCAATTGTTTCATTACACACACACACACACATATATATATATATATATATA  
 141 TATATATATATATATATATATATATATTTATTTCACACCAATCTCGAGTATCATTATCATTTCATCAAG  
 211 TGTTTACAATGATTCCACCTGAAAGAAAAAGAAAAAATGAAAAGAAAACCCCAACAAGGAACAAACG  
 281 TGTGATTGTGGTGTAATATAGTATATGTATGTAAATTGCCGTTGGGGTGGAGGACTTGGCCCGTTATGTA  
 351 AGAGGAGCATATGTTGCTAGCAAGAAACACGTAAGAGGGAGCAACTTCAATAGTTTGGCCAGCTCTCTTC  
 421 GTTGGCACAATCTGAACCAATCGAACCAATTGTTTCTCGTACATTGTTGCTACTATTAACGGAGTCCAGA  
 491 TGGGTCCAGGCGCAACCTAAAAGACCAAACTAGACTTGGATTAATAGGATCAACTATATCAATTTTTTTA  
 561 TATAACATATTTACGTTATAAAATGATTTTTTTTTTACTTTCTCTTCACTATAGATTTATTAGCGTATAT  
 631 ACTGATGTTTTAATATGAGTCCATTACAAAATTACATTATTTATATATTTTAATTTTTTAAATATTTTC  
 701 TAGTTATAATTTATTAATACACTCATAATAAAAAAATATATATACATTAAACAAAAATTTTATTAATA  
 771 TTTACTAACCTACACATCATTATTTTTTAGTATGAATGTTTAAACACACTCATAAAAAAATAATGGTATGT  
 841 ATGTTAAACAAATTTTCATATATGTATATATATATATATATATATATATATATATATATATATATATA  
 911 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA  
 981 TACTTTTTGTTCTTCAACAAACATAGCATGATAAGAGTACATGTGGTGTGATGGAGATGGAACCGGA  
 1051 CGGATAATAATGTTGTTCCAATAATTGATTATAGCGGGGTGACGCAGCCGGCGAGAATTGGCAGAAGGA  
 1121 TGTCTTGACGTGTTGTGTCGATTTCTTGACCGGAGAGAATTTAGAAATCTCCTTTGGATCAAAGTCAC  
 1191 AAAGTGAATCGAACTCGTTATAATTGACGTGTCATCGTGTACAAACCGTAAGATTATACAGAAAAGTA  
 1261 GAAGGCTCCATCCATCACAGAGATGAAATTGAAAAATAAGGAATTAATTAAGTTTGGTATATTATTGTAT  
 1331 ATGAATTAATTAATCAATTTAGAGTTAATATTACGTGGACAAACATTGATTTAGTTGATGAGAATATATA  
 1401 AATGTGTTTGATTTTCATTATCAATATAAAAAAATTATTGATAGATAATATATAAATATATTTATAAG  
 1471 TATTTACATAATAATAATAAAAAAGTTTTAAATTAATCCACCTAAATTTGCATAATAATAAAAAAT  
 1541 TTTTATAATCATATAACTAATTTAAATCTATCTAATGGATATCACTTTTAAATAATTATTTTTAAAAA  
 1611 TCATCTTATCATATATCAACTTATAATTGGATGAAAAAGTAAATATCGTTACACTATAATACATTATAA  
 1681 TTAAATTCATTAACCTGTGTTTGATCTAAGTATGAAGAACAAGAAAGAAATAGGTGGTTTGGACTCT  
 1751 GGATTTTGCTCGGTACAGCTGTATGCATTTGATTAATATTATCAGAACTAATTAAGTATAAGCAGTTGA  
 1821 TTAGTTTAATGAGAGCGAACCAGATCACGGATTAAGTTGTGACCTTGCATGTCCGCCACGTGGAAGATCA  
 1891 AGCACTTGAAACCTGTTACTACTTTAGGCACCAATTCTTGCTACTCTCCATCTCTTAATTCAGAGTTTA  
 1961 ATTCATCAATGGGTGTTTTCTTTGAGATTCAATCTCACC  
 ↳ ATG of *P39-1* gene

图 4 *P39-1-p* 序列的生物信息学分析Fig. 4 Bioinformatics analysis of DNA sequence of *P39-1-p*

### 3 讨 论

本研究通过 qRT-PCR 和生物信息学方法,对大

豆 *P39-1* 基因及其启动子序列进行初步分析。qRT-PCR 检测结果表明大豆 *P39-1* 基因在大豆花和种子中的表达活性极高,相对表达量是根、茎、叶中表达

量的 2 000 ~ 5 000 倍;说明大豆 *P39-1* 基因在大豆的花和种子中表达量极高。继续利用生物信息学软件预测该基因的启动子序列,序列中具有多种典型的种子特异表达元件和花特异表达元件,且多数顺式元件的数量多个,集中在 ATG 上游 1 ~ 630 bp 区域的顺式元件较多,这些顺式元件的存在及位置可能与大豆 *P39-1* 基因启动子调控下游基因大量表达在花和种子中有关。

本研究通过 qRT-PCR 检测 *P39-1* 基因在大豆各组织中的表达量,与 Xiang 等<sup>[2]</sup> 的 Northern 杂交实验相比,qRT-PCR 能够更明显地表明 *P39-1* 基因在大豆各组织中表达活性的差异。此外,通过 qRT-PCR 检测可知,*P39-1* 基因在大豆花中的表达活性也较高。

根据以上的研究结果,初步推测 *P39-1* 基因启动子在大豆花和种子中具有极高的调控活性,在转基因大豆的稳定表达中是否仍具有调控下游基因在花和种子中大量表达的活性,还需进一步验证。

## 参考文献

- [1] Verza N C, Esilva T R, Neto G C, et al. Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags[J]. *Plant Molecular Biology*, 2005, 59 (2): 363-374.
- [2] Xiang P, Baird L M, Jung R, et al. P39, a novel soybean protein allergen, belongs to a plant-specific protein family is present in protein storage vacuoles[J]. *Journal of Agricultural and Food Chemistry*, 2008, 56: 2266-2272.
- [3] Gu X, Beardslee T, Zeece M, et al. Identification of IgE-binding proteins in soy lecithin[J]. *International Archives of Allergy and Immunology*, 2001, 126(3): 218-225.
- [4] 肖月华, 罗明, 方卫国, 等. 利用 YADE 法进行棉花基因组 PCR 步行[J]. *遗传学报*, 2002, 29(1): 62-66. (Xiao Y H, Luo M, Fang W G, et al. PCR walking in cotton genome using YADE method[J]. *Acta Genetica Sinica*, 2002, 29(1): 62-66.)
- [5] Yoshino M, Nagamatdu A, Tsutsumi K I, et al. The regulatory function of the upstream sequence of the  $\beta$ -conglycinin  $\alpha$  subunit gene in seed-specific transcription is associated with the presence of the RY sequence[J]. *Genes and Genetic Systems*, 2006, 81(2): 135-141.
- [6] Zhao Y, Shao S L, Li X W, et al. Isolation and activity analysis of a seed-abundant *soyAP1* gene promoter from soybean[J]. *Plant Molecular Biology Reporter*, 2012, 30: 1400-1407.
- [7] Chung K J, Hwang S K, Hahn B S, et al. Authentic seed-specific activity of the *Perillaoleosin* 19 gene promoter in transgenic *Arabidopsis*[J]. *Plant Cell Reports*, 2008, 27: 29-37.
- [8] Kim M J, Go Y S, Lee S B, et al. Seed-expressed casein kinase I acts as a positive regulator of the *SeFAD2* promoter via phosphorylation of the *SeBHLH* transcription factor[J]. *Plant Molecular Biology*, 2010, 73(4-5): 425-437.
- [9] Xie C J, Zhang B B, Wang D, et al. Molecular cloning and characterization of an achene-seed-specific promoter from motherwort (*Leonurus japonicus* Houtt)[J]. *Biotechnology Letters*, 2011, 33(1): 167-172.
- [10] Kim M J, Kim J K, Shin J S, et al. The *SeBHLH* transcription factor mediates trans-activation of the *SeFAD2* gene promoter through binding to E- and G-box elements[J]. *Plant Molecular Biology*, 2007, 64(4): 453-466.
- [11] Gao M J, Lydiate D J, Li X, et al. Repression of seed maturation genes by a trihelix transcriptional repressor in *Arabidopsis* seedlings[J]. *The Plant Cell*, 2009, 21: 54-71.
- [12] Luo K, Zhang G F, Deng W, et al. Functional characterization of a cotton late embryogenesis-abundant *D113* gene promoter in transgenic tobacco[J]. *Plant Cell Reports*, 2008, 27: 707-717.
- [13] Hu X W, Liu S X, Guo J C, et al. Embryo and anther regulation of the mabinlin II sweet protein gene in *Capparis masaiikai* Lévl[J]. *Functional & Integrative Genomics*, 2009, 9: 351-361.
- [14] Zavallo D, Lopez B M, Hopp H E, et al. Isolation and functional characterization of two novel seed-specific promoters from sunflower (*Helianthus annuus* L.)[J]. *Plant Cell Reports*, 2010, 29(3): 239-248.
- [15] Eyal Y, Curie C, McCormick S. Pollen specificity elements reside in 30 bp of the proximal promoters of two pollen-expressed genes[J]. *Plant Cell*, 1995, 7: 373-384.
- [16] Zhang Q, Liu H, Cao J. Identification and preliminary analysis of a new PCP promoter from *Brassica rapa* ssp. *chinensis*[J]. *Molecular Biology Reports*, 2008, 35: 685-691.
- [17] Gupta V, Khurana R, Tyagi A K. Promoters of two anther-specific genes confer organ-specific gene expression in a stage-specific manner in transgenic systems[J]. *Plant Cell Reports*, 2007, 26: 1919-1931.
- [18] 张树珍, 杨本鹏, 刘飞虎. 花特异表达启动子 *PchsA* 的克隆及其序列分析[J]. *农业生物技术学报*, 2002, 10(2): 116-119. (Zhang S Z, Yang B P, Liu F H. Cloning and sequence analysis of a flower-specific expression promoter *PchsA*[J]. *Journal of Agricultural Biotechnology*, 2002, 10(2): 116-119.)