

提高外源基因在植物体内表达的策略

王彪, 武天龙*

(上海交通大学植物科学技术系, 上海 201101)

摘要: 介绍提高外源基因在植物体内表达的方法。从外源基因的优化、整合、转录、翻译、运输以及基因间的相互作用等方面, 总结提高外源蛋白在植物宿主体内表达的常用策略。

关键词: 外源蛋白; 外源基因; 基因表达; 综述

中图分类号: Q786

文献标识码: A

文章编号: 1005-3395(2005)01-0080-05

Strategies to Enhance Expression of Foreign Genes in Plants

WANG Biao, WU Tian-long*

(Department of Plant Science and Technology of Shanghai, Jiao Tong University, Shanghai 201101, China)

Abstract: Strategies to achieve high-level expression of foreign genes in plants are mainly introduced in this paper. The authors summarize general ways to improve amounts of foreign proteins in plants in respect of optimization, integration, transcript, translation, transport and interaction of genes.

Key words: Foreign proteins; Foreign genes; Gene expression; Review

在建立植物生物反应器的过程中, 提高外源蛋白在植物体内的表达水平是必须解决的首要问题。本文对如何增强外源基因在植物宿主体内的表达水平从以下几个方面进行综述, 以期能为植物转基因技术的应用提供参考。

1 寻找在宿主体内表达效率高的启动子和调节序列

植物种类繁多, 遗传多样性丰富。植物基因的启动子和调控序列类型也多种多样, 寻找能在宿主体内高效表达的特异性启动子是提高外源基因在植物体内表达途径之一。植物基因工程中常用的启动子有的来自质粒, 如章鱼碱型启动子和胭脂碱型启动子; 有的来自病毒, 如花椰菜花叶病毒 35S 启动子(CaMV35S); 有的来自动物, 如果蝇热激蛋白基因启动子; 但大部分启动子还是来源于植物, 如单子叶转基因植物主要使用的玉米泛素(Ubiquitin)启动子和水稻肌动蛋白(Actin1)启动子, 查尔酮合成

酶基因启动子, 1,5-二磷酸核酮糖羧化酶小亚基基因启动子, 豆科植物种子凝集素基因启动子, 大豆谷氨酸合成酶基因启动子^[1], β -伴大豆球蛋白基因 α 亚单位启动子^[2]等等。植物启动子类型也很多, 有双向启动子和可变启动子; 环境诱导表达型启动子: 如激素、创伤、光和真菌诱导表达的启动子; 植物发育中特异表达启动子等等^[3]。启动子的强弱是影响基因表达的重要因素, 但启动子种类繁多, 同一种启动子在不同宿主体内表达的效率有很大差别, 有的启动子表达还具有组织和器官特异性, 同时受到调节序列和其他启动子的影响。现在已经商业化的转基因大豆采用 CaMV35S 或增强型 35S 启动子, 表明 35S 启动子在大豆中表达效率较高^[4]。Kamo^[5]研究了 β -葡萄糖苷酸酶(GUS)基因在花椰菜花叶病毒基因启动子 (CaMV35S)、泛素启动子 (*ubq3*)、甘露碱合成酶基因启动子 (*mas2*) 和鸟氨酸环化脱氨酶基因启动子 (*rolD*) 等控制下, GUS 在烟草愈伤组织、茎杆和根中的表达情况, 发现在茎杆中以 CaMV35S 启动子为最高, 根中 *rolD* 启动子最

收稿日期: 2003-12-12 接受日期: 2004-03-17

基金项目: “863” 重大专项油料作物转基因项目(2001AA212141)资助

* 通讯作者 Corresponding author

高。比较温室和室外培养的结果发现, *GUS* 表达量在 CaMV35S、*mas2* 和 *ubq3* 三种启动子中没有区别, 但 *rolD* 启动子表现却不同, 室外和室内培养的植株茎杆中 *GUS* 表达量相差 4 倍, 根中相差 8 倍。总的来说, *GUS* 在茎杆中表达量比根中要低, 这可能与不同部位细胞的分生能力有关。Olga 等^[6]分别用马铃薯的马铃薯块茎储藏蛋白(Patatin)启动子和 CaMV35S 启动子构建单链 Fv 抗体基因的表达载体转化马铃薯, 结果发现用 CaMV35S 启动子构建的载体转化植株比 Patatin 启动子的表达量要高, 重组蛋白占可溶性总蛋白的 2%, 在 4℃ 下保存 18 个月仍有一半抗体具有活力。

研究发现, 植物基因启动子常利用其第一个 ATG 作为起始密码子, ATG 旁侧的保守序列与动物基因的 CCACCATG(G) 保守序列不同, 为 TAAA CAATGGCT^[7]。启动子旁侧及其上游的调节序列, 能够感受光、不同发育阶段或微生物感染引起的信号, 因而基因的表达受到信号因子的诱导和调节^[8]。起始位点上游的调节序列, 也对转录的效率有很大的影响, 如 Ω 序列等。在烟草和水稻中, 以 CaMV35S 为启动子, 插入 Ω 序列和内含子序列, 使 *GUS* 基因的活性提高 20–70 倍^[9]。秦红敏等^[10]利用 *GUS* 基因和双增强的 35S 启动子, 构建 GUS-NOS (胭脂碱合成酶终止子), Ω -GUS-NOS, Ω -GUS-3'-UTR-NOS 和 Ω -GUS-3'-UTR(3'非翻译端)组成 4 个不同的表达框架, 对转基因烟草 *GUS* 活性检测表明, 前 3 种载体表达量有显著差异, 以 Ω -GUS-3'-UTR-NOS 表达量为最高, 在整体植株上证明 Ω 序列能提高外源基因的表达水平。在转基因植物中, 3'-UTR 似乎不能起 poly(A)的作用, 也不能作为转录的终止子, 但 3'-UTR 和 Ω 序列共同作用, 能使外源基因的表达效果更好。甚至基因的内含子序列对基因的转录也有影响。Xu 等^[11]研究发现, 细胞质磷酸丙糖异构酶基因中的第一个内含子对磷酸丙糖异构酶基因在水稻、大麦和玉米叶子中的表达是必须的。还有其他常见表达调控元件, 如增强子、CCAAT 盒元件、G 盒元件等等, 都能控制下游的基因表达。Rieping 利用基因定点突变技术, 删去 CCAAT 元件, 能使目的基因表达量降低 5 倍^[12]。

2 构建融合基因

将外源基因与植物特异蛋白基因及其调控序列连接起来, 使重组蛋白在植物体内的表达具有器

官和组织特异性, 可大大提高外源蛋白的表达量。这种方法是利用在植物中能够高效表达基因的部分序列、启动子和调控序列与目的基因结合, 构建融合基因, 从而实现外源基因的表达和分泌, 利用融合蛋白的特性将目的蛋白富集和保护起来, 免受宿主体内蛋白水解酶的攻击, 从而提高外源蛋白的含量。Van Rooijen 和 Moloney^[13]将葡萄糖醛酸酶基因与油蛋白的前 800 bp 核苷酸序列连接, 构建融合蛋白载体, 导入油菜植株, 融合蛋白 80%能正确定位于油体, 提高了重组蛋白的表达量, 简化了外源蛋白的提取和纯化步骤。Hayashi 等^[14]将 2S 清蛋白基因与磷丝菌素转移酶(PPT)的基因构建成一个融合基因转入拟南芥中, 发现虽然成功转入 *ppt* 基因, 但转化的拟南芥仍然对磷丝菌素敏感, 进一步研究发现是融合蛋白将 PPT 蛋白引向 2S 清蛋白前体累积小泡, 避免外源蛋白由融合蛋白中切出或释放。Sojikul 等^[15]将大豆营养贮存蛋白 *vspA* 的信号肽与肝炎 B 表面抗原(HBsAg)构建成融合蛋白, 使用的是增强型花椰菜花叶病毒启动子, 虽然融合蛋白没有将信号肽正确地切除, 但表达的融合蛋白稳定性高, 融合蛋白的表达量也比没有修正的 HBsAg 表达水平高得多。

3 利用分泌途径将外源蛋白引向特定的部位

利用分泌途径来提高外源蛋白的含量。将外源蛋白导向特定的部位或分泌到细胞外, 避免宿主细胞将外源蛋白视为异己蛋白而降解, 从而提高了外源蛋白的积累。真核生物分泌蛋白 N 端一般都有信号肽序列, 引导蛋白质定位于内质网膜, 经高尔基体加工后分送到细胞的各个部位, 特别是膜蛋白和细胞外蛋白。Sijmons 等^[16]用人清蛋白(HSA)作为目的基因, 使用人的前导序列和烟草 PR-S 序列, 用 pMOG285 作对照, 转化烟草, 发现在积累 HSA 蛋白方面有明显的差异, 并推测表达的差异是由于非分泌型 HSA 不稳定造成的, 而这些特异的序列能将与之相连的蛋白引向特定目标, 最后这些引导序列能被正确切除。研究表明, 蛋白质 C 端 KDEL 序列能将蛋白质定位在内质网膜上。通过改变蛋白质的细胞定位和分泌途径, 能明显提高外源蛋白在宿主体内的表达水平^[17,18]。Olga 等^[6]构建含有 KDEL (Lys-Asp-Glu-Leu)和不含有 KDEL 的载体分别转化马铃薯, ScFv 抗体的表达具有明显差异。Udo 等^[19]

比较烟草叶子和种子中在有或无 KDEL 信号情况下 ScFv-ox 表达量的差别,结果表明:以 CaMV35S 为启动子,ScFv-ox 在叶子中的表达量从无 KDEL 信号时的 0.1%–0.2%提高到有 KDEL 信号时的 1.1%–4.0%;在种子可溶性蛋白中,以 LeB4 为启动子,ScFv-ox 表达量从无 KDEL 信号时的 0.2%–0.6%提高到有 KDEL 信号时的 1.0%–1.9%。

4 反义基因技术

反义基因技术是利用某基因 mRNA 与其反义基因的 RNA 能特异结合的特点,在翻译水平上阻断该基因表达。反义基因技术已经在植物基因工程中广泛应用,主要用于分析相应基因的功能,改良农作物的品质,提高重要的有经济价值的农艺性状,提高农作物的抗病性能等方面^[20]。运用反义技术抑制植物体内相应基因的活性,提高作物的品质特别是种子贮存蛋白的品质。Kohno-Murase 等^[21]将 *napin* 的反义基因转入油菜,并在种子中特异表达,发现种子中 Napin 蛋白含量减少,同时 Cruciferin 蛋白含量增加 1.4 倍以上。但在转高油作物(如油菜)时发现,用反义基因转化的植物总蛋白和总油脂的相对含量并没有改变。反义技术能提高外源蛋白表达,在 1999 年,Alain 等人^[22]用拟南芥种子 2S 清蛋白的反义基因和 *arcclin(arc)5-I* 基因构建载体,转化马铃薯,结果发现 2S 清蛋白在种子中累积减少,但 ARC-5 在种子中积累大大增加,表达量占种子总蛋白量的 24%,这表明反义基因对提高外源基因在异体表达中有很大的促进作用。反义基因技术用于大豆的研究中,Anthony 等人^[23]用 β -伴大豆球蛋白的反义基因转化大豆发现,反义基因极大地抑制 β -伴大豆球蛋白的表达,种子中 7S 球蛋白的含量明显增加。同时,植物种子中蛋白的累积生理代谢途径也发生明显的改变,蛋白体的含量增加,而液泡贮存蛋白则减少,反义基因诱导了内质网来源蛋白体的形成。更为重要的是反义抑制并不会损伤植物的生产能力,Staswick 等^[24]用反义 *vspA* 转化大豆,营养贮存蛋白降至原来的 2%,VSP α 降至 1%,VSP β 降至 10%。田间试验表明,种子产量、蛋白质含量、油份和氨基酸组成在 $\alpha=0.05$ 水平上并没有显著的变化;在开花期间,种子的含氮水平也没有显著的变化,但叶子可溶性蛋白减少 3%。

5 利用叶绿体基因工程

植物细胞,特别是绿色组织细胞中,叶绿体含

量丰富。叶绿体转化具有位点专一的同源重组、高拷贝数、母系遗传等优点。将外源基因转入细胞器叶绿体,也是实现外源基因在植物宿主体内高效表达的途径之一。该方法有很多优点,包括外源蛋白表达量高,没有后生效应,转基因的叶绿体呈现出母性遗传,转入的基因很少有机会逃逸到环境中去,不会发生通过花粉传播而引起基因污染问题,有利于环境的保护^[25]。Jeffery 等^[26]用人生长素基因转化烟草的叶绿体,表达量达到可溶性蛋白的 7%,比转入核基因组方法提高了 300 倍,且叶绿体表达外源蛋白没有复杂的蛋白修饰过程,却能够正确折叠成有生物活性的蛋白,是一种极有前途的提高外源蛋白表达的新方法。但这种方法在非绿色组织如果实、块茎或其他贮藏器官中的表达量较低,近年来在这些方面的研究也取得一些进展,Ruf 等^[27]采用一种新的质体转化系统,转化西红柿叶片中的叶绿体,提高了外源蛋白在果实中的表达水平,表达量达到叶绿体可溶性蛋白的 50%左右。

6 其他策略

提高外源蛋白在宿主体内的表达水平方法还有很多。对这些方法的深入研究,必将推动植物基因工程的进一步发展。

早在 1994 年,Conceicao 等^[28]就发现,多拷贝 DNA 表达量异常的高,但多拷贝的 DNA 容易引起 DNA 之间的相互干扰和基因沉默。

外源基因在受体植物的基因组内插入位点不同造成转录活性的差异,这就是所谓的“位置效应”。为了消除位置效应,构建表达载体时通常会应用核基质结合区以及定点整合技术。真核细胞含有核骨架,研究发现,附着在核骨架上的染色体往往是活跃转录的基因,将细胞核中核基质结合序列(matrix attached region, MAR)置于目的基因的两侧,构建成 MAR-Gene-MAR,可创建一个独立的结构域,在转基因烟草中,用酵母 MAR 序列可使表达量提高 12 倍,如果使用烟草的 MAR 序列,可使表达量提高 60 倍^[29-31]。目前还没有找到大幅提高外源基因与植物核基因组之间同源重组频率的方法,限制了定点整合技术在植物基因工程中的应用。

细菌、植物、动物各有自己偏好的密码子,对已知氨基酸序列的蛋白质来说,可以通过计算机优化密码子,使用宿主偏好的密码子,也可以提高外源蛋白的表达水平。Elizabeth 等^[32]将抗生物素蛋白密码子进行优化设计,使用玉米偏好的密码子,使抗

生物素蛋白在玉米种子中的表达量占可溶性总蛋白 2% 以上, 实现抗生物素蛋白的商业化生产。

基因的排列也影响蛋白质的表达水平。在细菌中, 利用双顺反子的结构提高重组人粒细胞-巨噬细胞集落刺激因子(hGM-CSF)蛋白的表达水平, 将表达量从 5% 提高到 30%^[33]。对多价载体来说, 基因间的排列也会影响基因表达的效率。

Poly(A)对成熟 mRNA 的稳定性及促进其翻译有重要的作用。新霉素磷酸转移酶 II (*npt II*) 的 3' 末端序列可以显著提高 *npt II* 基因在转化植物中的表达。对马铃薯蛋白酶抑制剂 II (PI- II) 研究发现, 聚腺苷酸信号的下流 100 bp 左右有一段 DNA 序列, 它对基因的高效表达是必须的, 这段序列含有一段 8 bp 的保守序列 CGTGT TTT 并广泛存在动植物基因的 3' 末端^[34]。

总之, 外源基因在宿主体内表达受多种因素影响, 外源基因在基因组中的整合位置, 外源基因的转录、翻译、加工以及分泌等过程都会影响该蛋白最终在宿主体内的积累, 必须根据研究目的综合运用以上的策略才会获得最佳效果。

参考文献

- [1] Reisdorf-Cren M, Carayol E, Terce-Laforgue T, et al. A novel HMG A-like protein binds differentially to the AT-rich regions located in the far distal and proximal parts of a soybean glutamine synthetase gene (*GS15*) promoter [J]. *Plant Cell Physiol*, 2002, 43 (9):1006-1016.
- [2] Cahoon E B, Marillia E F, Stecca K L, et al. Production of fatty acid components of meadowfoam oil in somatic soybean embryos [J]. *Plant Physiol*, 2000, 124(1):243-251.
- [3] Li Y K (李一琨), Wang J F (王金发). Advances in the studies on plant promoter [J]. *Chin Bull Bot (植物学通报)*, 1998, 15(Suppl): 1-6. (in Chinese)
- [4] Qin W (覃文), Zhu S F (朱水芳). Genetically modified oil crops [J]. *China Oil Fat (中国油脂)*, 2003, 28(5):48-52. (in Chinese)
- [5] Kamo K K. Long-term expression of the *uidA* gene in *Gladiolus* plants under control of either the ubiquitin, *rolD*, mannopine synthase, or cauliflower mosaic virus promoters following three seasons of dormancy [J]. *Plant Cell Rep*, 2003, 21(8):797-803.
- [6] Olga A, Barbara K, Ulrike F, et al. Potato tubers as a biofactory for recombinant antibodies [J]. *Mol Breed*, 1998, 4:313-319.
- [7] Joshi C P. An inspection of the domain between putative TATA box and translation start site in 79 plant genes [J]. *Nucl Acids Res*, 1987, 15(16):6643-6653.
- [8] Kozak M. Completion and analysis of sequences upstream from translational site in eukaryotic mRNAs [J]. *Nucl Acids Res*, 1984, 12(2):857-872.
- [9] Mitsuhashi I, Ugaki M, Hirochika H, et al. Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants [J]. *Plant Cell Physiol*, 1996, 37(1):49-59.
- [10] Qin H M (秦红敏), Guo H N (郭洪年), Jia Y T (贾燕涛), et al. Expression of foreign gene in the whole plant affected by untranslated region of TMV-RNA [J]. *Chin Sci Bull (科学通报)*, 2000, 45 (6):619-621. (in Chinese)
- [11] Xu Y, Yu H, Hall T C. Rice triosephosphate isomerase gene 5' sequence directs β -glucuronidase activity in transgenic tobacco but requires an intron for expression in rice [J]. *Plant Physiol*, 1994, 106(2):459-467.
- [12] Rieping M, Schoffl F. Synergistic effect of upstream sequences CCAAT box elements and HSE sequences for enhanced expression of chimeric heat shock genes in transgenic tobacco [J]. *Mol Gen Genet*, 1992, 23(2):226-232.
- [13] Van Rooijen G J, Goloney M M. Plant seed oil bodies as carriers for foreign proteins [J]. *Biotechnology*, 1995, 13(1):72-77.
- [14] Hayashi M, Toriyama K, Kondo M, et al. Accumulation of a fusion protein containing 2S albumin induces novel vesicle in vegetative cells of *Arabidopsis thaliana* [J]. *Plant Cell Physiol*, 1999, 40(3): 263-272.
- [15] Sojikul P, Buehner N, Mason H S. A plant signal peptide-hepatitis B surface antigen fusion protein with enhanced stability and immunogenicity expressed in plant cells [J]. *Proc Natl Acad Sci*, 2003, 100(5):2209-2214.
- [16] Sijmons P C, Deller B M, Schrammeijer B, et al. Production of correctly processed human serum albumin in transgenic plants [J]. *Biotechnology*, 1990, 8:18-22.
- [17] Munroe W D, Pelham H A. C-terminal signal prevents secretion of luminal ER proteins [J]. *Cell*, 1987, 48:899-907.
- [18] Andres D A, Rhodes J D, Meisel R L, et al. Characterization of the carboxyl-terminal sequences responsible for protein retention in the endoplasmic reticulum [J]. *J Biol Chem*, 1991, 266 (22): 14277-14282.
- [19] Udo C, Ulrike F, Olga A, et al. High-level and stable accumulation of single-chain Fv antibodies in plant storage organs [J]. *J Plant Physiol*, 1998, 152:708-711.
- [20] Bourque J E. Antisense strategies for genetic manipulations in plant [J]. *Plant Science*, 1995, 105:125-149.
- [21] Kohno-Murase J, Murase M, Ichikawa H, et al. Effects of an antisense *napin* gene on seed storage compounds in transgenic *Brassica napus* seeds [J]. *Plant Mol Biol*, 1994, 26(4):1115-1124.
- [22] Alain G, Marc V M, Geert A. Co-introduction of an antisense gene for an endogenous seed storage protein can increase expression of a transgene in *Arabidopsis thaliana* [J]. *FEBS Letter*, 1999, 456(1): 160-164.
- [23] Anthony J K, Rudolf J, Eliot M H. Cosuppression of the subunits of β -conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum derived protein bodies [J]. *Plant Cell*, 2001, 13:1165-1178.
- [24] Staswick P E, Zhang Z, Clemente T E, et al. Efficient down-regulation of the major vegetative protein genes in transgenic

- soybean does not compromise plant productivity [J]. *Plant Physiol*, 2001, 127:1819-1826.
- [25] Xiao N Z(肖乃仲), Bai Y F(白云峰), Liu J X(刘锦绣), et al. Plant as bioreactor for the production of pharmaceutical proteins [J]. *Hereditas(遗传)*, 2003, 25(1):107-112.(in Chinese)
- [26] Jeffery M S, Bradley G, Julie G, et al. High yield production of a human therapeutic protein in tobacco chloroplasts [J]. *Natl Biotechnol*, 2000, 18:333-338.
- [27] Ruf S, Hermann M, Berger I J, et al. Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit [J]. *Natl Biotechnol*, 2001, 19(9):870-875.
- [28] Conceicao A S, Van Vliet A, Krebbers E. Unexpectedly higher expression levels of a chimeric 2S albumin seed protein transgene from a tandem array construction [J]. *Plant Mol Biol*, 1994, 26(3): 1001-1005.
- [29] Allen G C, Hall G E, Childs L C, et al. Scaffold attachment regions increase reporter gene expression in stably transformed plant cells [J]. *Plant Cell*, 1993, 5:603-613.
- [30] Allen G C, Hall G E, Michalowski S, et al. High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco [J]. *Plant Cell*, 1996, 8:899-913.
- [31] Spiker S, Thompon W F. Nuclear matrix attachment regions and transgene expression in plants [J]. *Plant Physiol*, 1996, 110:15-21.
- [32] Elizabeth E H, Derrick R W, Sheila M, et al. Commercial production of avidin from transgenic maize: characterization of transformation, production, processing, extraction and purification [J]. *Mol Breed*, 1997, 3:291-306.
- [33] Liu L(刘丽), Zheng L(郑莉), Xie B S(谢宝树), et al. Expression level increase of human granulocyte-macrophage colony stimulating factor by in *E. coli* using two-cistron system [J]. *J N. Bethune Univ Med Sci(白求恩医科大学学报)*, 1996, 22(4):336-338.(in Chinese)
- [34] Wang G L(王关林), Fang H J(方宏筠). *Plant Gene Engineering [M]*. Second ed. Beijing: Science Publishing Company, 2002. 680. (in Chinese)