

大豆种子贮存蛋白基因及其遗传转化的研究进展^{*}

谈建中¹ 楼程富²

(1. 苏州大学生物技术学院, 苏州 215151 2. 浙江大学蚕蜂科学系, 杭州 310029)

STUDY ON SEED STORAGE PROTEIN GENE AND ITS GENITIC TRANSFORMATION IN SOYBEAN

Tan Jianzhong¹ Lou Chengfu²

(1. *Suzhou University of Bioelectronics College Suzhou, 215151*
2. *Zhejiang University of Sericulture Department, Hangzhou, 310029*)

种子蛋白包括贮存蛋白、结构蛋白、酶蛋白及其他功能性蛋白等。其中,贮存蛋白的含量最为丰富,占种子总蛋白的 70–80%。种子贮存蛋白基因具有特异性表达的功能,在种子成熟的中后期,瞬时性地大量合成并贮存于蛋白体中。因此,种子贮存蛋白基因是研究高等植物基因表达调控的极好材料。同时,种子贮存蛋白又是粮食或饲料的重要蛋白质源,在研究基因表达调控的基础上,利用基因操作技术,将种子贮存蛋白基因或经过人工改造后再转移到同种或异种植物中,可望改良种子蛋白品质,增加蛋白产量或生产有用物质。

1 大豆种子贮存蛋白质的种类

大豆 (*Glycine max*) 种子富含蛋白质,一般品系含量 40% 左右,高蛋白品系可达 45%,野生大豆 (*Glycine soja*) 最高可达 50%。根据溶解性的不同,大豆种子蛋白可分为溶于水的清蛋白 (albumin)、溶于盐溶液的球蛋白 (globulin)、溶于醇溶液的醇溶蛋白 (prolamine) 和溶于酸或碱溶液的谷蛋白 (glutelin)。上述四种蛋白质占大豆种子干重的平均含量分别为 28.56%、7.50%、5.71% 和 1.81%,但不同品种之间有一定的差异^[1]。大豆种子贮存蛋白都能用盐溶液提取,当用密度梯度离心时,可分离到沉降系数为 11S、7S 和 2S 的三个主要组分,分别称为 11S 球蛋白 (大豆球蛋白)、7S 球蛋白 (主要成分为 β -伴大豆球蛋白) 和 2S 球蛋白。

1.1 大豆球蛋白

^{*} 收稿日期 1999-02-22
Received on Feb. 22, 1999

大豆球蛋白 (glycinin)有 8种酸性亚基 (A_{1a} A_{1b} A_2 A_3 A_4 A_5 A_6 A_7)和 5种碱性亚基 (B_{1a} B_{1b} B_2 B_3 B_4),两类亚基以特定组合、通过二硫键结合成亚基复合物。迄今已知存在着 $A_{1a}B_{1b}$ A_2B_{1a} $A_{1a}B_3$ $A_5A_4B_3$ 和 A_3B_4 等几种亚基复合物。由于复合物中的某些亚基含有比其他亚基较多的含硫氨基酸,如 A_3B_4 $A_5A_4B_3$ 复合物都只含有 1个含硫氨基酸,而 A_2B_{1a} 复合物含有 7- 8个。因此,大豆球蛋白的含硫氨基酸含量可以通过遗传工程或种质资源选择而得到提高^[2]。

1.2 β -伴大豆球蛋白

β -伴大豆球蛋白 (β -Conglycinin)是 7S球蛋白的主要组分,由 α' 、 α 和 β 三种亚基所构成。与大豆球蛋白相比,含硫氨基酸含量更少,特别是不含半胱氨酸。由于大豆球蛋白与 β -伴大豆球蛋白都存在于种子子叶的蛋白体中,因此,11S球蛋白与 7S球蛋白的比值,与大豆种子营养价值有密切关系^[3]。

1.3 2S大豆球蛋白

2S球蛋白并非均一组分,而是由物理化学性质相异而又比较接近的 5种亚组分所构成,作为大豆种子贮存蛋白的一个小组分,与主要贮存蛋白(11S球蛋白、7S球蛋白)最明显的差别是含有蛋白酶抑制剂,且含硫氨基酸含量特别丰富^[4]。

在上述三种球蛋白中,主要成分是 11S球蛋白,占种子贮存蛋白的 70- 80%。其次为 7S球蛋白,一般占 10- 15%。2S球蛋白为 6- 15%,变幅较大。不同类型大豆的球蛋白组成有很大差别,从野生、半野生的栽培大豆的演化趋势表现为 11S球蛋白逐渐增加,7S球蛋白逐渐减少,而 2S球蛋白在含量上表现为多样性^[5]。

2 大豆种子贮存蛋白基因的结构特征

2.1 大豆球蛋白基因

自从 Marco^[6]等首次克隆大豆球蛋白亚基 A_2B_{1a} 基因的部分序列以来,其它亚基如 $A_{1a}B_{1b}$ $A_{1a}B_3$ A_3B_4 $A_5A_4B_3$ 及 A_2B_{1a} 的 cDNA 序列或基因全序列也相继被克隆^[7- 11]。DNA 序列分析及相关研究表明,大豆球蛋白基因具有若干相同的结构特征。

首先,大豆球蛋白基因属于多基因家族。Morero 等认为大豆球蛋白前体多肽至少有七个不同的基因编码。另外,由于大豆球蛋白存在 8种酸性亚基和 5种碱性亚基,而这些酸性亚基或碱性亚基的一级结构之间存在着保守的同源序列,编码这些蛋白质的基因也比较相似,从而形成许多分子种。其次,大豆球蛋白亚基基因都含有三个内含子,其中一个较大两个较小。如 A_2B_{1a} 亚基前体基因三个内含子大小分别为 238bp、292bp 及 624bp。这些内含子的插入位点没有一定的规则可循,但内含子边界序列符合 GT... AG 拼接原则。在信号肽 (18- 24个氨基酸)编码序列的 5'端上游,存在着 CAAT box、TATA box 及 Legumin box 等顺式作用元件,它们对大豆球蛋白基因的特异性表达具有重要作用。

2.2 β -伴大豆球蛋白基因

β -伴大豆球蛋白的 α 、 α' 和 β 亚基之间具有较高的同源性^[12- 13]。与大豆球蛋白相似, β -伴大豆球蛋白基因也属于多基因家族,每个基因族中又有多个成员。根据 Tierney 等^[14]的分析, α' 和 β 亚基基因在基因组中分别有 1- 2 和 8- 13 个拷贝数。其次,从大豆基因组 DNA 文库中克隆的 β -伴大豆球蛋白 (α' 亚基)基因在编码区域内存在 5 个内含子,与大豆球蛋白基因的内含相比,插入片段较小,仅 80- 200bp。第三,在 α 亚基基因的 5'端侧翼区

域,除了一般的 TATA box 和 CAAT box 外,在转录起始位点上游约 - 560 bp 处,存在一个保守序列 GTGGATAG,非常类似于 SV40 核心增强子。并且,这样的序列同样存在于菜豆球蛋白 β 型亚基因中^[15],推测这一保守序列对 β 伴大豆球蛋白基因的表达可能有着重要作用,但目前尚不清楚它的确切功能。

3 大豆种子贮存蛋白基因的表达调控

真核生物基因的表达调控取决于顺式作用元件与反式作用因子相互作用的结果。顺式作用元件是存在于基因调控区域内的特定核苷酸序列,反式作用因子是认识这些序列并能特异性结合的蛋白质因子,若能探明这些因子的作用规律就可以人为地控制基因的表达。

3.1 顺式作用元件的鉴定

Kitamura^[16]克隆了 A_2B_{1a} 与 $A_{1a}B_{1b}$ 亚基因相连的侧翼序列,并对 A_2B_{1a} 2.2 kb 的启动子进行一系列缺失试验,结果表明 A_2B_{1a} 基因 5'端至上游 - 903 的区域包含了启动子的主要部分,在 - 903~ - 427 区域内存在着调控基因表达量的 DNA 序列。Iida 等^[17]的研究证实, $A_{1a}B_{1b}$ 亚基因 3'端终止区域也与基因的表达活性有关;在启动子的 - 620~ - 380 区域内存在着正调控元件,而 *Lgeumin* box 是大豆球蛋白基因表达的负调控元件。在大多数豆科作物的种子贮存蛋白(豆球蛋白)基因启动子中都含有 5'-CATGCAT-3' 序列,这些片段对基表达量的调控具有重要作用,并且,在其更远的上游还存在其他的调控元件^[18]。此外,内含子对基因的表达量或特异性表达也具有一定的调控作用。Itoh 等(1993)在 A_2B_{1a} 亚基因的启动子下游,分别连接了 A_2B_{1a} 的 cDNA 和基因组 DNA,两者在转基因烟草种子内虽然都能准确表达,并产生成熟形态的大豆球蛋白,但它们的表达量存在明显差异,含 cDNA(缺少内含子序列)的融合基因的表达量只有 1/5 左右。

相对于大豆球蛋白基因顺式作用元件,对反式作用因子的研究报导较少。Itoh^[19]发现了几个新的反式作用因子——胚转录因子,它们能够特异性地结合到 A_2B_{1a} 基因 5'侧翼序列的三个亚区域 - 653~ - 527 - 526~ - 422 - 427~ - 321,这些因子的作用是组织特性的,只在胚发育的早、中期(授花后 24 天左右)出现。这些胚转录因子与 - 657~ - 327 调控区域内的共有基序 5'-ATA/TATTT CN- /CTA-3'之间具有相互作用。并且,在大豆球蛋白另一亚基 $A_{1a}B_{1b}$ 基因的 5'侧翼序列中也存在这样的基序。这暗示在大豆种子发育过程中,可能是由于这些胚转录因子与保守基序的相互作用而激活大豆球蛋白基因的转录。

为了鉴定 β 伴大豆球蛋白基因的重要顺式作用元件,Chen 等^[20,21]对 α' -亚基的启动子进行了一系列的缺失突变,结果表明在转录起始位点上游位于 - 257~ - 159 之间的 DNA 序列对 α' -亚基基因的表达具有重要的调控功能。在 - 257~ - 159 之间的 DNA 序列内,发现存在着 2 个 28bp 的重复片段—— R_y 元件,其中含有 4 个 6bp 的重复序列 (AA/CCCCA)。28bp 的 R_y 元件也存在于 β -亚基基因、菜豆球蛋白 β 亚基基因及小麦的麦谷蛋白基因中。进一步研究表明 - 78~ - 257 内 170bp 的 DNA 序列可能是一个强的顺式作用元件,它控制基因的种子特异性和发育特异性表达。这些种子特异性和发育调控元件 (SDRE)通过与其他顺式作用元件的协同作用,调控种子特异性基因的表达,而 R_y 元件在这过程中起着关键的作用^[22]。

Chamberlan 等^[23]首次证实,大多数种子豆球蛋白(如大豆球蛋白)基因中存在的 Legumin box 对种子豌豆球蛋白(如 β 伴大豆球蛋白)基因的转录也具有调控作用。Bernier 等^[24]的研究进一步表明豆科植物的几种 Legumin box 作为顺式作用元件,都能提高 α' 基因在转基因烟草胚中的转录水平。

Allen 等^[21, 25]分离到一个蛋白因子能够与 α' 亚基基因启动子内一个 70bp 的区域特异性结合,该区域包含了上述 4 个 6bp 的重复序列。同时,还发现另外若干个蛋白因子可以结合在 -300 和 -860 等几个位点。并且,这些具有 DNA 结合活性的蛋白因子只有在种子 α' 和 β 亚基基因开始转录时才检测到,因而推测 α' 和 β 亚基基因的表达,既存在两者共有的调控因子,也存在各自独特的调控因子,这些反式作用因子特异性地结合于结构基因 5'侧翼区域的不同序列(顺式作用元件),在转录阶段起调控作用。

3.2 大豆种子贮存蛋白基因的表达调控

大豆球蛋白和 β 伴大豆球蛋白具有特异性表达的特征,即在大豆种子发育的特定阶段才大量合成和积累,一般品系在授粉后 15-20 天就可检测到 mRNA 的积累,20-30 天时达到最大。其中,大豆球蛋白的积累比 β 伴大豆球蛋白稍迟一些,而 β 伴大豆球蛋白中 α 亚基的积累又早于 β 亚基。另外, β 亚基积累水平受硫营养等环境因子的影响较大。

在转基因矮牵牛和烟草中,只有在种子中才能检测到 α' 和 β 亚基基因的表达(mRNA),而在叶片中检测不到其存在。从表达时期看, β 伴大豆球蛋白也受阶段性发育调控,即在种子发育的早中期表达,并且能正常装配到蛋白体中。这表明 β 伴大豆球蛋白基因的顺式作用元件也能被烟草或矮牵牛细胞中的调控因子所识别,这些因子存在于大豆和烟草或矮牵牛种子的相似区域。虽然在非种子组织中也能检测到基因的表达(mRNA),并且各组织与种子之间没有大的差异,但蛋白质的积累水平依次是种子 > 花组织 > 叶 > 根、茎,其中种子与叶片相差几十倍,其原因在于不同组织之间蛋白质的稳定性存在差异^[21, 26, 27]。

4 大豆种子贮存蛋白基因的遗传转化

4.1 基因的修饰及遗传转化

大豆种子贮存蛋白基因在转基因植物中的准确表达,为研究蛋白质翻译后的修饰加工及稳定性以及通过修饰特殊位点的氨基酸来改变蛋白质功能提供了新的途径。研究表明,大豆等豆科作物的种子贮存蛋白缺乏含硫氨基酸,而大豆球蛋白前体存在一个可变区域,可以插入含 4 个甲硫氨酸的寡肽,构建的表达载体也能在大肠杆菌中合成被修饰的多肽。在转基因烟草的叶片和茎中,同样可检测到天然或修饰的大豆球蛋白,两者的表达量未见差异。不仅如此,在各种组织中积累的两种蛋白质都具有准确的分子量,能加工为成熟型多肽,在种子中还可装配成六聚体^[28]。这说明有可能通过基因工程培育全新的大豆品种,生产营养价值和功能特性更优的大豆蛋白。

Ustumi 等^[29]还尝试用马铃薯表达大豆球蛋白。他们用马铃薯块茎蛋白(potatin)基因的启动子及 NOS 的 3'末端分别与大豆球蛋白 A1aB1b 前体 cDNA 和它的修饰序列(IV + 4Met V+ 4Met)构建了二个融合基因,结果两者都能在转基因马铃薯块茎的薄壁细胞中特异性表达,得到了天然的和修饰的大豆球蛋白基因的转录产物。这说明 Potatin 启动子可使大豆球蛋白基因在马铃薯块茎中特异性表达,并且, IV+ 4Met 和 V+ 4Met 的修

饰不会影响球蛋白在薄壁组织中的生物合成和翻译后的加工过程,这将为生产高经济价值的蛋白质提供新的途径

β 伴大豆球蛋白与大豆球蛋白相比,含硫氨基酸更少,尤其不含半胱氨酸,因此,应用定点突变的方法,修饰编码 β 伴大豆球蛋白基因的核苷酸序列以增加含硫氨基酸密码子,对于改良大豆品质同样具有重要意义。Lelievre 等^[30]在体外表达正常的和修饰的 β 伴大豆球蛋白亚基,除了 NH_2 -端增加一个甲硫氨酸外,体外合成的 α' 和 β 亚基与从大豆种子分离到的亚基具有相同的分子量,蛋白质特征也比较相似,这些亚基因能装配到多聚体中。但装配过程因修饰区域不同而有差别, COOH -端高度保守的疏水区域被修饰后,与未修饰的相比,亚基不能装配或被装配的比例极低。相反,在 NH_2 -端的亲水区域修饰后,亚基装配到多聚体的比例大于或者等于未修饰的,这表明 β 伴大豆球蛋白的 NH_2 -端亲水区域更容易进行结构性修饰,这是应用基因工程改良大豆品质的理论依据。

4.2 大豆总 DNA 的遗传转化

雷勃钧等^[31-32]应用花粉管通道技术将野生大豆的总 DNA 导入栽培大豆中,在鉴定的转基因植株中,有一个株系的蛋白质含量比标准品种约提高 10%,并且与大豆加工品质密切相关的 11S 球蛋白组分所占比例超过了 70%,另有 3 个株系的蛋白质和脂肪总含量达到 66% 以上^[31-32]。此外,用直接转化法将大豆 DNA 导入茄子及玉米^[33]、用花粉管通道技术将花生总 DNA 导入栽培大豆^[34]或应用微注射技术将玉米和大豆的 DNA 注射到另一个栽培大豆^[35]等的转基因研究表明,导入后代在蛋白质和赖氨酸含量以及很多农艺性状方面产生了可遗传的变异。

综上所述,有关大豆种子贮存蛋白生物合成、基因结构及其表达调控的研究均取得了较大的进展,目前已开始应用于农作物的遗传转化方面。由于单子叶和双子叶植物的种子蛋白中各种氨基酸的成分和比例不完全相同,而且,其中必需氨基酸的含量也不一样,因此,通过植物转基因技术,可以将双子叶植物如大豆种子蛋白质基因转移到一些重要的禾谷类作物中去,以提高它们的赖氨酸含量。同样,也可以将单子叶植物如玉米的贮存蛋白基因转移到大豆或其他双子叶植物中去,培育含硫氨基酸含量丰富的转基因大豆品种。在上述转基因大豆的种子蛋白中^[31-35],大豆球蛋白或必需氨基酸含量的增加,显示了种子贮存蛋白基因的遗传转化在改良作物品质方面的应用前景。可以相信,随着对种子蛋白基因表达调控机制的深入研究和植物转基因技术的不断完善,种子贮存蛋白基因在作物品质改良和特殊蛋白质生产方面的应用将会取得更大的进展。

参 考 文 献

- 1 周新安等,大豆科学,1992,11: 191- 197
- 2 赵文明,种子蛋白质基因工程,陕西科学技术出版社,1995,1- 31,97- 107
- 3 林忠平等,大豆科学,1985,4: 327- 336
- 4 张崇本等,生物化学杂志,1991,7: 230- 236
- 5 雷勃钧等,大豆科学,1984,3: 36- 40
- 6 Marco YA, et al., J Biol Chem, 1984, 259: 13436- 13441
- 7 Utsumi S, et al., Agric Biol Chem, 1987, 51: 3267- 3273

- 8 Fukazawa C, et al., *Nucleic Acids Res*, 1987, 15 8117
- 9 Cho T J, et al., *Plant Cell*, 1989, 1 339– 350
- 10 Momma T, et al., *Eur J Biochem*, 1985, 149: 491– 496
- 11 Xue ZT, et al., *Plant Mol Biol*, 1992, 18 897– 908
- 12 Schuler MA, et al., *Nucleic Acids Res*, 1982, 10 8225– 8244
- 13 Schuler MA, et al., *Nucleic Acids Res*, 1982, 10 8245– 8261
- 14 Tierney M, et al., *Planta*, 1987, 172 356– 363
- 15 Doyle JJ, et al., *J Bio Chem*, 1986, 26 9228– 9238
- 16 Kitamura Y, et al., *Nucleic Acids Res*, 1990, 18(4): 339– 350
- 17 Iida A, et al., *Plant Cell Reports*, 1995, 14 539– 544
- 18 Lelievre M, et al., *Plant Physiology*, 1992, 98 387– 391
- 19 Itoh Y, et al., *Mol Gen Genet*, 1994, 243 353– 357
- 20 Chen ZL, et al., *EMBO J*, 1988, 7 297– 302
- 21 Chen ZL, et al., *植物基因工程研究*, 北京大学出版社, 1993, 116– 137
- 22 Fujiwara T, et al., *Plant Mol Biol*, 1994, 24 261– 272
- 23 Chamberland S, et al., *Plant Mol Biol*, 1992, 19 937– 949
- 24 Benier F, et al., *Proceedings of the Fourth International Workshop on Seeds: basic and applied aspects of seed biology*, Angers, France, 20– 24 July 1992, Volume 1, 1993, 13– 20
- 25 Allen RD, et al., *Plant Cell*, 1989, 1 623– 631
- 26 Naito S, et al., *Plant Mol Biol*, 1988, 11 119– 123
- 27 Lawton MA, et al., *Plant Mol Biol*, 1987, 9 315– 325
- 28 Utsumi S, et al., *Plant Science Limerick*, 1993, 92 191– 202
- 29 Utsumi S, et al., *Plant Science Limerick*, 1994, 102 181– 188
- 30 Lekievre JM, et al., *Plant Mol Biol*, 1992, 18 259– 274
- 31 雷勃钧等, *中国科学(辑)*, 1994, 25 576– 601
- 32 雷勃钧等, *大豆科学*, 1995, 14(3): 203– 207
- 33 雷勃钧等, *大豆科学*, 1990, 9(3): 262– 263
- 34 许守民等, *大豆科学*, 1995, 14(4): 316– 320
- 35 张国栋等, *大豆科学*, 1994, 13(3): 268– 274