

大豆贮藏蛋白基因及其表达调控研究进展^{*}

杨辉霞^{1,2} 王芳^{1,2} 单雷² 毕玉平²

(1. 山东师范大学生命科学学院, 济南 250014; 2. 山东省农业科学院, 济南 250100)

摘要 简要介绍了大豆的两类主要贮藏蛋白大豆球蛋白和 β -伴大豆球蛋白基因的组织、结构和表达调控的研究进展。编码大豆球蛋白 5 种亚基的基因均已被克隆, Gy₁ 和 Gy₂ 紧密连锁, Gy₃、Gy₄ 和 Gy₅ 分别位于不同染色体上, 每个基因都包含四个外显子和三个内含子, 其外显子的同源性很高, 其调控序列也具较强的保守性。 β -伴大豆球蛋白的基因为至少 15 个成员的大家族, 至少分布在两条染色体上, 主要集中于三个区域, 它们之间的同源性很高, 但各亚基都有其独特的表达调控方式。

关键词 贮藏蛋白; 大豆球蛋白; β -伴大豆球蛋白; 基因表达调控

中图分类号 S 565.103 文献标识码 A 文章编号 1000-9841(2003)04-0296-05

大豆种子贮藏蛋白的含量远远高于其他作物, 占种子干重的 40% 以上, 是人类食物和动物养殖的主要植物蛋白来源, 具有极高的经济价值^[1,2]。同时, 发育中的大豆种子也是研究高等植物基因表达调控的极好系统, 因为开花这一过程会激发一些组织特异性基因的表达, 导致种子中少数几种贮藏蛋白的大量合成。有人估算一粒大豆种子在发育过程中一天内就合成 1mg 蛋白质^[3]。大豆贮藏蛋白中 11S 大豆球蛋白(glycinin)和 7S β -伴大豆球蛋白(β -conglycinin)占总蛋白的 70%, 其它蛋白还包括脲酶(urease)、植物凝集素(phytochmaglutinins)和胰蛋白酶抑制剂(trypsin inhibitor)等^[4-6]。贮藏蛋白的主要功能是提供种子萌发时所需的氮源和碳源, 而没有催化活性^[7,8]。

1 大豆球蛋白基因及其表达调控

1.1 大豆球蛋白(glycinin)

大豆球蛋白是大豆种子的主要贮藏蛋白, 占种子干重的 20% 左右, 积累于子叶细胞的蛋白体中^[9]。从种子中提取的大豆球蛋白为六聚体复合物, 分子量介于 320—375KD 之间。每个六聚体复合物包含两个三聚体(trimer), 每个三聚体由三个单

体(monomer)组成, 每一单体又包含一条酸性多肽链和一条碱性多肽链, 这两条多肽链来自同一前体, 系前体在蛋白体中被剪切形成, 两条多肽链之间由二硫键相连^[10]。现已鉴定出五种单体, 归为两大类(见表 1)^[11]。

1.2 大豆球蛋白基因

大豆基因组中存在多个大豆球蛋白基因(glycinin genes)及大豆球蛋白相关基因(glycinin-related genes), 它们之间具有很高的同源性^[12]。Tae-Ju Cho 等分析了这些基因的遗传规律及组织形式, 认为 Gy₁、Gy₂ 紧密连锁于一个遗传位点, 两基因之间相隔 3kb, Gy₃、Gy₄ 和 Gy₅ 分别位于基因组的 3 个位点, 遵循孟德尔分离规律独立遗传^[8,11]。Nielsen C. Nielsen 等通过异源双链分析(heteroduplex analysis)证明各基因所在区域的同源性也很高。Gy₁/Gy₂ 和 Gy₃ 都与在叶中表达而胚中不表达的基因连锁^[7]。大豆球蛋白基因具有复杂的结构, N. C. Nielsen 等人通过 S1 作图法、R-loop 分析以及测序分析证明各基因均包含四个外显子和三个内含子, 不同基因内含子的保守性较低, 同源性只有 40% 左右, 而外显子的同源性达 80% 以上。5' 一端的 UTR 区(untranslated region)具有顺式作用元件(cis-elements), 其 TATA 盒位于转录起始点上游 25—30

* 收稿日期: 2003-03-19

项目来源: 中国农科院油料所农业部国家重点实验室开放课题 200102 号

作者简介: 杨辉霞(1974—), 女, 发育生物学专业。E-mail: yhx7404@hotmail.com.

bp, CAAT-盒位于上游100bp, 种子蛋白特异的共感序列 (consensus sequence) 也位于5'-端, 如8-bp的5'-C ATGCATG-3'和28-bp的Leg-box序列。3'-端具有多个保守的5'-AATAAA-3'加多聚腺苷

酸尾巴的信号^[13], 各大豆球蛋白基因的高度同源性是多基因家族的共同特征, 据此可以推断这些基因的进化关系。

表1 大豆球蛋白各亚基的特征

Table 1 Characteristics of glycinin subunit

基因 Gene	类型 Group	亚基 Subunit	分子量 Mr	酸性多肽链 Acid polypeptide		碱性多肽链 Basic polypeptide	
				分子量 Mr	等电点 IEP	分子量 Mr	等电点 IEP
Gy1	1	G1(A _{1a} B ₂)	53.624	33.161	5.37	20.481	7.99
Gy2	1	G2(A ₂ B _{1a})	52.445	32.136	5.03	20.327	7.97
Gy3	1	G3(A _{1b} B _{1b})	52.311	331.718	5.09	20.611	5.69
Gy4	2	G4(A ₅ A ₄ A ₃)	61.294	40.687	4.72	20.626	9.90
Gy5	2	G5(A ₃ B ₄)	55.677	36.392	4.97	19.050	9.51

1.3 大豆球蛋白基因的表达调控

D. W. Meinke 等通过 SDS-PAGE 和 Northern 杂交等方法研究了 11S 大豆球蛋白基因的表达规律, 发现子叶中 11S 各亚基的 mRNA 在果针入土 14~18 天(DAP, days after pegging)时可以检测到, 各亚基的出现和积累则在 18~20 DAP, 胚轴中有极少 11S 各亚基和其 mRNA 积累。他们还证明 45~85DAP mRNA 最丰富的时期也是蛋白合成最活跃的时期, 说明基因表达的调控主要发生在转录水平上^[11]Niels C. Nielsen 等也得到同样的结论, 他们还用幼嫩子叶的细胞核进行体外转录实验, 证明大豆球蛋白基因的表达除了受转录水平的调控之外还受转录后水平的影响, 因为特定时期 mRNA 的含量并不总是与转录速率绝对一致^[7]。

大豆球蛋白各亚基基因与其他种子蛋白基因相似, 它们的 5'-UTR 区均具有保守的顺式作用元件 (cis-elements), 包括 28-bp Leg-box、5'-CATG-CATG-3'RY 重复元件以及 CACA 成分等, 对这些元件功能的研究已较为深入。Jean-Marc Lelievre 等人将 5'-UTR 区的一些可能顺式作用元件与 GUS 报告基因连接转化烟草, 证明 5'-CATGCAT-3'元件对于控制基因表达的量具关键作用, 属于增强子成分^[14]。Yoshifumi Itoh 等在研究 A₂B_{1a} 亚基的表达调控时发现包含 Leg-box 在内的 -326~-102 区域和 -657~-327 区域参与基因表达的数量调控, 还鉴定出与反式作用因子 (trans-elements) 相互作用的区域, 同时鉴定了一些与 A₂B_{1a} 的顺式作用元件作用的反式作用因子, 也称 SEF (soybean embryo factor), 这些因子只在发育中期的早些时候出现^[15]。

大豆球蛋白各亚基的 mRNA 是在与内质网结合的核糖体上转译合成蛋白质。与其它贮藏蛋白一样, 首先合成一段带信号肽的多肽链, 信号肽在内质

网腔内被切除, 然后继续翻译加工成为原大豆球蛋白 (proglycinin)。原大豆球蛋白在内质网腔内折叠并形成三聚体, 在内质网小泡内被运送至高尔基复合体中进一步加工, 在高尔基复合体的反面形成前液泡 (prevacuole) 并在特定的 Asp-Gly 处被切开, 产生酸性和碱性多肽链, 两条多肽链由二硫键相连。原大豆球蛋白被切开这一过程对于它形成六聚体贮藏蛋白是必需的^[16]。前液泡与蛋白质贮存液泡融合, 贮藏蛋白在其内沉积^[18]。这种途径是典型的“内质网-高尔基复合体-液泡”内膜加工途径。分析各亚基的序列可以发现其 Asp-Gly 及形成二硫键的半胱氨酸也具保守性。最近, A. J. Kinny 等用转基因沉默手段抑制 β -伴大豆球蛋白 α 亚基的表达, 结果大豆球蛋白的表达量增加, 这也许是一种补偿效应, 其调控机理还不清楚。他们同时发现了正常内膜加工途径以外的蛋白质加工途径, 即该转基因大豆子叶中的蛋白质可以由内质网小泡直接形成蛋白体 (protein body), 然后蛋白体与蛋白质贮藏液泡融合^[19]。原大豆球蛋白在蛋白质贮藏液泡中被切开, 产生酸性和碱性多肽链, 两条多肽链由二硫键相连。

2 β -伴大豆球蛋白基因及其表达调控

2.1 β -伴大豆球蛋白 (β -conglycinin)

β -伴大豆球蛋白也是大豆种子的一种主要蛋白, 占总蛋白的 25%, 在种子成熟过程中迅速积累, 与大豆球蛋白一样贮存于子叶细胞蛋白体中, 种子萌发时作为氮源和碳源。 β -伴大豆球蛋白由多个亚基组成, 主要是 α' 、 α 和 β , 分子量分别为 76KD、72KD 和 52-54KD, 另外还有一些含量较低的 β' 、 γ

和 δ 亚基等^[1]。这些亚基通过相互作用形成组成各异同源或异源三聚体。 β -伴大豆球蛋白是由两类大小悬殊的 mRNA 分子编码的,一类为 2.5kb-mRNA, 编码 α' 和 α 亚基,另一类为 1.7kb-mRNA, 编码 β 亚基。这两类 mRNA 代表不同的基因亚家族 (gene subfamily)。

2.2 β -伴大豆球蛋白基因

编码 β -伴大豆球蛋白基因的数目比大豆球蛋白的要多,是至少包含 15 个成员的大家族,称为 CG-1~CG-15^[2]。这些成员可以分为两组,一组合成 2.5kb-mRNA,一组合成 1.7kb-mRNA。它们位于染色体的不同区域,存在不同程度的同源性。Harada 等分析了这些基因在基因组中的分布情况,证明它们至少分布在基因组的六个区域,占 175kb,大多集中于三个区域,至少分布在两条染色体上,某些区域中有紧密连锁的 β -伴大豆球蛋白基因,如 A 区就包含 5 个连锁的 α'/α 和 β 亚基的基因,这些基因的转录方向有的相同,有的相反,说明各基因都有其独特的调控方式^[20]。 β -伴大豆球蛋白基因也具有复杂的结构。以 CG-4 为例,该基因包含 6 个外显子和 5 个内含子,据其核苷酸序列推导的氨基酸序列与实际测得的 β -亚基的氨基端序列一致。用 CG-4 转化烟草,可检测到胚中特异转录的 1.7kb-mRNA。CG-1 合成 2.5kb- α' 亚基 mRNA,通过对比 CG-1 与 CG-4 序列发现,CG-1 也包含 6 个外显子和 5 个内含子,这两种基因的外显子的同源性可达 85% 以上,而内含子的同源性只有 50%。它们翻译的多肽链的氨基酸序列一致性达 75%。这两个基因的差异在于 CG-1 的 5'-端的第一个外显子中有一段 0.5kb 的编码序列,正是这段序列造成两类 mRNA 的大小悬殊以及 α'/α 和 β 亚基的大小差异。进一步分析表明该片段是由多个高度同源的重复序列构成的,在 CG-4 的第一个内含子中也有一个 27bp 的该重复序列单元。

2.3 β -伴大豆球蛋白基因的表达调控

Meinke 等还研究了 β -伴大豆球蛋白基因的表达调控,发现 α'/α 亚基的 mRNA 在 18-20 DAP 开始积累, β 亚基的 mRNA 则要晚 1-2 周^[4],进一步表明各亚基是由不同的基因独立编码调控。与大豆球蛋白基因一样其调控也是主要发生在转录水平上,同时转录后水平的调控也很重要^[21]。Harada 等人的研究表明 2.5kb 与 1.7kb-mRNA 的积累和降解的时期不同,但这两类基因却在相同时期被激活。

现已鉴定了许多 β -伴大豆球蛋白基因的顺式

作用元件和反式作用因子。以 α 为例其反式作用因子及其结合部位见表 2。Randy D. Allen 等还研究了 SEF₃ 在发育不同时期的 DNA 结合活性,发现只有在种子成熟的中后期才能检测到它与 DNA 的结合活性,表明它对该基因的表达起正调控作用。

与大豆球蛋白的各亚基不同, β -伴大豆球蛋白各亚基基因调控序列之间的保守性较差,这反映在各亚基及其 mRNA 在胚发育的不同阶段积累、在子叶及胚轴中所占比例不同以及受外界硫源及甲硫氨酸缺乏的影响不同等方面^[22]。Elizabeth A. Bray 等研究了 ABA 对体外培养的子叶发育过程中 β -伴大豆球蛋白各亚基的影响,发现 ABA 能增加 β -亚基的含量,而对 α -和 α' -亚基的影响较小^[23]。各亚基的独立调控对于人们利用基因工程手段改变各亚基含量以改良大豆蛋白的品质提供了理论基础。

表 2 β -伴大豆球蛋白 α' 亚基基因的顺式作用元件和反式作用因子

Table 2 Cis-elements and trans-elements of β -conglycinin α' subunit

反式作用因子 Soybean embryo factor	DNA 结合部位 Location of DNA binding site(s)	识别基序 Recognition motif	参考文献 Reference
SEF1	-670 ~ -640 -790 ~ -765	ATATTTAT/AA/T	a b
SEF3	-186 ~ -130	AACCCA.AACCCA	c
SEF4	-115 ~ -105 -340 ~ -325 -690 ~ -670	A/GTTTTTA/G	a

注: a. R. D. Allen and R. N. Beachy, unpublished data

b. Consensus SEF 1 binding site; an analogous binding site was reported by Jofuku etc (1987)

c. Nuclear factor interact with a soybean beta-conglycinin enhancer by R. D. Allen etc (1989)

3 展望

人们已经根据大豆贮藏蛋白基因的结构及其表达调控规律对蛋白质品质的改良进行了许多有益的探索,并取得了一些成功。例如通过对比各亚基的基因序列找到保守区和可变区,然后对可变区尤其是高变区进行碱基替换等遗传操作,增加含硫氨基酸的含量,提高了其营养价值,而不影响蛋白质的折叠过程^[24]。还有人研究了贮藏蛋白基因 5'-端上游控制表达蛋白质量的顺式作用元件,然后对其改造从而提高或抑制某个亚基的表达,达到提高含硫氨基酸含量的目的。大豆的基因组研究也取得了巨大成就,如大豆的 EST 数据库收集了 285,000 条 EST 序列,平均长度大于 400bp,为研究组织或器官的基因

表达奠定了良好基础。大豆的功能基因组学也逐步展开,该项目的目标是利用现代分子生物学技术如DNA微阵列、基因表达的系列分析(SAGE)等建立大豆的“非冗余基因套”(unigene set),全面了解大豆的基因表达谱。^[25]

由于大豆基因组庞大,生长周期长等特点,还有许多问题有待解决,如各顺式作用元件与反式作用因子互作的详细机制,蛋白质合成、装配以及修饰过程和 Related 因子的分析鉴定,提高大豆品质的高效方法等。随着拟南芥、水稻等基因组测序工作的完成,可以借助比较基因组学、功能基因组学等各种手段从整体上研究大豆的生命活动规律,为研究高等植物生命活动规律和利用其他豆科作物奠定基础。

参 考 文 献

- 1 Jeff J. Doyle, Mary A. Schuler, W. Dean Godette et al. The Glycosylated Seed Storage Proteins of Glycin max and Phaseolus vulgaris [J]. J. Biol. Chem. 1986, 261, 9228-9238.
- 2 Hari B. Krishnan, Guoqiao Jiang, Ammulu Hari Krishnan et al. Seed storage protein composition of non-nodulating soybean and its influence on protein quality [J]. Plant science, 2000, 157-199.
- 3 Nilgum Erenken Tumen, Vu Huu Thanh, and N. C. Nielsen. Purification and Characterization of mRNA from Soybean Seeds [J]. J. Biol. Chem. 1981, 256, 8756-8760.
- 4 D. W. Meinke, J. Chen, R. N. Beachy. Expression of Storage Protein Genes during Soybean Seed Development [J]. Planta, 1981, 153: 130-139.
- 5 许月,朱长甫,石连旋,等.大豆种子贮藏蛋白的研究概况 [J]. 大豆科学, 1998, 17: 262-267.
- 6 Linda Walling, Gary N. Drews, R. B. Goldberg. Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels [J]. Proc. Natl. Acad. Sci. 1986, 83, 2123-2127.
- 7 N. C. Nielsen, C. D. Dickinson, Tae-Ju Cho et al. Characterization of the Glycinin Gene Family in Soybean [J]. Plant Cell, 1989, 1, 313-328.
- 8 Tae-Ju Cho, Corinne S. D., N. C. Nielsen. Inheritance and Organization of Glycinin Genes in Soybean [J]. Plant Cell, 1989, 1, 329-337.
- 9 郑易之,何孟元,郝水.大豆子叶中蛋白体的形成与贮藏蛋白质积累的关系 [J]. 植物学报, 1992, 34(8): 641-644.
- 10 Robert W. Yalick. Beta-conglycinin and glycinin in high-protein seeds [J]. J. Agric. Food Chem. 2001, 49, 729-735.
- 11 V. Beilinson, Z. Chen, R. C. Shoemaker et al. Genomic Organization of Glycinin Genes in Soybean [J]. Theor Appl Genet, 2002, 104: 1132-

- 1140.
- 12 Trevor L. Wang, Claire Domoney, Cliff L. Hedley et al. Can we improve the nutritional quality of legume seeds? [J]. Plant physio. 2003, 131, 886-891.
- 13 Shigen Utsumi, Mitsutata Kohno, Tomohiko Mori et al. An alternate cDNA encoding glycinin A1aBx subunit [J]. J. Agric. Food Chem. 1987, 35, 210-214.
- 14 Jean-Marc Lelievre, Luiz O. Oliveira and N. C. Nielsen. 5'-CATG-CAT-3 Elements Modulate the Expression of Glycinin Genes [J]. Plant Physiol, 1992, 98: 387-391.
- 15 Yoshifumi Itoh, Yoshiaki Kitamura, Masaomi et al. Cis-acting Regulatory Regions of the Soybean Seed Storage 11s Globulin Gene and Their Interactions with Seed Embryo Factors [J]. Plant Mol. Bio., 1993, 21: 973-984.
- 16 Jung, R., Scott, M. P., Nam, Y. W., et al. The Role of Proteolysis in the Processing and Assembly of 11S Seed Globulins [J]. Plant Cell. 1998, 10: 343-357.
- 17 Dickinson C. D. Hussein, E.H.A., Nielsen, N. C. Role of post-translational cleavage in glycinin assembly [J]. Plant cell, 1989, 1, 459-469.
- 18 Eliot M. Herman, B. A. Larkins. Protein Storage Bodies and Vacuoles [J]. Plant Cell, 1999, 11: 601-603.
- 19 Anthony J. Kinney, Rudolf Jung, E. M. Herman. Cosuppression of the α Subunits of β -Conglycinin in Transgenic Soybean Seeds Induces the Formation of Endoplasmic Reticulum-Derived Protein Bodies [J]. Plant Cell (2001) 13: 1165-1178.
- 20 John J. Harada, Susan J. Barker, R. B. Goldberg. Soybean beta-conglycinin Genes Are Clustered in Several DNA Regions and Are Regulated by Transcriptional and Posttranscriptional Processes [J]. Plant Cell, 1989, 1: 415-425.
- 21 R. D. Allen, Francois Benier, Philip A. Lessard et al. Nuclear Factors Interact with a Soybean beta-conglycinin Enhancer [J]. Plant Cell. 1989, 1: 623-631.
- 22 Lorraine P. Holowach, James T. Madison, J. F. Thompson. Studies on the Mechanism of Regulation of the mRNA Level for a Soybean Storage Protein Subunit by Exogenous L-Methionine [J]. Plant Physiol. 1986, 80: 561-567.
- 23 Elizabeth A. Bray, Roger N. Beachy. Regulation by ABA of beta-conglycinin expression in cultured developing soybean cotyledons [J]. Plant physio. 1985, 79, 746-750V.
- 24 Tomoyuki Katsube, Andrew Barde, Gidamis, Jirokanamori et al. Modification Tolerability of the Hypervariable Region of Soybean Proglycinin [J]. J. Agric. Food Chem., 1994, 42: 2639-2645.
- 25 VandenBosch KA, Stacey G. Summaries of legume genomics projects from around the globe: community resources for crops and models [J]. Plant Physio, 2003, 131: 840-865.

THE PROGRESS OF THE STUDIES ON SOYBEAN SEEDS PROTEIN GENES AND THE REGULATION OF THEIR EXPRESSION

Yang Huixia^{1,2} Wang Fang^{1,2} Shan Lei² Bi Yuping²

(1. *School of Life Science, Shandong Normal University, Jinan 250014*

2. *Shandong Academy of Agricultural Sciences, Jinan 250100*)

Abstract The progress of the studies on two major soybean storage proteins, their genomic organization and the regulation of their expression are briefly discussed. Five major genes encoding the predominant glycinin subunits have been cloned and studied. Gy1 and Gy2 are tandemly linked. Gy3, Gy4 and Gy5 locate on different chromosomes. Each glycinin gene contains four exons and three introns like genes that encode related proteins in other legumes. The level of identity among the exons of the five genes are very high. Glycinin genes are coordinately expressed. The β -conglycinin gene family contains at least 15 members grouped into two subfamilies and are highly homologous along their entire lengths. The expression of each gene subfamily is independently regulated.

Key words Storage Protein; Glycinin; β -conglycinin; Gene expression regulation

欢迎订阅《中国粮油学报》

《中国粮油学报》是中国粮油学会主办的学术性刊物,已纳入全国食品工业类中文核心期刊,属全国重点期刊之一,并被美国《化学文摘》收录。《中国粮油学报》主要登载谷物、油脂化学方面的学术论文;报道优质粮油品质资源选育、贮藏、加工利用以及品质检测方法方面的研究成果,它对指导粮油学科的发展、提高粮油资源的深度开发利用水平,具有实用价值。

《中国粮油学报》是国内外公开发行的一级刊物,国内统一刊号:CN11-2864/TS,国际标准连续出版物刊号:ISSN 1003-0174。双月刊,逢双月出版,胶版印刷,大16开,88页,每期定价10元,全年定价60元(含邮费)。

地址:北京市宣武区报国寺1号《中国粮油学报》编辑部

邮政编码:100053 帐户:中国粮油学会

开户银行:北京市商业银行报国寺支行 75000

帐号:201050385-70

编辑部联系电话:010-63038765 传真:010-63183883

E-mail: lyxuebao@public.bta.net.cn Lyxuebao@ccoa.info