

贮藏蛋白的基因表达和调控

宋松泉 傅家瑞

(中山大学生物系, 广州 510275)

提 要

种子贮藏蛋白质是一类多聚蛋白质, 由多基因家族编码。贮藏蛋白基因内部和周围具有高度保守的调节序列。其基因表达具有组织特异性, 受植物发育过程、植物激素 (ABA 和细胞分裂素) 和营养成分的调节控制。

关键词 种子, 贮藏蛋白, 基因结构, 基因表达及其调节

贮藏蛋白质 (storage protein) 是一类多聚蛋白质, 是种子的主要成分之一, 是人类和其他动物赖以生存的主要蛋白营养来源。在种子萌发过程中, 贮藏蛋白水解, 为幼苗生长提供氮源和能量。在某种意义上说, 贮藏蛋白的积累, 为种子活力提供了分子基础。此外, 贮藏蛋白本身又是基因活动的直接产物, 容易在各个不同发育时期获得大量的实验材料, 且蛋白种类少, 含量高, 容易分离和鉴定, 这就使得贮藏蛋白基因成为研究高等植物特异基因选择性表达的较好系统。

一、贮藏蛋白基因表达的发育变化

高等植物的胚胎发生有明显的发育特征。根据不同的形态发生和生理活动, 可将大豆的胚胎发育过程分为球形期、心形期、子叶期、成熟期和休眠期五个阶段^[1]。在不同的发育时期, 贮藏蛋白 mRNA 的种类和量都发生变化。将不同发育时期胚细胞的 mRNA 进行电泳分析, 子叶期的 mRNA 看不到分开的带; 到成熟后期除了一条带外, 在成熟中期出现的那些带, 或者已看不到或者已不占优势^[1]。

Goldberg 等^[2]用克隆的种子蛋白 cDNA 定量研究了大豆胚胎发育过程中几种种子蛋白 mRNA 含量的变化 (图 1)。贮藏蛋白 mRNA

从子叶期开始积累, 到成熟中期达到高峰, 然后下降。Sánchez-Martínez 等^[3]用两种孔径的凝胶电泳分析了不同胚胎发生阶段和萌发早期玉米离体胚中合成的多肽。他们把这些多肽描述为三个部分: (a) 胚性部分, 这些多肽只在幼胚和成熟胚中合成, 在萌发早期不合成。 (b) 成熟部分, 这部分多肽不存在于幼胚中, 但在成

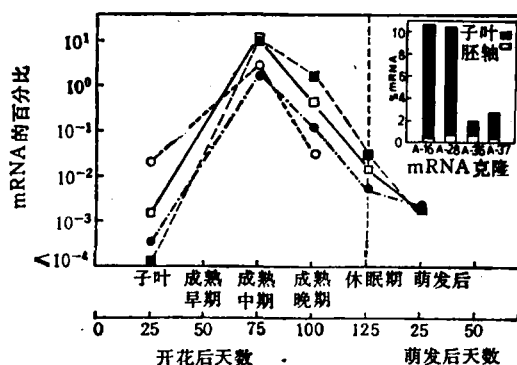


图 1 大豆胚胎中种子蛋白 mRNA 的发育变化

□ A-16, β 伴大豆球蛋白; ■ A-28, 大豆球蛋白; ● A-36, 15kD 蛋白质;
○ A-37, Kunitz 胰蛋白酶抑制剂

熟期间出现。 (c) 萌发部分, 这些多肽在成熟期间不被表达, 在干胚短暂吸水后出现。通过根癌农杆菌 (*Agrobacterium tumefaciens*) 把一种编码 Mr 19000 玉米醇溶蛋白的基因引进矮牵牛属植物。利用玉米醇溶蛋白 mRNA 的

5' 和 3' 探针的 SI 核酸酶图谱证明, 在转基因植物的种子中基因转录被正确地起始和终止。在授粉后 10 天的矮牵牛植物种子中 mRNA 首先被发现, 在授粉后 17 天消失^[4]。稻胚凝集素 (RGL) 存在于胚中, 胚乳中没有测得凝集素活性。水稻开花后 7—21 天胚中 RGL 活性与含量迅速增加、积累, 基本上达到成熟胚的最高水平^[5]。这些事实说明贮藏蛋白基因的表达受发育时期的调节。

另外, 成熟中期高度丰富的 mRNA 在不同的组织部位表现也不同。图 1 内的插图比较了存在于胚轴和子叶中的种子蛋白 mRNA, 所研究的四种 mRNA 占子叶 mRNA 总量的 25%, 但只占胚轴 mRNA 总量的 3%^[4]。Fischer 等的研究表明, 大豆贮藏蛋白 mRNA 在子叶细胞中的浓度为胚轴细胞中浓度的 7—20 倍^[6]。叶子细胞核 RNA 和克隆的种子蛋白 cDNA 探针的杂交研究证明, 叶子细胞核 RNA 中未检测到种子蛋白的转录本^[4]。菜豆蛋白 mRNA 和蛋白质产物在种子中大量存在, 但在叶子中不能检测到。在蚕豆中, 菜豆蛋白唯一地在胚蛋白体中发现。烟草种子贮藏蛋白只存在于胚和胚乳中^[7]。这表明编码贮藏蛋白 mRNA 的那些基因受细胞组织特异性的调节。

二、贮藏蛋白的基因结构

贮藏蛋白由两个或多个亚基组成, 目前只

有少数几种清楚其基因结构。

1. 玉米醇溶蛋白基因

玉米醇溶蛋白 (zein) 由 Z₁ 和 Z₂ 两个家族组成。Z₁ 家族有 4 个亚家族: Z₁A, Z₁B, Z₁C 和 Z₁D。Z₂ 家族有 3 个亚家族: Z₂A, Z₂B, Z₂C。这些亚家族由相应的基因编码 (表 1)。表 1 是来自 cDNA 克隆的杂交资料, 基本上概括了玉米醇溶蛋白基因的特点^[8]。此外, 基因和 cDNA 克隆的比较还证明了玉米醇溶蛋白基因不含插入顺序^[9-11], 其基因长度为 0.75kb, mRNA 长度为 0.75kb, 基因/mRNA 的长度比为 1.0^[9]。

2. 大豆贮藏蛋白基因

大豆种子中, 贮藏蛋白似乎主要包装在 2S, 7S 和 11S 三种复合体中。在 SDS-聚丙烯酰胺凝胶电泳中 7S 蛋白可分为三个亚基, α' (83kD), α (76kD), β (53kD)。11S 蛋白可分为酸性和碱性两类亚基, 但亚基总数至少多达 10 个。酸性亚基 A₁, A₂, A₄ 均约为 37 kD, A₃ 为 42 kD; 碱性亚基 B₁, B₂, B₃ 和 B₄ 均为 20 kD^[12]。

大豆 7S 蛋白的 α 和 α' 基因的 cDNA 克隆分析表明它们是由两个密切相关的多基因家族编码^[13]。在家族内, 各个成员彼此之间有核苷酸同源性。而且在整个 3' 端那一半的编码区域和非编码区域上显示很高的保守性。然而家族之间 (即 α 和 α' 基因之间) 的核苷酸差异也基本发生在 3' 端, 在这两个基因编码的蛋白质

表 1 玉米醇溶蛋白多基因家族的醇溶谷蛋白部分^[11]

亚 家 族	Z ₁ (非还原条件)				Z ₂ (还原条件)		
	Z ₁ A	Z ₁ B	Z ₁ C	Z ₁ D	Z ₂ A(ASG)	Z ₂ B	Z ₂ C
代表每个亚家族的克隆	A ₂₀	A ₃₀	B ₄₉	B ₅₉	B ₃₆	Z ₂₂ A	—
Mr × 1000(kD)	大部分 19	大部分 19 小部分 22	大部分 22 小部分 19	大部分 19	27	15	10
染色体定位	4, 7, 10	4, 7	4	?	?	?	?
优势氨基酸残基	Gln	Gln	Gln	Gln	Pro	Cys	Met
表达的时间选择 (dap)	ca 12	ca 12	ca 18	ca 12	ca 18	ca 18	Ca 18
每个单倍体基因组的数目	<25	<20	<15	<5	2	2	?

Mr: 分子量; dap: 在授粉后的天数。

(α 和 α' 亚基)中,氨基酸的主要差异部分接近羧基末端。 α 和 α' cDNA 彼此之间的核苷酸差异只有 7%。就 mRNA 的结构而言,发现 α 和 α' cDNA 顺序内含有一段 155 个核苷酸的最高保守顺序。它即能和 α 、 α' 亚基 mRNA 杂交,也能和中早熟期出现在种子中的其他 mRNA 杂交。将 α' cDNA 的顺序和 α' 基因组克隆中的 DNA 顺序比较,发现 α' 基因组内含四个内含子。其长度分别为 85, 115, 132 和 40 个核苷酸。

Fisher 等^[6]研究了两个协调表达的 11S (glycinin) 系统的基因组 DNA 克隆,分别编码 11S 蛋白和另一个功能未知的 15.5kD 多肽。发现前者家族约有 3 个非等位基因,后者有 2 个非等位基因。这些基因家族,在胚胎发生期间没有选择性扩增或重排现象。编码 11S 蛋白和 15.5kD 多肽的这两个基因家族之间,于染色体上在 10—15kb 的长度之内不发生连锁关系(也就是说两者的距离起码超过 10—15 kb)。R-环异源双股杂交反应等试验证明,11S 蛋白基因至少含有一个(很可能是 2 个)内含子,然而 15.5kD 多肽基因不含内含子^[6]。

3. 豆球蛋白基因

豆球蛋白由 8 个基因编码。豆球蛋白基因位于染色体 7 上。它含有 3 个内含子 (IVS₁, IVS₂, IVS₃)。其中 2 个内含子存在于豆球蛋白 α 亚基编码区,这 2 个内含子具有 88 个碱基对;另一个内含子存在于豆球蛋白 β 亚基编码区,具有 99 个碱基对^[14]。

三、贮藏蛋白基因表达的调控

贮藏蛋白基因表达的调控是复杂的,包括转录的调节,转录产物的加工调节,转译的调节

以及激素的调节等等。

1. 贮藏蛋白基因内可能的调节顺序

玉米醇溶蛋白基因的调节信号与其他真核生物系统相似,在转录起始位点上游的 30bp 处有一个 TATA 盒,有时两个信号系统相距 10 bp。已经报道在转录起始调节位点之前 900 bp 处还有第二个活性启动子^[8]。在大多数动物中 CAAT 盒存在于 RNA 起始位点上游的大约 70bp 处,在一些植物基因中尚未被发现。当玉米醇溶蛋白基因和一些其他植物基因比较时,在位置 80 附近能够发现一种不同的交感作用顺序 (consensus sequence): $\begin{matrix} C & A_2- & G \\ T & & T \end{matrix}$ $\begin{matrix} CC^{[10]} \\ TT \end{matrix}$ 。这种顺序是否有调节功能还不清楚。它可能代替 CAAT 盒,或者它可能相当于一种增强子顺序的植物版本 (version),因为它具有一些类似于动物核心增强子的交感作用顺序 GTGG $\begin{matrix} A & A & A \\ T & T & T \end{matrix}$ 。尽管 AGGA 盒不存在于所有植物基因中,但是它存在于许多目前已发表的顺序中^[8]。另外,在玉米醇溶蛋白基因终止多肽合成的终止密码子后面 24bp 处有 AATAAA 顺序^[11]。

大豆血红蛋白基因在 5' 端上游 30bp 处有 TATAAA 顺序,90bp 处有 CCAAT 顺序,基因 3' 端上游 30bp 处有 AATAAA 顺序,并含有图 2 所示的 GT/AG 规则的内含子^[15]。

豆球蛋白基因起始密码子之前的核苷酸序列相当于 TATA 盒和 CAT 盒。而且 TATA 盒是在起始密码子 AUG 上游区域的 60bp 处,与其他植物和动物基因启动子区域的位置相符。在豆球蛋白的 3' 端具有 AATAAATAAA 重复排列的多聚腺苷化序列。其内含子的边界

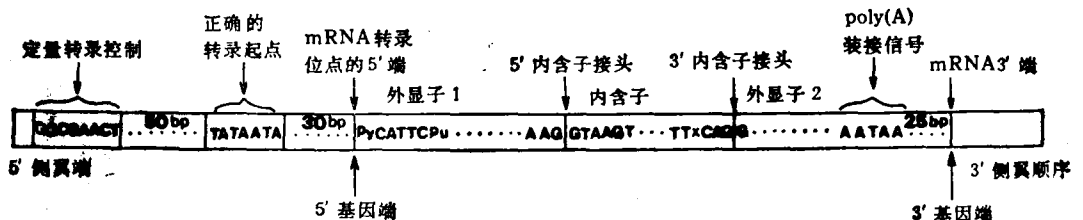


图 2 贮藏蛋白基因内部和周围可能调节的 DNA 顺序

(intron boundary) 也符合 GT/AG 规则^[14]。

2. 贮藏蛋白基因表达的激素调节

许多研究证明在胚胎形成中脱落酸(ABA)阻止过早的萌发,调节贮藏蛋白的合成。在体外培养的油菜籽、大豆和小麦胚中,外源 ABA 促进胚专一的 mRNA 和蛋白质的合成^[16]。

Goday 等^[17]鉴定了一组 23—25kD 的蛋白质,其等电点(pI)范围为 6.2—8.2,这组蛋白质在正常的玉米胚胎形成过程中逐渐积累,在萌发早期消失。在未成熟的胚中 ABA 处理能较早地诱导这些多肽的生成。在发育的培养大豆子叶中已经研究了外源 ABA 对子叶专一基因表达的调节。当未成熟的子叶在修改的 Thompson 培养基中培养时,添加 ABA 导致 β -伴大豆球蛋白 β -亚基浓度的增加,而 β -伴大豆球蛋白的 α' -和 α -亚基的量相对地不受 ABA 处理的影响。fluridon (一种类胡萝卜素生物合成的抑制剂)能降低植物组织中 ABA 的含量,当把它添加到培养基中时,子叶中 ABA 和 β -亚基的含量降低。在体外,用 β -伴大豆球蛋白的抗血清免疫沉淀 RNA 的转译产物时,来自 ABA 培养子叶的 RNA 与对照相比增加前 β -亚基多肽的量。从 fluridon 处理的子叶分离的 RNA 转译产物中,不能检测到前 β -亚基多肽。与对照相比较,ABA 处理的子叶中含有较高的 β -亚基 mRNA 水平,而 fluridon 处理的子叶中则较低^[18]。这就证明在发育的大豆子叶中,外源 ABA 能够调节 β -伴大豆球蛋白 β -亚基的积累。

用 10^{-4} mol/L 苄基腺嘌呤培养南瓜子叶,其 poly(A)⁺mRNA 的量大约是对照的 4 倍以上。子叶中羟基丙酮酸还原酶的活性随苄基腺嘌呤处理浓度的增加(10^{-9} — 10^{-4} mol/L)而提高^[19]。这说明细胞分裂素的作用和基因表达有密切的关系。

3. 营养成分对贮藏蛋白基因表达的影响

在发育的大豆种子中已经研究了硫的缺乏对蛋白质的影响。硫的缺乏使大豆球蛋白的量降低 40%,而 β -伴大豆球蛋白的量反而明显提高,这些蛋白质亚基的组成也受影响。特别是

在缺硫的种子中, β -伴大豆球蛋白的 β -亚基增加了 3 倍,这种增加主要是以 β -球蛋白的 B 异构体积累。B 异构体实际上是缺乏 Met 和 Cys,但保留正常 7S 贮藏蛋白的物理性质^[20]。

Holowach 等^[21]证明当大豆子叶培养在补充 Met 的培养基中时,7S 蛋白 β -亚基和 β -mRNA 是缺乏的。为了进一步研究 Met 的作用机制,他们分别用含和不含 Met 的培养基培养子叶 16 天。4 天后一些子叶从补充 Met 的培养基转移到基本培养基(不含 Met),或者反过来把另一些子叶从基本培养基中转移到补充 Met 的培养基。在基本培养基中,第 4 天能够检测到 β -亚基,而在补充 Met 的培养基中没有 β -亚基存在。当子叶从基本培养基转移到补充 Met 的培养基中时,在 4 天内 β -亚基增加,然后保持恒定(以每个子叶为单位)。这个结果表明 Met 不是通过促进 β -亚基的降解起作用。子叶从补充 Met 的培养基转移到基本培养基后 4 天,子叶含有 β -亚基,这样说明抑制作用是可逆的。在这期间未结合的 Met 从每克鲜重 $7\mu\text{mol/L}$ 降低到 $1.5\mu\text{mol/L}$ 。当用体外翻译系统检测 β -mRNA 时,在不积累 β -亚基的组织中功能 β -mRNA 是缺乏的。在 Met 存在时, β -亚基的 mRNA 大约有一天的半衰期。子叶 mRNA 与互补 β -mRNA 的 cDNA 杂交也证明 1700 核苷酸的 β -mRNA 不存在于补充 Met 的子叶中。这些资料说明营养元素 S 及其含 S 氨基酸能够调节贮藏蛋白基因的表达。至于其他营养元素如 Ca, K, Zn 等在贮藏蛋白基因表达中的作用也值得研究。

四、结 语

种子贮藏蛋白为人类和其他动物提供了主要的蛋白营养物质。然后种子中的必需氨基酸如赖氨酸,色氨酸等含量较低,不能满足人体正常代谢的需要。因此弄清楚贮藏蛋白的基因结构和基因表达的调控机理,利用分子生物学的手段对贮藏蛋白基因进行改造和利用,提高贮藏蛋白的质和量都具有重大的理论意义和现实意义。

但是对贮藏蛋白基因表达的研究尚处于起始阶段,对许多贮藏蛋白基因的结构,以及基因表达的调控机理仍不清楚。应用转化实验研究贮藏蛋白基因在发育过程中特异性表达的有关顺序,把关于控制植物基因表达过程的理论同把基因引入植物细胞并产生可遗传的变异的植物的技能结合起来,必定会在不久的将来为培养和选择具有高质量和数量的贮藏蛋白的植物开辟美好的前景。

参 考 文 献

- 1 李向辉等编著. 植物遗传操作技术. 北京: 科学出版社, 1988; 105
- 2 Goldberg R B, Hoschek G, Ditta G S *et al.* *Developmental Biology*, 1981; 83: 218
- 3 Sánchez-Martínez D, Puigdomènech P, Pagés M. *Plants Physiol*, 1986; 82: 543.
- 4 Ueng P, Galili G, Sapanra V *et al.* *Plants Physiol*, 1988; 86: 1281
- 5 缪国华, 唐锡华. 植物生理学报, 1985; 11(4): 336
- 6 Fischer R L, Goldberg R B. *Cell*, 1982; 29: 65.
- 7 Kublemeier C, Green P J, Chua N H. *Ann Rev Plant*

- Physiol*, 1987; 38: 221
- 8 Heidecker G, Messing J. *Ann Rev Plant Physiol*, 1986; 37: 439
- 9 Hu N-T, Peifer M A, Heidecker G *et al.* *EMBO J*, 1982; 1: 1337
- 10 Messing J, Geraghty D, Heidecker G *et al.* In: Kesuge T, Meredith CP, Hollander A eds, *Genetic engineering of plants*, New York: Plenum, 1983; 211
- 11 Pedersen K, Devereux J, Wilson DR *et al.* *Cell*, 1982; 29: 1015
- 12 Moreira M A, Hermodson M A, Larkins B A, *et al.* *The Journal of Biological Chemistry*, 1979; 254: 9921
- 13 Schuler M A, Schmitt E S, Beachy R N, *Nucleic Acids Res*, 1982; 10: 8225
- 14 赵文明. 生命与化学, 1986; 6(4): 32
- 15 Hyldig-Nielsen J J, Jensen E O, Wiborg O *et al.* *Nucleic Acid Res*, 1982; 10: 689
- 16 Zeevaart J A D, Creelman R A. *Ann Rev Plant Physiol Plant Mol Biol*, 1988; 39: 439
- 17 Goday A, Sánchez-Martínez D, Gomez J *et al.* *Plants Physiol*, 1988; 88: 564
- 18 Bray E A, Beachy P N. *Plants Physiol*, 1985; 79: 746
- 19 Chen C-M, Leisner S M. *Plants Physiol*, 1985; 77: 99
- 20 Gayler K R, Sykes G E. *Plants Physiol*, 1985; 78: 582
- 21 Holowach L P, Madison J T, Thompson J F. *Plants Physiol*, 1986; 80: 561

[本文于1990年4月16日收到, 6月12日修回]

(上接第 176 页)

活力水平的性质,是与它在体内可能参与蛋白质生物合成,催化巯基二硫键交换,形成天然二硫键的功能相一致的。

参 考 文 献

- 1 Epstein C J, Goldberger R F, Anfinsen C B. *Cold Spring Harbor Symp Quant Biol*, 1963; 28: 439
- 2 Wetlaufer D B, Ristow S S. *Annu Rev Biochem*, 1973; 42: 135
- 3 Creighton T E. *Prog Biophys Mol Biol*, 1978; 33: 231
- 4 Goldberger R F, Epstein C J, Anfinsen C B. *J Biol Chem*, 1963; 238: 628
- 5 Venetainer P, Straub F B. *Biochim Biophys Acta*, 1963; 67: 166
- 6 Freedman R B, Brockway B E, Lambert N. *Biochem Soc Trans*, 1984; 12: 929
- 7 Morin J E, Dixon J E. *Meth in Enzymol*, 1984; 113: 541
- 8 Kaetzel C S, Rao C K, Lamm M E. *Biochem J*, 1987;

- 241: 39
- 9 Roth R A, Pierce S B. *Biochemistry*, 1987; 26: 4179
- 10 Bjelland S, Wallevik K, Kroll J *et al.* *Biochim Biophys Acta*, 1983; 747: 197
- 11 Lambert N, Freedman R B. *Biochem J*, 1983; 213: 235
- 12 Tang J G, Wang C C, Tsou C L. *Biochem J*, 1988; 255: 451
- 13 Tang J G, Tsou C L. *Biochem J*, 1990; 268: 429
- 14 Creighton T E, Hillson D A, Freedman R B. *J Mol Biol*, 1980; 142: 43
- 15 Edman J C, Ellis L, Blacher R W *et al.* *Nature*, 1985; 317: 267
- 16 Hawkins H C, Forster S J, Murant S J *et al.* *Biochem Soc Trans*, 1986; 14: 756
- 17 Pihlajaniemi T, Heleakoski T, Tasanen K *et al.* *The EMBO J*, 1987; 6: 643
- 18 Koivu J, Myllyla R, Kivirikko K I. *Biochem J*, 1987; 237: 237
- 19 Yamauchi K, Yamamoto, T, Hayashi H *et al.* *Biochem Biophys Res Commun*, 1987; 146: 1485
- 20 Geetha-Habib M, Noiva R, Kaplan H A *et al.* *Cell*, 1988; 54: 1053

[本文于1990年2月20日收到, 6月15日修回]