

种子贮藏蛋白质表达调控及应用研究进展

杨学军¹, 喻方圆^{1*}, 张欢喜²

(1. 南京林业大学森林资源与环境学院, 南京 210037; 2. 安徽省泾县林业局, 安徽泾县 242500)

摘要: 种子在成熟过程中大量积累贮藏蛋白质, 且不同植物种子中贮藏蛋白质的种类不同。为了解种子贮藏蛋白质的表达调控模式, 对近年来表达调控机制的理论研究和在植物分类研究、分子育种及表达药用蛋白质方面的应用研究进行了综述。理论研究表明, 种子贮藏蛋白质的特异性表达是受精确调控的, 这些调控包括启动子中的顺式作用元件和反式作用因子。应用研究表明, 通过对种子贮藏蛋白质表达的合理调控可提高作物的产量和营养价值, 同时种子贮藏蛋白质具有稳定表达的特性, 可用于研究植物分类、植物系统进化和表达药用蛋白质。

关键词: 表达; 分子育种; 植物分类; 调控; 种子贮藏蛋白质

中图分类号: Q946.1

文献标识码: A

文章编号: 1000-470X(2008)02-0203-10

Progress in Research and Application on Expression and Regulation of Seed Storage Protein

YANG Xue-Jun¹, YU Fang-Yuan^{1*}, ZHANG Huan-Xi²

(1. College of Forest Resources and Environment, Nanjing Forestry University, Nanjing 210037, China;
2. Forestry Bureau of Jingxian, Jingxian, Anhui 242500, China)

Abstract: Maturing seed accumulates a large number of storage proteins and different plant species store different seed storage proteins. To understand the process of expression pattern of seed storage protein, the theoretical studies on the expression and regulation of seed storage protein and its application in the plant taxonomy, molecular breeding and expression of pharmaceutical protein in recent years were reviewed. The theoretical studies showed that the expression pattern of seed storage protein was accurately regulated, involving *cis*-elements and *trans*-acting factors. The applied studies showed that the yield and nutritional quality of crop could be greatly improved through the proper regulation of the expression of seed storage protein. Due to its stable expression pattern, seed storage protein could be used in the studies on plant taxonomy and plant systemic evolution, as well as expression of pharmaceutical proteins.

Key words: Expression; Molecular breeding; Plant taxonomy; Regulation; Seed storage protein

种子是植物发育形成的一种特异器官, 可为下一代植物生长提供营养物质, 并能保护遗传物质的稳定和促进种群的合理分布。种子作为一种贮藏器官, 成熟时大量合成蛋白质、淀粉和脂类等贮藏化合物。种子贮藏蛋白质(seed storage protein)一般是在种子发育后期大量合成和积累, 根据溶解特性可分为清蛋白(albumin)、球蛋白(globulin)、谷蛋白(glutenin)和醇溶蛋白(prolamin)。球蛋白是种子中含量最丰富的贮藏蛋白质, 而豆球蛋白(沉降系数11S)和豌豆球蛋白(沉降系数7S)是双子叶植物种子中最主要的球蛋白, 约占成熟种子总蛋白质含量的70%。谷蛋白是高分子量的聚合物, 分子间含有二硫键, 其亚基可根据分子量分为高分子量

(HMW)和低分子量(LMW)两种。醇溶蛋白只以分子内的二硫键结合但并不聚集, 醇溶蛋白可在低pH时用聚丙烯酰胺凝胶电泳进行分离, 可以分为4种, 即: α -、 β -、 γ -和 ω -醇溶蛋白, 其中 α -和 γ -醇溶蛋白中含硫氨基酸的含量很高。清蛋白一般可溶于水, 其中以2S清蛋白的研究最为广泛, 2S清蛋白在植物中常具抗真菌等特殊作用。在中性pH条件下, 绝大部分种子贮藏蛋白质不溶于水或稀释的缓冲液。除少数种子贮藏蛋白质如植物凝聚素、壳多糖酶和arcelin(菜豆中一种具有抗虫机制的蛋白质)具有生理活性外, 多数种子贮藏蛋白质不具有生理活性。

种子在发育中以蛋白质的形式贮藏氮源, 并在

收稿日期: 2007-05-27, 修回日期: 2008-03-04。

基金项目: 国家自然科学基金资助项目(30271076)。

作者简介: 杨学军(1977-), 男, 在读农学硕士研究生, 研究方向为林木种苗学(E-mail: xjyang_jx@126.com)。

* 通讯作者(E-mail: fyuu@njfu.edu.cn)。

种子萌发过程中逐渐降解,作为发芽后幼苗迅速生长发育所需碳、氮和硫等营养元素的来源。同时,豆类、谷物和牧草种子中含有丰富的贮藏蛋白质,是人类和动物的主要氮素营养来源。此外,一些种子贮藏蛋白质(如 2S 清蛋白)还具有药理方面作用。因此,研究种子贮藏蛋白质对深入了解种子发育的生物学规律、改善人类的营养结构和开发新的医药用途都很有意义。对种子贮藏蛋白质的研究一直倍受各国研究者的关注,笔者就种子贮藏蛋白质表达调控机制及其应用的研究进展情况作一介绍。

1 种子贮藏蛋白质的表达和调控

1.1 种子贮藏蛋白质的编码基因和位点

编码种子贮藏蛋白质的基因是一个多基因家族。Stevens 等^[1]用限制性核酸内切酶 *Bam* HI(26 880 个克隆)和 *Eco* RI(48 000 个克隆)构建两个藜谷(*Chenopodium quinoa* Willd.)BAC 文库,分离了编码 11S 球蛋白的同源基因并用 BAC 文库进行基因鉴定,Southern 杂交表明,BAC 文库中有 4 个克隆(为单一的 4 kb *Eco* RI 片断和单一的 12 kb *Hind* III 片断)存在 11S 球蛋白基因序列,且所获得的 11S 球蛋白基因可能是 BAC 文库中的克隆基因位点中含有的单拷贝,属于多基因家族的一部分。

研究者多以麦属植物为对象研究谷蛋白编码基因的位点。研究已确定,编码谷蛋白的基因位于第 1 组染色体上,其中 HMW 亚基的编码基因位于第 1 组染色体的长臂上,而 LMW 亚基的编码基因位于第 1 组染色体的短臂上。Payne 等^[2]首先报道,麦谷蛋白 HMW 亚基由第 1 组染色体长臂上的 1AL、1BL 和 1DL 上同源位点 *Glu-A1*、*Glu-B1* 和 *Glu-D1* 编码。同样,西班牙单粒小麦(*Triticum monococcum* L. ssp. *monococcum*)中的谷蛋白 HMW 亚基由第 1 组染色体长臂上部分同源的位点控制^[3],但编码 LMW 亚基的位点是染色体 1AS 上的 *Glu-A3*、1BS 上的 *Glu-B3*、1DS 上的 *Glu-D3*^[4]。小麦(*Triticum aestivum* L.)中 B 和 C 两个 LMW 麦谷蛋白亚基由紧连 *Glu-1* 多基因位点的 *Glu-3* 多编码基因位点相同,LMW 麦谷蛋白 D 亚基的编码基因与染色体 1D 上的 *Glu-D1* 位点紧密相连^[5]。Sreeramulu 和 Singh^[6]却发现,小麦中存在另外两种 LMW 麦谷蛋白亚基(相对分子量为 30 kD 和 32 kD),它们的编码基因分别在染色体 1D(*Glu-D4*)和 7D(*Glu-D5*)上,其中染色体 7D 与 α -醇溶蛋白有关。Nieto-Taladriz 等^[7]和 Joppa 等^[8]分别报道,硬质小麦(*Triticum turgidum*)谷蛋白

LMW-1 亚基包括 5 条带(5 + 8 + 9 + 13 + 16),第 5 条带由染色体 1A 短臂上的等位基因 *Glu-A3b* 编码,而 8 + 9 + 13 + 16 由染色体 1B 短臂上的等位基因 *Glu-B3b* 编码。Alvarez 等^[9]研究单粒小麦中麦谷蛋白的结果还表明,LMW 亚基由 *Gli-1* 位点紧密相连的 *Glu-3* 位点控制,与小麦的研究结果相同。

研究已证实,醇溶蛋白的编码基因位于第 1 组和第 6 组染色体的短臂上。Payne 等^[10]报道,小麦醇溶蛋白主要编码基因位于第 1 组和第 6 组染色体的短臂上,且 *Gli-1* 位点与谷蛋白 LMW 亚基的编码基因(*Glu-3*)位点紧密相连。Dubcovsky 等^[11,12]进一步研究指出,小麦染色体 1A、1B 和 1D 短臂上末端的 *Glu-1* 多基因位点编码 γ -、 ω -和一些 β -醇溶蛋白,而另一部分 ω -醇溶蛋白的编码基因位于 *Glu-3* 位点上,全部 α -醇溶蛋白和大部分 β -醇溶蛋白是由位于染色体 6A、6B 和 6D 短臂上的多基因位点编码。Payne^[3]研究指出,单粒小麦中控制醇溶蛋白的 *Gli-1* 位点分别位于第 1 组和第 6 组染色体短臂上。Xu 等^[13]也报道,硬质小麦醇溶蛋白由染色体 1B 上的 *Gli-B1* 位点编码,此位点与染色体 1B 短臂上谷蛋白 LMW 亚基编码基因 *Glu-B3* 紧密相连。

研究者还对球蛋白、清蛋白等种子贮藏蛋白质的编码基因进行了研究。 β -伴大豆球蛋白是一种异质性的三聚体大豆球蛋白,其亚基由序列上相邻的 2.5 kb 或 1.7 kb mRNAs 编码,但它们起源于 β -伴大豆球蛋白多基因家族的不同成员。研究已经发现了两种 β -伴大豆球蛋白的基因,即 CG4(1.7 kb mRNAs 亚族)和 CG1(2.5 kb mRNA 亚族)^[14],且这两种基因在表达上有差异,在转基因烟草中,CG4 只在胚中表达,但 CG1 在胚腔和胚乳中都可表达^[15]。Anisimova 等^[16]对一年生向日葵(*Helianthus annuus*)杂交分析表明,种子贮藏蛋白质 helianthinin 的合成至少受到 *HelA*、*HelB* 和 *HelC* 3 个位点的控制。Hou 等^[17]研究鼠耳芥属(*Arabidopsis*)的结果却表明,种子贮藏蛋白质由小基因家族编码,编码 5 个 2S 清蛋白的基因有 5 个(称为 at2S1 ~ at2S5),at2S1 ~ at2S4 在第 4 组染色体的下半区紧密相连,而 at2S5 则位于第 5 组染色体的下半区;同时发现 12S 球蛋白由 3 个基因编码(*CRA1*、*CRB* 和 *CRC*),分别位于第 5、第 1 和第 4 组染色体上。

值得注意的是,编码种子贮藏蛋白质的各基因位点不是独立的,而是可能存在相互作用。Dubcovsky 等^[12]用 DNA 克隆限制片断杂交和 SDS-PAGE 研究二倍体、四倍体和六倍体小麦中 HMW

及 LMW 麦谷蛋白位点、醇溶蛋白位点和小麦糖位点之间的关系,结果表明,在四倍体和六倍体小麦中,含有 LMW 麦谷蛋白克隆的 DNA 片断,位于染色体臂 1AS、1BS 和 1DS 图谱末端区域的 *XGlu-3* 位点,第 2 个位点(*XGlu-B2*)位于染色体臂 1BS 上,且与醇溶蛋白位点 *XGli-B3* 完全相连,在四倍体小麦的 SDS-PAGE 图谱上这个位点的限制性片断表现出与 LMW 麦谷蛋白的 B 亚基共分离的特点,表明 *XGlu-B2* 是一个 LMW 麦谷蛋白的活性位点。该项研究还观察到,在二倍体小麦的 7A^M 染色体长臂上有一个与低分子量克隆片断杂交的新位点,没有麦谷蛋白可与其共分离;小麦糖位点位于染色体臂 1AS、1BS 和 1DS 上,且没有发现与 DNA 片断共分离的小麦糖蛋白,表明 *XTri-B1* 位点是没有活性的;没有发现 *Gli-A4* 的存在,可能与 *Gli-A3* 同义。

同时,基因的位点上存在许多的等位基因,它们对种子贮藏蛋白质合成数量和质量都有作用,如在小麦中,谷蛋白 HMW 亚基就决定面粉的质量。Payne 等^[18]指出,“5 + 10”亚基是由 *Glu-D1d* 编码,可以提高面粉的质量。Xu 等^[13]用荧光基因组原位杂交(FISH)和微卫星标记研究 2 个硬质小麦 [*Triticum turgidum* L. ssp. *durum* (Desf.) Husn.] 家系[由普通小麦和硬质小麦 2 个品种(LDN 和 Reville)进行杂交得到]的结果表明,它们都含有由 *Glu-D1d* (L252 和 S99B34) 编码的谷蛋白 HMW 亚基 1Dx5 + 1Dy10,但是产量低的家系中含有 LMW-1 亚基(由 *Glu-A3k* 和 *Glu-B3s* 基因编码)和 1B(由硬质小麦 LDN 的醇溶蛋白基因编码),而产量高的 3 个家系含有 LMW-2 亚基和 1B(由硬质小麦 Reville 的醇溶蛋白基因编码),表明硬质小麦品种 LDN 中含有影响产量的基因,并且由于硬质小麦在 D- 染色体组上缺少谷蛋白和醇溶蛋白的编码基因,所合成的蛋白质质量不好。

总之,研究种子贮藏蛋白质的编码基因是了解种子贮藏蛋白质的首要和重要的一步。编码种子贮藏蛋白质的基因属于多基因家族,其序列和位点千差万别。不同种类的种子贮藏蛋白质有不同的编码基因,不仅贮藏蛋白编码各基因之间存在相互作用,而且蛋白编码基因与其它物质的编码基因之间也存在着相互作用,这从根本上决定着种子贮藏蛋白质表达模式的多样性。

1.2 种子贮藏蛋白质的表达与调控

种子贮藏蛋白质的基因只在种子发育的特定阶段表达,并指导种子贮藏蛋白质的合成和积累。

Boulter 等^[19]报道,豌豆 (*Pisum sativum*) 中贮藏蛋白质的基因仅在种子发育过程中被激活转录,可为研究植物组织特异表达和发育调控提供很好的实验系统。Thompson 等^[20]也发现,豌豆中 *legA* 基因编码一种豆球蛋白亚基的前体蛋白,在豌豆种子子叶细胞增大期,该基因的转录水平升高, *legA* mRNA 丰度增加。

研究已确定,7S 和 11S 球蛋白启动子存在保守序列,并分别命名为豌豆球蛋白盒(vicilin box)和豆球蛋白盒(legumin box)。豆球蛋白盒序列 CATGCATG 存在于许多豆类种子贮藏蛋白基因的启动子中,被命名为 RY 重复序列^[21]。对转基因植株的研究表明,RY 重复序列对 7S 球蛋白和 β -伴大豆球蛋白在种子中特异表达是必须的。Yamauchi^[22]报道,一些核蛋白可同时与刀豆球蛋白基因 5'-上游区域中 CACA 及富含 A/T 的序列产生相互作用。Bobb 等^[23]发现, *PvAlf* (*Phaseolus vulgaris* *abscisic acid-insensitive 3-like factor*) 与拟南芥 (*Arabidopsis thaliana*) 中的蛋白编码基因 *ABI3* (*abscisic acid-insensitive 3*) 相似,它可激活扁豆种子中一些贮藏蛋白质、菜豆蛋白和植物凝聚素基因的转录; *PvAlf* 还可调节启动子中的 RY 重复序列,但启动子与亮氨酸拉链型转录因子 ROM1 和 ROM2 相互作用则可抑制 *PvAlf* 对转录的激活作用。

种子贮藏蛋白质基因表达的初始步骤是转录的起始,许多研究表明,贮藏蛋白质编码基因的起始序列与相邻序列存在相互作用,且一些蛋白质可起始和增强转录。Shirsat 等^[24]研究缺失突变(deletion mutant)的结果表明,指导转基因烟草植株中豆球蛋白正确表达的起始序列位于帽子结构的 -549 ~ -97 之间,其表达水平明显受到相邻序列的调控。Roeder^[25]认为,蛋白质可与重复序列(-404 ~ -367) 和 USR1-II (-482 ~ -405) 内的成分相结合,之后再与 RSR2 中另一个结合位点发生相互作用,进而装配成转录起始复合物。Meakin 和 Gatehouse^[26]用酶裂解豌豆中 *legA* 基因的 873 bp 起始片断(-833 ~ +40),然后分析产物与豌豆种子核蛋白提取物的结合能力,结果也表明, -316 ~ +40 之间的起始序列与豌豆种子核提取物不形成稳定复合物,而 -549 ~ -316 和 -833 ~ -582 之间的 2 个序列则可以与种子核提取物产生强烈作用并形成结合物(LABF1),但与叶核提取物却没有这种作用;用 SDS-PAGE 分离种子核蛋白,再进行洗提和复性后发现, LABF1 是分子量分别为 84 ~ 116 kD 和 48 ~

58 kD 的多肽,它尽管仅临时作用于豆球蛋白基因表达,但这种作用既可以作为增强子又可以作为起始因子。Howley 和 Gatehouse^[27]用迁移率试验 (electromobility shift assay) 和竞争性试验 (competition assay) 研究豌豆细胞扩展时期出现的 DNA 结合蛋白质的目标序列时发现,这种蛋白质结合于 *legA* 启动子的转录起始位点 -404 ~ -367 的不完全序列上,被称为 USR1 (-549 ~ -316),另有一种相同或相似的蛋白质结合于 USR2 (-833 ~ -582) 上,同时还有一种没有重复序列的蛋白质与目标位点结合于 USR1 相近的上游区域;该项研究同时指出,这些 DNA 结合蛋白质之间的相互作用是装配转录起始复合物的事件,由于这些结合作用只存在于豌豆子叶细胞增大期的种子蛋白提取物中,且与豆球蛋白 A 增长相平行,因此这些蛋白质可能是起始转录或增强转录的因子。

种子特异性基因的表达是受精确调控的,这些调控包括启动子中的顺式作用元件和反式作用因子。研究已确定,在大麦、小麦和水稻的胚乳特异启动子有 3 种保守性顺式基序,分别命名为 GLM、醇溶谷蛋白盒 (PB) 和 5'-AACAT-3'。其中,在大麦和水稻中,5'-AACAT-3' 基序是 GAMYB 和 OsMYB5 的结合位点^[28]。

GLM (GCN4 类基序, 5'-ATGAG/CTCAG-3') 与 bZIP (属于 *Opaque2* 亚族) 蛋白相结合, 因而它与 O2 蛋白的作用有关系。O2、OHP1、OsBZIPPA、SPA、CPRF2 和 RITA1 这 5 种蛋白都属于碱性亮氨酸拉链家族。O2 基因最初是从玉米 (*Zea mays*) 中分离的, 是 α -和 β -醇溶谷蛋白基因的转录起始因子, 它可导致胚乳中玉米蛋白 (主要是 22 kD 的 α -玉米蛋白) 总量明显下降。Schmidt 等^[29]指出, 玉米中隐性突变基因 O2 编码亮氨酸拉链 DNA 结合蛋白, 可调节玉米蛋白的合成, 但由于这种蛋白的赖氨酸含量很低, 基因突变后玉米蛋白的合成减少却引起种子蛋白中赖氨酸的含量上升。Pater 等^[30]对 O2 基因进行克隆后发现, 它编码一种 DNA 结合蛋白, 属于碱性亮氨酸拉链区域内的起始因子。Schmidt 等^[31]报道, O2 蛋白通过识别启动子中特殊序列而控制玉米中 22 kD α -玉米蛋白和 *b-32* 基因的表达。Cord-Neto 等^[32]进一步研究后指出, O2 蛋白除控制 22 kD α -玉米蛋白基因的表达外, 还控制玉米和一些薏苡属 (*Coix*) 种子中 β -醇溶谷蛋白基因的表达; 此外, O2 蛋白还调控玉米种子成熟过程中蛋白质的合成、氮和糖的代谢^[33]。Pirovano 等^[34]报道, O2 蛋

白的同源基因还存在于高粱属 (*Sorghum*) 植物中; Vettore 等^[35]进一步研究后发现, 薏苡属、高粱属和玉米种子中的贮藏蛋白质都是由同源基因编码的, 它们在胚乳中可暂时性表达, 并且发现薏苡属 O2 基因中有 5 个内含子, 编码一种含有 408 个氨基酸的多肽, 与玉米中的 O2 基因相同。

醇溶谷蛋白盒 (PB, 5'-TGTAAAG-3') 由 DOF (DNA binding with one finger) 转录因子识别。缺失和点突变试验表明, PB 在调节胚乳特异性基因表达中很重要^[36, 37]。小麦中的 PBF (WPBF) 首先是从小麦胚乳中发现的, Vicente-Carbajosa 等^[38]在进行序列分析后认为, 小麦 α -醇溶蛋白基因的启动子区域中存在一个 PB-盒基序, 可与 *Hor2* (大麦醇溶谷蛋白基因) 启动子的 PB-盒结合, 可能在醇溶谷蛋白的表达中充当启动子^[36]。Dong 等^[39]报道, 小麦的所有组织中都有 WPBF 基因表达, 并发现小麦根部的 cDNA 中存在一种 TaQM 蛋白, 可与 WPBF 中 DOF 区发生作用, 且 TaQM 蛋白基因与 WPBF 基因的表达相似; 该项研究还在共侵染的 BY-2 原型细胞进行瞬时表达实验, 结果表明, WPBF 与 PB-盒结合后, 可以激活 α -醇溶蛋白启动子的转录, 特别是在 WPBF 和 *TaQM* 基因共侵染时, α -醇溶蛋白启动子的转录活性比 WPBF 单独作用时高得多。另外, WPBF 基因还存在于转基因鼠耳芥属植物的种子和维管系统中, 所以它不仅在小麦种子发育中起作用, 而且在其它植物生长发育过程中也起作用。Diaz 等^[28]研究大麦 (*Hordeum vulgare* L.) 种子发育的结果表明, 大麦中 PBF (BPBF) 的 C-末端可与 HvGAMYB 相互作用, 从而激活胚乳特异性基因的表达, 并认为 BPBF 与 WPBF 具有同源性。Diaz 等^[37]进一步用双分子荧光络合 (BiFC) 方法证实, 大麦中的 DOF 蛋白 (SAD) 也可与核中 HvGAMYB 相互作用。Yanagisawa^[40]也报道, 在玉米种子发育过程中, 玉米蛋白的表达包括 PBF 和 O2 蛋白之间的相互作用, 且它们在谷类种子许多贮藏蛋白质基因的启动子中都是保守的。研究已证实, DOF 蛋白不仅可与顺式作用元件相结合, 而且它的 DOF 区或 C-末端可调节蛋白质之间的相互作用, 因此在种子发育过程中包括 bZIP 和 DOF 家族在内的反式作用因子共同调节贮藏蛋白质的表达^[38]。

可见, 种子贮藏蛋白质的基因仅在种子发育中的特定时期才被激活转录, 并且基因的时空表达模式受多种因素的综合调控, 包括识别特异性序列的 DNA 结合蛋白和相应的顺式作用元件与蛋白质之

间的相互作用^[41];因此,种子贮藏蛋白质的表达同样也受各种顺式作用元件和反式作用因子的调控,包括各种特异性表达的启动子、增强子、DNA结合蛋白质以及它们之间的相互作用。种子贮藏蛋白质的表达调控机制纷繁复杂,是目前种子分子生物学研究的一个重点领域。

2 种子贮藏蛋白质表达调控机制的应用

2.1 种子贮藏蛋白质可用于物种鉴定与植物分类

Singh 等^[42]认为,同种植物的不同表现型中蛋白质组成的变化非常有限或是没有变化,这意味着,在特定情况下,蛋白质组成的变化可以对同种植物的不同表现型进行分类。因而,蛋白质的 SDS-PAGE 分析可对表型特殊的类型进行鉴定和分类,特别是鉴定不同表现型之间的杂交种,对蛋白质的化学分析(SDS-PAGE 等)和同功酶分析有可能提供比 RAPD 分析更准确的分类。

大豆(*Glycine max* L. Merr.)种子中的贮藏蛋白很容易用 SDS-PAGE 进行分离和分析,可作为基因型的标记,在植物分类和物种鉴定上应用很广泛。Mennella 等^[43]对意大利坎帕尼亚地区 9 种生态型的普通大豆用盐溶液进行提取,然后用离子交换高效液相色谱(AE 和 CE-HPLC)、SDS-PAGE、等电聚焦(IEF)进行分析,结果表明,分析种子贮藏蛋白质可从质和量上对不同生态型进行划分。Thorne 等^[44]报道,菜豆(*Phaseolus vulgaris* L.)不同野生基因型中的菜豆蛋白也有很高的多态性,认为分析菜豆蛋白可以对不同地理起源进行分类。Mennella 等^[45,46]用高效液相色谱分别分析园艺植物茄属(*Solanum*)及番茄(*Lycopersicon esculentum* L. cv. Cherry)等种子胚乳中的种子贮藏蛋白,并对不同种和栽培变种进行了分类。Vladova 等^[47]对 34 个辣椒(*Capsicum annum*)栽培种的种子贮藏蛋白质进行电泳分析发现,2 种表型基因的差异是由 11S 球蛋白的中间亚基或它们的低聚物引起的,并根据蛋白质电泳图谱的明显不同将这些栽培种分为 2 类。Syros 等^[48]根据克里特黑檀树(*Ebenus cretica*)总蛋白质的电泳图谱,将其划分为 4 种生态型。Bianchi-Hall 等^[49]和 Singh 等^[50]也分别指出,研究落花生属(*Arachis*)植物中的种子贮藏蛋白质可对落花生属的野生种质、落花生(*Arachis hypogaea* L.)栽培种的特点和同种植物内可能存在的不同表现型进行鉴定。Bertozzi 和 Valls^[51]用 SDS-PAGE 研究平托落花生(*Arachis pintoi*)和 *A. repens* 种子贮藏蛋白质时发现,

它们的种子贮藏蛋白质存在很大差异,且不同表现型之间种子贮藏蛋白质的差异比表现型内的差异要大。Alvarez 等^[52]对圆柱小麦(*Triticum dicoccum*)贮藏蛋白质分析也表明,谷蛋白 HMW 亚基差异很大,新的等位基因在植物材料中出现的频率很低,表明有多种基因型。Alvarez 等^[9]用 SDS-PAGE 和 A-PAGE 研究西班牙单粒小麦种子蛋白质组成的结果表明,谷蛋白的 *Glu-A1^m* 有 3 个不同的等位基因,而 *Glu-A3^m* 有 6 个不同的等位基因;醇溶蛋白的 *Gli-A1^m* 有 7 个等位基因, *Gli-A2^m* 则有 14 个等位基因,并根据起源和种子贮藏蛋白组成的不同划分为 48 种不同基因型。

此外,通过分析子代所表达的种子贮藏蛋白质还可预测杂种优势,如分析同功酶和同种异型酶可预测粮食作物的杂种优势。Han 等^[53]使用基因组原位杂交(GISH)和多色 GISH(mcGISH)技术进行细胞发生,对小麦 × 中间偃麦草(*Thinopyrum intermedium*)进行杂交获得了 5 个部分二倍体,并分析麦谷蛋白 HMW 亚基和醇溶蛋白,结果表明,其中 2 个二倍体有与小麦相同的带型,有 1 个二倍体有中间偃麦草的特异性带型并有良好的 HMW 亚基,适合小麦育种。Yu 等^[54]还报道,基于形态学特征、同功酶和种子贮藏蛋白质估算的遗传距离(GDs)和随机扩增多态性 DNAs(RAPD)可用于预测油菜(*Brassica napus* L.)各杂交组合的性状和杂种优势。

2.2 种子贮藏蛋白质可用于研究系统演化

研究种子贮藏蛋白质氨基酸序列和编码基因的相互关系,还有可能阐明不同物种的系统演化过程和系统发生关系。Shutov 等^[55]对豆球蛋白和豌豆球蛋白基因的内含子位置进行对比后发现,编码现有的豆球蛋白和豌豆球蛋白亚基的基因可能是由原始的单一区域进化而来的,为用种子贮藏蛋白质分析来研究系统演化提供了理论依据。为了研究种子贮藏蛋白质的演化过程,Shutov 等^[56]还对裸子植物的一些属,如买麻藤属(*Gnetum*)、麻黄属(*Ephedra*)和百岁兰属(*Welwitschia*)(它们形态不同却在大多数分类系统中归为买麻藤目)代表种的成熟种子蛋白质进行部分测序,通过克隆得到基于 N-末端氨基酸序列的 cDNA,并分析千岁兰(*Welwitschia mirabilis*)编码豆球蛋白核基因的结构,结果表明,买麻藤目植物第 1 个核基因结构在 4 个保守位置上仅有 2 个具有内含子,对外显子和内含子的序列进行分析表明,买麻藤目植物的豆球蛋白基因与现存的银杏、松柏目及所有被子植物均不同,可能属于单起源的

进化分枝。Bianchi-Hall^[57]等也指出,研究落花生属植物中的种子贮藏蛋白质也可阐明其系统发生关系。

2.3 种子贮藏蛋白质可用于分子育种

随着分子生物学技术的发展,分子育种(molecular breeding)在育种工作中已日趋重要。利用种子贮藏蛋白质开展分子育种有两种方法:一种是蛋白质工程方法,即调节种子蛋白质的编码基因;另一种是转基因方法,即从外源物种转入编码基因。Teraishi等^[58]发现一种引起所有 β -伴大豆球蛋白亚基基因不能表达的优势突变,证明了在自然种群中种子定性和定量突变可以遗传。种子贮藏蛋白质基因在种子发育中高度表达,许多研究表明,这种基因在转基因植株也能正确表达。调节种子特异表达序列位于贮藏蛋白质基因的启动区,对种子特异表达启动子的正向调控区域(RY重复序列)进行复制建立合适的启动子,可提高种子蛋白质的营养价值。

利用种子贮藏蛋白质开展分子育种已成功应用于粮食作物中,并取得了良好的效果。例如在大豆中,很早就发现种子贮藏蛋白质具有良好性状,并可用来培育新品种提高种子蛋白质的产量,但豆类种子中含硫氨基酸的比例很低,因而营养价值比动物蛋白质低。为提高菜豆蛋白的营养品质,Hoffman等^[59]将包含有6个甲硫氨酸密码子的寡核苷酸(45 bp)插入到菜豆蛋白基因的第3个外显子中以获得目标蛋白质,结果却发现,转基因烟草种子中仅产生少量的目标蛋白质;Pueyo等^[60]进一步用免疫电子显微技术进行观察发现,种子发育早期在内质网、高尔基体、运输小泡及液泡中都能观察到目标蛋白质的积累,但是在种子成熟后期在贮藏液泡中却没有检测到目标蛋白质,因此他们认为加入甲硫氨酸残基的种子蛋白质因发生错误折叠,在液泡中不稳定。Kim等^[61]报道,11S球蛋白的氨基酸序列中有非保守区域(称为可变区),可以进行调节;因此,Utsumi等^[62]将含有4个甲硫氨酸密码子的合成DNA插入大豆球蛋白cDNA的可变区,结果发现,目标蛋白质在烟草植株中可以表达,并可在烟草种子的液泡中正常装配和积累。Kjemtrup等^[63]对植物凝聚素的基因进行定点突变,在转基因烟草种子的液泡中也获得了目标植物凝聚素。

用转基因方法提高种子蛋白的质量,首先必须融合种子特异性的启动子和含硫氨基酸的蛋白质基因组成嵌合基因,然后将嵌合基因转入栽培作物中,进而提高种子贮藏蛋白质的营养价值。为了提高烟

草种子蛋白质中甲硫氨酸含量,Altenbach等^[64]融合菜豆蛋白的启动子与2S清蛋白的cDNA,并将嵌合基因转入烟草植株中,结果使转基因烟草种子中2S清蛋白的甲硫氨酸含量提高了30%。Nordlee等^[65]报道,巴西坚果(*Bertholletia excelsa*)中的2S清蛋白在转基因大豆种子中也可以合成。Karchi等^[66]报道,将融合种子特异性的启动子和对赖氨酸及苏氨酸的反馈抑制不敏感的大肠杆菌(*Escherichia coli*)天冬氨酸激酶基因组成嵌合基因,并将其转入烟草中,结果发现,种子中自由苏氨酸和甲硫氨酸的含量上升,但种子蛋白质中这两种氨基酸含量却没有明显上升。Sharma等^[67]对 δ -玉米蛋白(δ -zein,一种玉米中的富硫种子贮藏蛋白质)的编码基因进行修饰后转入白三叶草(*Trifolium repens*)中,并用花椰菜病毒(*cauliflower mosaic virus, CaMV*)的二倍35S起始子和胭脂仙人掌属植物的基因转录终止子控制该种基因的表达,结果表明,转基因植株可以表达特异的mRNA,并在叶片、叶柄、节、节间、根和种子中积累 δ -玉米蛋白,且修饰后基因所表达的 δ -玉米蛋白N-末端的序列与玉米种子中的相同,因此他们认为,可用这种贮藏蛋白质改善白三叶草的草料质量,并可防治某些疾病。

2.4 种子贮藏蛋白质可用于医药工程

由于种子具有较长的贮藏时间,用种子作为生物反应器生产具有药用价值的蛋白质和多肽很有潜力。用种子贮藏蛋白质表达药用蛋白质,首先必须找到一种合适的启动子。花椰菜病毒35S起始子是植株中表达外来蛋白质最常用的启动子,但其活性在植物不同组织中有差异,在叶片中活性最高。Fiedler等^[68]和Hoffman等^[69]分别用双子叶植物的豆球蛋白和 β -菜豆蛋白基因的起始子在烟草种子中表达了工程抗体和玉米醇溶蛋白;Firek等^[70]还在转基因烟草种子中成功表达了免疫球蛋白。Hood等^[71]和Witcher等^[72]分别用玉米中泛基因(ubiquitin gene)的启动子在玉米胚中也成功表达了高含量的抗生素蛋白和 β -葡萄糖苷酸酶(β -glucuronidase)。另一种表达重组蛋白质和多肽的策略是以自然种子蛋白质为载体,Vandekerckhove等^[73]和Parmenter等^[74]分别融合种子贮藏蛋白napin及亲脂蛋白oleosin,获得了Leu-脑啡肽和水蛭素。Leite等^[75]用 γ -高粱醇溶蛋白基因,在转基因烟草种子中获得了人类生长激素,并观察到,在种子成熟过程中这种激素由信号肽作用定向运输到内质网,其含量占可溶性蛋白质总量的0.16%,且该种

激素与人体激素有相同的氨基酸序列和与受体结合的性质。

大豆球蛋白 *GI* 基因编码一种含量很高的种子贮藏蛋白质,其启动子可启动药用蛋白质在种子中的特异表达,因而具有良好的应用前景。Ding 等^[76]用大豆球蛋白 *GI* 基因的启动子启动人类纤维原生长因子(bFGF)在转基因大豆中的特异表达,即在 18 kD bFGF 编码区嵌入 *GI* 基因的启动子或启动子的信号肽序列,再将其转入大豆中,然后对转基因植株的分析表明,在种子中 bFGF 可起始转录,表达的蛋白质占可溶性蛋白质总量的 2.3%;大豆合成的 bFGF 可促进 Balb/c 3T3 细胞的有丝分裂,并和人体合成的 bFGF 具有相同的性质。该项研究还发现, *GI* 基因启动子可促进一种 β -葡糖昔酸酶在种子中的特异表达,因此他们认为 *GI* 基因启动子含有种子特异性表达所需的顺式作用元件,可用于药用蛋白质在大豆种子中的表达。

综合这方面的研究可以看出,种子贮藏蛋白质表达调控的应用研究也是一个十分活跃的领域,且很有意义。首先,研究种子贮藏蛋白质可解决物种鉴定、植物分类和系统演化的问题,并能准确评估遗传变化;蛋白质分析与其它生物化学及分子标记技术相结合,可为种间和种内的分类提供一种新的工具,可更好地估计出遗传多样性;电泳图谱分析已成为育种上用于栽培种鉴定的有利工具。其次,研究种子贮藏蛋白质可从蛋白质变化的角度揭示生物进化的规律。再次,研究种子贮藏蛋白质可明晰蛋白质表达的调控机制,通过调控表达过程,提高作物的产量和营养价值。最后,研究种子贮藏蛋白质可获取人类需要的药用蛋白质和其它生物活性物质,促进医药卫生事业的发展。

3 结语

综上所述,研究种子贮藏蛋白质可深入了解种子发育的生物学规律,进而可对表达进行有效调控,对改善人类生活很有意义。目前,种子贮藏蛋白质研究颇受各国研究者瞩目,并取得了重大的进展。对种子贮藏蛋白质表达调控机制的研究表明,其特异性表达是受精确定控的,这些调控包括启动子中的顺式作用元件和反式作用因子。此外,通过对种子贮藏蛋白质表达的合理调控可提高作物的产量和营养价值,由于种子贮藏蛋白质具有稳定表达的特性,可用于研究植物分类、植物系统进化和表达药用蛋白质。

尽管近年来种子贮藏蛋白质的研究取得了一定进展,但由于种子贮藏蛋白质的研究内容复杂,还有很多问题有待深入研究。具体来说,以下几个方面的研究可能很重要:(1)探索研究种子贮藏蛋白质的新方法;(2)扩大种子贮藏蛋白质的物种研究范围;(3)进一步对种子贮藏蛋白质表达和调控机制进行深入的研究,特别是蛋白质翻译后修饰的机制和非生物因子对蛋白质表达的作用机制;(4)研究蛋白质之间相互作用及其生物学功能;(5)寻找适合分子育种的启动子和能在转基因植株中稳定表达的新基因;(6)基因安全问题应加以高度重视,对转基因植株所合成蛋白质的安全性进行深入研究。

参考文献:

- [1] Stevens M R, Coleman C E, Parkinson S E, Maughan P J, Zhang H B, Balzotti M R, Kooyman D L, Arumuganathan K, Bonifacio A, Fairbanks D J, Jellen E N, Stevens J J. Construction of a quinoa (*Chenopodium quinoa* Willd.) BAC library and its use in identifying genes encoding seed storage proteins [J]. *Theor Appl Genet*, 2006, 112: 1593-1600.
- [2] Payne P I, Holt L M, Worland A J, Law C N. Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin. Part 3. Telocentric mapping of the subunit genes on the long arms of the homoeologous group 1 chromosomes [J]. *Theor Appl Genet*, 1982, 63: 129-138.
- [3] Payne P I. Genetics of wheat storage proteins and the effects of allelic variation on bread-making quality [J]. *Ann Rev Plant Physiol*, 1987, 38: 141-153.
- [4] Pogna P E, Autran J C, Mellini F, Lafiandra D, Feillet P. Chromosome 1B-encoded gliadins and glutenin subunits in durum wheat: genetics and relationship to glutenin strength [J]. *J Cereal Sci*, 1990, 11: 15-34.
- [5] Pogna N E, Redaelli R, Vaccino P, Biancardi A M, Peruffo A D B, Curioni A, Metakovsky E V, Pagliaracci S. Production and genetic characterization of near-isogenic lines in the bread-wheat cultivar Alpe [J]. *Theor Appl Genet*, 1995, 90: 650-658.
- [6] Sreeramulu G, Singh N K. Genetic and biochemical characterization of novel low-molecular-weight glutenin subunits in wheat (*Triticum aestivum* L.) [J]. *Genome*, 1997, 40: 41-48.
- [7] Nieto-Taladriz M T, Ruiz M, Martínez M C, Vázquez J F, Carrillo J M. Variation and classification of B low-molecular-weight glutenin subunit alleles in durum wheat [J]. *Theor Appl Genet*, 1997, 95: 1155-1160.
- [8] Joppa L R, Klindworth D L, Hareland G A. Transfer of high molecular weight glutenins from spring wheat to durum wheat [A]. In: Slinkard A E ed. *Proceedings of the 9th International Wheat Genetic Symposium* [M]. Saskatoon, SK, Canada: University of Saskatchewan Ext. Press, 1998, 257-260.
- [9] Alvarez J B, Moral A, Martín L M. Polymorphism and genetic diversity for the seed storage proteins in Spanish cultivated einkorn

- wheat (*Triticum monococcum* L. ssp. *monococcum*) [J]. *Genet Resour Crop Evol*, 2006, 53: 1061–1067.
- [10] Payne P I, Jackson E A, Holt L M, Law C N. Genetic linkage between endosperm storage protein genes on each of the short arms of chromosome 1A and 1B in wheat [J]. *Theor Appl Genet*, 1984, 67: 235–243.
- [11] Dubcovsky J, Luo M-C, Dvorak J. Differentiation between homoeologous chromosomes 1A of wheat and 1A^m of *Triticum monococcum* and its recognition by the wheat *Ph1* locus [J]. *Proc Natl Acad Sci USA*, 1995, 92: 6645–6649.
- [12] Dubcovsky J, Echaide M, Giancola S, Rousset M, Luo M C, Joppa L R, Dvorak J. Seed-storage-protein loci in RFLP maps of diploid, tetraploid, and hexaploid wheat [J]. *Theor Appl Genet*, 1997, 95: 1169–1180.
- [13] Xu S S, Faris J D, Cai X, Klindworth D L. Molecular cytogenetic characterization and seed storage protein analysis of 1A/1D translocation lines of durum wheat [J]. *Chromosome Res*, 2005, 13: 559–568.
- [14] Harada J J, Barker S J, Goldberg R B. Soybean β -conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and post-transcriptional processes [J]. *Plant Cell*, 1989, 1: 415–425.
- [15] Barker S J, Harada J J, Goldberg R B. Cellular localization of soybean storage protein mRNA in transformed tobacco seeds [J]. *Proc Natl Acad Sci USA*, 1988, 85: 458–462.
- [16] Anisimova I N, Gavrilova V A, Loskutov A V, Rozhkov V T, Tolmachev V V. Polymorphism and inheritance of seed storage protein in sunflower [J]. *Russ J Genet*, 2004, 40: 995–1002.
- [17] Hou A, Liu K, Catawatcharakul N, Tang X, Nguyen V, Keller W A, Tsang E W T, Cui Y. Two naturally occurring deletion mutants of 12S seed storage proteins in *Arabidopsis thaliana* [J]. *Planta*, 2005, 222: 512–520.
- [18] Payne P I, Corfield K G, Holt L M, Blackman J A. Correlations between the inheritance of certain high-molecular weight subunits of glutenin and bread-making quality in progenies of six crosses of bread wheat [J]. *J Sci Food Agric*, 1981, 32: 51–60.
- [19] Boulter D, Evans I M, Ellis J R, Shirsat A, Gatehouse J A, Croy R R D. Differential gene expression in the development of *Pisum sativum* [J]. *Plant Physiol Biochem*, 1987, 25: 283–289.
- [20] Thompson A J, Evans I M, Boulter D, Croy R R D, Gatehouse J A. Transcriptional and post-transcriptional regulation of seed storage-protein gene expression in pea (*Pisum sativum*) [J]. *Planta*, 1989, 179: 279–287.
- [21] Dickinson C D, Evans R P, Nielsen N C. RY repeats are conserved in the 5'-flanking regions of legume seed protein genes [J]. *Nucleic Acids Res*, 1988, 16: 371.
- [22] Yamauchi D. Interaction of seed embryo factors with the 5'-upstream region of seed storage 7S globulin gene from *Canavalia gladiata* [J]. *Plant Sci*, 1997, 126: 163–172.
- [23] Bobb A J, Eiben H G, Bustos M M. PvAlf, an embryo-specific acidic transcriptional activator enhances gene expression from phaseolin and phytohemagglutinin promoters [J]. *Plant J*, 1995, 8: 331–343.
- [24] Shirsat A, Wilford N, Croy R, Boulter D. Sequences responsible for the tissue-specific promoter activity of a pea legumin gene in tobacco [J]. *Mol Gen Genet*, 1989, 215: 326–331.
- [25] Roeder R G. The complexity of eukaryotic transcription initiation: regulation of preinitiation complex assembly [J]. *Trends Biochem Sci*, 1991, 16: 402–408.
- [26] Meakin P, Gatehouse J A. Interaction of seed nuclear proteins with transcriptionally enhancing regions of the pea (*Pisum sativum*) legA gene promoter [J]. *Planta*, 1991, 183: 471–477.
- [27] Howley P M, Gatehouse J A. A 38 bp repeat sequence within the pea seed storage protein promoter of legA is a binding site for a nuclear DNA-binding protein [J]. *Plant Mol Biol*, 1997, 33: 175–180.
- [28] Diaz I, Vicente-Carabajosa J, Abraham Z, Martinez M, Isabel-La Moneda I, Carbonero P. The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development [J]. *Plant J*, 2002, 29: 453–464.
- [29] Schmidt R J, Burr F A, Aukerman M J, Burr B. Maize regulatory gene opaque-2 encodes a protein with a “leucine-zipper” motif that binds to zein DNA [J]. *Proc Natl Acad Sci USA*, 1990, 87: 46–50.
- [30] Pater S, Katgiri F, Kijne J, Chua N H. bZIP proteins bind to a palindromic sequence without an ACGT core located in a seed-specific element of the pea lectin promoter [J]. *Plant J*, 1994, 6: 133–140.
- [31] Schmidt R J, Ketudat M, Aukerman M J, Hoschek G. Opaque-2 is a transcriptional activator that recognizes a specific target site in 22-kD zein genes [J]. *Plant Cell*, 1992, 4: 689–700.
- [32] Cord-Neto G, Yunes J A, da Silva M J, Vettore A L, Arruda P, Leite A. The involvement of Opaque 2 in β -prolamin gene regulation in maize and Coix suggests a more general role for this transcriptional activator [J]. *Plant Mol Biol*, 1995, 27: 1015–1029.
- [33] Gallusci P, Varott S, Matsuoka M, Maddaloni M, Thompson R D. Regulation of cytosolic pyruvate, orthophosphate dikinase expression in developing maize endosperm [J]. *Plant Mol Biol*, 1996, 31: 45–55.
- [34] Pirovano L, Lanzini S, Hartings H, Lazzaroni N, Rossi V, Joshi R, Thompson R D, Salamini F, Motto M. Structural and functional analysis of an opaque-2-related gene from sorghum [J]. *Plant Mol Biol*, 1994, 24: 515–523.
- [35] Vettore A L, Yunes J A, Neto G C, Da Silva M J, Arruda P, Leite A. The molecular and functional characterization of an Opaque2 homologue gene from Coix and a new classification of plant bZIP proteins [J]. *Plant Mol Biol*, 1998, 36: 249–263.
- [36] Mena M, Vicente-Carabajosa J, Schmidt R J, Carbonero P. An endosperm-specific DOF protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin-box of a native B-hordein promoter in barley endosperm [J]. *Plant J*, 1998, 16: 53–62.
- [37] Diaz I, Martinez M, Isabel-LaMoneda I, Rubio-Somoza I, Carbonero P. The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development [J]. *Plant J*, 2002, 29: 453–464.

- ro P. The DOF protein, SAD, interacts with GAMYB in plant nuclei and activates transcription of endosperm-specific genes during barley seed development [J]. *Plant J*, 2005, 42: 652–662.
- [38] Vicente-Carbajosa J, Moose S P, Parsons R L, Schmidt R J. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2 [J]. *Proc Natl Acad Sci USA*, 1997, 94: 7685–7690.
- [39] Dong G Q, Ni Z F, Yao Y Y, Nie X L, Sun Q X. Wheat Dof transcription factor WPBF interacts with TaQM and activates transcription of an alpha-gliadin gene during wheat seed development [J]. *Plant Mol Biol*, 2007, 63: 73–84.
- [40] Yanagisawa S. The Dof family of plant transcription factors [J]. *Trends Plant Sci*, 2002, 7: 555–560.
- [41] Levine M, Tjian R. Transcription regulation and animal diversity [1]. *Nature*, 2003, 424: 147–151.
- [42] Singh A K, Gurtu S, Jambunathan R. Phylogenetic relationships in the genus *Arachis* based on seed protein profiles [J]. *Euphytica*, 1994, 74: 219–255.
- [43] Mennella G, Sanajà V O, D'Alessandro A, Milone M, Perrone D. HPLC analyses of seed storage proteins reveal polymorphism in Italian common bean (*Phaseolus vulgaris* L.) ecotypes [J]. *Euphytica*, 2003, 134: 85–95.
- [44] Thorne J, Toro C O, Vargas J, Debouck D G. Variability in Andean *nuña* common beans (*Phaseolus vulgaris*, Fabaceae) [J]. *Econ Bot*, 1995, 49: 78–95.
- [45] Mennella G, Sanajà V O, Tonini A, Magnifico V. Seed storage protein characterization of *Solanum* species and of cultivars and androgenetic lines of *S. melongena* L. by SDS-PAGE and AE-HPLC [J]. *Seed Sci Technol*, 1999, 27: 23–35.
- [46] Mennella G, Sanajà V O, D'Alessandro A, Coppola R, Poma I. Tomato ecotype characterization by anionic exchange-high performance liquid chromatography analysis of endosperm seed proteins [J]. *Acta Hortic*, 2001, 546: 453–457.
- [47] Vladova R, Pandeva R, Petcolicheva K. Seed storage proteins in *Capsicum annuum* cultivars [J]. *Biol Plant*, 2000, 43: 291–295.
- [48] Syros T, Yupsanis T, Economou A. Fractionation and electrophoretic patterns of storage protein of *Ebenus cretica*. A preliminary survey as a tool in taxonomy [J]. *Biol Plant*, 2002, 46: 435–443.
- [49] Bianchi-Hall C M, Keys R D, Stalker H T. Use of seed protein profiles to characterize peanut cultivars [J]. *Peanut Science*, 1994, 21: 152–159.
- [50] Singh A K, Subrahmanyam P, Gurtu S. Variation in a wild groundnut species, *Arachis duranensis* Krapov. & W. C. Gregory [J]. *Genet Res Crop Evol*, 1996, 43: 135–142.
- [51] Bertozzi M R, Valls J F M. Seed storage protein electrophoresis in *Arachis pintoi* and *A. repens* (Leguminosae) for evaluating genetic diversity [J]. *Genet Resour Crop Evol*, 2001, 48: 121–130.
- [52] Alvarez J B, Caballero L, Ureña P, Vacas M, Martín L M. Characterisation and variation of morphological traits and storage proteins in Spanish emmer wheat germplasm (*Triticum dicoccum*) [J]. *Genet Resour Crop Evol*, 2006, 00: 1–8.
- [53] Han F P, Liu B, Fedak G, Liu Z H. Genomic constitution and variation in five partial amphiploids of wheat—*Thinopyrum intermedium* as revealed by GISH, multicolor GISH and seed storage protein analysis [J]. *Theor Appl Genet*, 2004, 109: 1070–1076.
- [54] Yu C Y, Hu S W, Zhao H X, Guo AG, Sun G L. Genetic distances revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.) [J]. *Theor Appl Genet*, 2005, 110: 511–518.
- [55] Shutov A D, Kakhovskaya I A, Braun H, Bäumlein H, Müntz K. Legumin-like and vicilin-like seed storage proteins: evidence for a common single-domain ancestral gene [J]. *J Mol Evol*, 1995, 41: 1057–1069.
- [56] Shutov A D, Braun H, Chesnokov Y V, Horstmann C, Kakhovskaya I A, Bäumlein H. Sequence peculiarity of gnetalean legumin-like seed storage proteins [J]. *J Mol Evol*, 1998, 47: 486–492.
- [57] Bianchi-Hall C M, Keys R D, Stalker H T, Murphy J P. Diversity of seed storage protein patterns in wild peanut (*Arachis*, Fabaceae) species [J]. *Plant Syst Evol*, 1993, 186: 1–15.
- [58] Teraishi M, Takahashi M, Hajika M, Matsunaga R, Uematsu Y, Ishimoto M. Suppression of soybean b-conglycinin genes by a dominant gene, *Scg-1* [J]. *Theor Appl Genet*, 2001, 103: 1266–1272.
- [59] Hoffman L M, Donaldson D D, Herman E M. A modified storage protein is synthesized, processed, and degraded in the seeds of transgenic plants [J]. *Plant Mol Biol*, 1988, 11: 717–729.
- [60] Pueyo J J, Chrispeels M J, Herman E M. Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole [J]. *Planta*, 1995, 196: 586–596.
- [61] Kim C S, Kamiya S, Sato T, Utsumi S, Kito M. Improvement of nutritional value and functional properties of soybean glycinin by protein engineering [J]. *Protein Eng*, 1990, 3: 725–731.
- [62] Utsumi S, Kitagawa S, Katsume T, Kang I J, Gidamis A B, Takaiwa F, Kito M. Synthesis, processing and accumulation of modified glycinins of soybean in the seeds, leaves and stems of transgenic tobacco [J]. *Plant Sci*, 1993, 92: 191–202.
- [63] Kjemtrup S, Herman E M, Chrispeels M J. Correct post translational modification and stable vacuolar accumulation of phytohemagglutinin engineered to contain multiple methionine residues [J]. *Eur J Biochem*, 1994, 226: 355–391.
- [64] Altenbach S B, Pearson K W, Meeker G, Staraci L C, Sun S S M. Enhancement of the methionine content of seed proteins by the expression of a chimeric gene encoding a methionine-rich protein in transgenic plants [J]. *Plant Mol Biol*, 1989, 13: 513–522.
- [65] Nordlee J A, Taylor S L, Townsend J A, Thomas L A, Bush R K. Identification of a Brazil-nut allergen in transgenic soybeans [J]. *N Engl J Med*, 1996, 334: 688–692.
- [66] Karchi H, Shaul O, Galili G. Seed-specific expression of a bacterial desensitized aspartate kinase increases the production of seed

- threonine and methionine in transgenic tobacco [J]. *Plant J*, 1993, 3:721 – 727.
- [67] Sharma S B, Hancock K R, Ealing P M, White D W R. Expression of a sulfur-rich maize seed storage protein, δ -zein, in white clover (*Trifolium repens*) to improve forage quality [J]. *Mol Breed*, 1998, 4:435 – 448.
- [68] Fiedler U, Phillips J, Artsenko O, Conrad U. Optimization of scFv antibody production in transgenic plants [J]. *Immunotechnology*, 1997, 3:205 – 216.
- [69] Hoffman L M, Donaldson D D, Bookland R, Rashka K, Herman E M. Synthesis and protein body deposition of maize 15-kD zein in transgenic tobacco seeds [J]. *EMBO J*, 1987, 6:3213 – 3221.
- [70] Firek S, Draper J, Owen M R L, Gandeche A, Cockburn W, Whitelam G C. Secretion of a functional single-chain Fv protein in transgenic tobacco plants and cell suspension cultures [J]. *Plant Mol Biol*, 1993, 23:861 – 870.
- [71] Hood E E, Witcher D R, Maddock S, Meyer T, Basczynski C, Bailey M, Flynn P, Register J, Marshall L, Kulisek E, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehagh R J, Herman R, Kappel W K, Ritland D, Li C P, Howard J A. Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification [J]. *Mol Breed*, 1997, 3:291 – 306.
- [72] Witcher D R, Hood E E, Peterson D, Bailey M, Bond D, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehagh R, Kappe W, Register J, Howard J A. Commercial production of β -glucuronidase (GUS): a model system for the production of protein in plants [J]. *Mol Breed*, 1998, 4:301 – 312.
- [73] Vandekerckhove J, Van Damme J, Van Lijsebettens M, Botterman J, de Block M, Vandewiele M, de Clerq A, Leemans J, Van Montagu M, Krebbers E. Enkephalins produced in transgenic plants using modified 2S seed storage proteins [J]. *Biotechnology (N Y)*, 1989, 7:929 – 932.
- [74] Parmenter D L, Boothe J G, Van Rooijen G J H, Yeung E C, Moloney M M. Production of biologically active hirudin in plant seeds using oleosin partitioning [J]. *Plant Mol Biol*, 1995, 29: 1167 – 1180.
- [75] Leite A, Kemper E L, Da Silva M J, Luchessi A D, Siloto R M P, Bonaccorsi E D, El-Dorry H F, Arruda P. Expression of correctly processed human growth hormone in seeds of transgenic tobacco plants [J]. *Mol Breed*, 2000, 6:47 – 53.
- [76] Ding S H, Huang L Y, Wang Y D, Sun H C, Xiang Z H. High-level expression of basic fibroblast growth factor in transgenic soybean seeds and characterization of its biological activity [J]. *Biotechnol Lett*, 2006, 28:869 – 875.