Plant Physiol. (1995) 109: 777–786

The Bean Seed Storage Protein β-Phaseolin Is Synthesized, Processed, and Accumulated in the Vacuolar Type-II Protein Bodies of Transgenic Rice Endosperm[1]

Zhenwei Zheng[2], Kazuhiko Sumi, Kunisuke Tanaka, and Norimoto Murai*

Department of Plant Pathology and Crop Physiology, Louisiana State University (LSU), and LSU Agricultural Center, Baton Rouge, Louisiana 70803–1720 (Z.Z., N.M.); and Department of Agricultural Chemistry, Kyoto Prefectural University, Kyoto 606, Japan (K.S., K.T.)

The seed storage protein β-phaseolin of the common bean (*Phaseolus vulgaris* L.) was expressed in the endosperm of transgenic rice (*Oryza sativa* L.) plants. The 5.1- or 1.8-kb promoter fragment of the rice seed storage protein glutelin Gt1 gene was fused transcriptionally to either the genomic or cDNA coding sequence of the β-phaseolin gene. The highest quantity of phaseolin estimated by enzyme-linked immunosorbent assay was 4.0% of the total endosperm protein in the transgenic rice seeds. The phaseolin trait was segregated as a single dominant trait with a positive gene dosage effect and was stably inherited through three successive generations. Both phaseolin genomic and cDNA coding sequences were used to synthesize four isoforms of mature phaseolin protein with apparent molecular masses of 51, 48, 47, and 45 kD. Enzyme deglycosylation experiments indicated that the 51-kD form contains high-mannose N-glycans; the 48- and 47-kD forms have further modified N-glycans; and the 45-kD form is a nonglycosylated protein. Immunolabeling studies using light and electron microscopy demonstrated that phaseolin accumulates primarily in the vacuolar type-II protein bodies located at the periphery of the endosperm near the aleurone layer. We discuss the implications of these results on nutritional improvement of rice grains.

Cereal and legume grains are major sources of nutrition for humans, cattle, and other animals. However, simultaneous consumption of cereal and legume grains is necessary to complement the proper composition of essential amino acids for human nutrition. Legume seeds are deficient in sulfur-containing essential amino acids (Met and Cys), whereas cereals have low contents of Lys and Ile. The traditional breeding approach has not been very successful in overcoming the amino acid limitations of cereal and legume grains (Mertz et al., 1964; Loesch et al., 1976; Delaney and Bliss, 1991). An alternative approach using protein engineering and gene transfer for amino acid enhancement has been initiated recently (Ohtani et al., 1991; Dyer et al., 1993).

Phaseolin is a 7S globulin and is a major component of common bean seeds. In developing bean cotyledons, phaseolin is synthesized on polysomes of the RER (Bollini and Chrispeels, 1978). The nascent polypeptide enters the lumen of the ER, where the signal peptide is cleaved cotranslationally (Chrispeels, 1991). The polypeptide is glycosylated by a high-Man glycan, folded into the proper tertiary structure, and assembled into a trimer. The last two processes may be mediated by a binding protein-like chaperonin (D'Amico et al., 1992; Vitale et al., 1995). Phaseolin trimer is transported by bulk flow from the ER to the Golgi apparatus, where some high-Man glycans are converted to complex glycans (Sturm et al., 1987). Glycosylated phaseolin trimer is targeted to the vacuole and packaged into the protein body.

Rice (*Oryza sativa* L.) seed storage proteins consist of 80% glutelin, 5% prolamine, and others. Rice is unique in that two types of protein bodies coexist in the endosperm (Yamagata et al., 1982). The type-I protein body is spherical with a concentric ring structure under EM and contains prolamine, whereas the type-II protein body does not have this ring structure and is rich in glutelin and globulin. Prolamine and glutelin transcripts are spatially segregated into two distinct RER membranes (Li et al., 1993a). Prolamine mRNAs are attached directly to the membrane of prolamine protein bodies (the so-called protein body ER). Binding protein facilitates prolamine folding and assembly into the type-I protein bodies (Li et al., 1993b). Glutelin mRNAs are translated on the cisternal ER surface. The nascent glutelin polypeptides are then processed in the ER lumen, transported to the Golgi apparatus, and targeted to the type-II vacuolar protein bodies. Although developing rice endosperm produces similar steady-state levels of prolamine and glutelin mRNAs, glutelin mRNAs seem to be much more efficiently translated, leading to much higher accumulation of glutelin than prolamine (Okita et al., 1989).

Rice is the first cereal crop in which fertile transgenic plants were generated with ease and certainty (Shimamoto et al., 1989; Hayashimoto et al., 1990). Progressive 5′-deletion analysis of the Gt1 gene in transgenic rice demonstrated that the 5.1- and 1.8-kb promoters directed tempo-

[1] This work was supported in part by grants from the Louisiana Education Quality Support Fund (1987–90)-RD-A-6 and (1991–94)-RD-B-7 to N.M.

[2] Present address: Department of Stomatology, University of California-San Francisco, San Francisco, CA 94143–0517.

* Corresponding author; e-mail nmurai@lsuvax.sncc.lsu.edu; fax 1–504–388–1415.

Abbreviations: Endo H, endoglycosidase H; FITC, fluorescein isothiocyanate; PNGase F, peptide N-glycosidase F.

rally regulated and endosperm-specific expression of the GUS reporter (Zheng et al., 1993). Since the overall content of Lys (3.4%) in rice grains makes it the first limiting amino acid, as in other cereals, the bean β-phaseolin, with a relatively high Lys content (6.0%), was introduced into transgenic rice to improve the nutritional quality of rice grains. Here we report the genetic, biochemical, and subcellular localization studies of the introduced phaseolin in transgenic rice endosperms. The phaseolin constituted up to 4% of the total rice endosperm protein. The phaseolin gene and its high expression were inherited through three successive generations as a single dominant trait with a positive gene dosage effect.

MATERIALS AND METHODS

Construction of Phaseolin Fusion Genes

Both genomic and cDNA coding sequences of the common bean (*Phaseolus vulgaris* L.) β-phaseolin gene were placed under the control of either the β-phaseolin promoter (Slightom et al., 1983) or the rice (*Oryza sativa* L.) glutelin Gt1 gene promoter (Okita et al., 1989). A 1.3-kb *Hin*dIII-*Eco*RI fragment containing the phaseolin 782-bp promoter was ligated to the *Eco*RI/*Bam*HI fragment containing the whole genomic or cDNA coding sequence and downstream region from AG-pPVPh3.8 or AG-pPVPh3.8/cDNA, resulting in pTRA321 and pTRA322, respectively (Fig. 1). The plasmids AG-pPVPh3.8 and AG-pPVPh3.8/cDNA were kindly provided by Dr. J.L. Slightom (Upjohn Co., Kalamazoo, MI) (Cramer et al., 1985).

The phaseolin and glutelin promoter fragments were fused transcriptionally to the phaseolin coding sequences via a *Bgl*II site. The *Bgl*II site was introduced at a point 13 or 19 bp in front of the initiation codon of the glutelin and phaseolin coding sequences, respectively (Zheng et al., 1993). The 1.8-kb Gt1 promoter was isolated as an *Eco*RI-*Bgl*II fragment and fused to the phaseolin genomic or cDNA coding sequences in a *Bgl*II/*Bam*HI fragment, forming pTRA323 and pTRA324, respectively. The 1.8-kb Gt1 promoters in pTRA323 and pTRA324 were replaced by the 5.1-kb Gt1 promoters in *Kpn*I-*Bgl*II fragments, resulting in pTRA327 and pTRA328, respectively. A hygromycin-resistance gene from pTRA141 (Zheng et al., 1991) was cloned into these transformation vectors as a selectable marker for rice transformation.

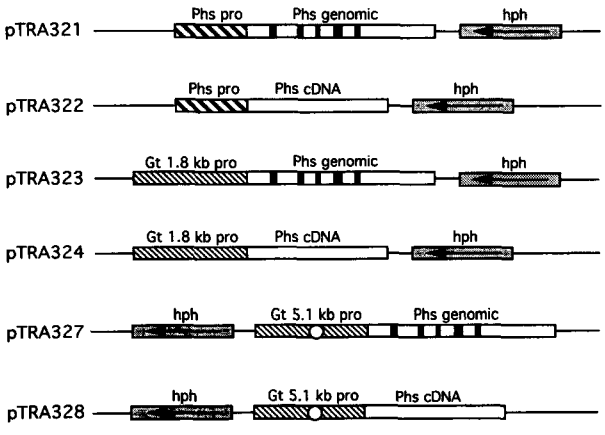

Figure 1. Abbreviated structures of six phaseolin fusion genes used for rice transformation. The phaseolin genomic or cDNA coding sequence was placed under the control of the native bean β-phaseolin or resident rice glutelin promoters. Black boxes indicate introns in coding sequences. Arrows indicate transcription direction of the *hph* gene. Gt 1.8 or 5.1 kb pro, The 5.1- or 1.8-kb promoter fragment from the rice seed storage protein glutelin gene (Gt1); Phs pro, the 783-bp promoter fragment from the bean seed storage protein β-phaseolin gene; Phs genomic or cDNA, the genomic or cDNA coding sequence and downstream region of the phaseolin gene; hph, hygromycin phosphotransferase gene.

Transgenic Rice Production

The above-described vectors were used to transform rice protoplasts (cv Nipponbare) after PEG treatment as described by Li et al. (1990). Nurse cells of rice cv Taipei 309 or cv OC were used in protoplast culture.

ELISA Estimates of Phaseolin Quantity

Salt-soluble globulins, including phaseolin, were extracted from approximately 10 mg of rice endosperm powder by homogenization in 1.0 mL of PBS (137 mM NaCl, 3 mM KCl, 12 mM PO_4, pH 7.4) with a Brinkmann homogenizer. After vigorous shaking at 4°C for 30 to 60 min, the homogenate was centrifuged. The recovered supernatant was collected as a salt-soluble protein fraction and stored for later use at −20°C. Pellets containing prolamine and glutelin were solubilized in 1 N NaOH and used for protein assay.

The amount of phaseolin in PBS extracts was estimated by ELISA using rabbit anti-phaseolin antibody conjugated to alkaline phosphatase (Engvall and Perlmann, 1972). Total protein quantity was estimated by the Bradford method using BSA as a standard (Bradford, 1976).

Southern Blot Hybridization Analysis

Total DNA was extracted from seedlings or mature leaves of transgenic rice plants (Rogers and Bendich, 1985). Crude DNA preparations were purified by CsCl equilibrium density centrifugation (Hayashimoto et al., 1990). Ten micrograms of rice DNA were digested with *Eco*RI or/and *Hin*dIII, extracted with a phenol/chloroform mixture, and precipitated by ethanol. The amount of recovered DNA was estimated spectrophotometrically. Five micrograms of digested DNA were fractionated electrophoretically on a 0.7% (w/v) agarose gel. Fractionated DNA was transferred to a Nytran membrane using procedures suggested by Schleicher & Schuell. The 0.75-kb *Eco*RI-*Xba*I fragment representing the 5′ half of the phaseolin coding sequence was labeled with ^{32}P by an oligonucleotide labeling procedure (Sen and Murai, 1991) and used as a probe for Southern blot hybridization.

Western Blots and Two-Dimensional Gel Electrophoresis

For western blots, protein samples were fractionated by SDS-PAGE on a 12% (w/v) gel and electrophoretically

blotted onto a nitrocellulose membrane. Phaseolin was identified using rabbit anti-phaseolin antibody conjugated to alkaline phosphatase and visualized using 5-bromo-4-chloro-3-indolyl phosphate and nitroblue tetrazolium (Ey and Ashman, 1986).

Two-dimensional gel electrophoresis was performed using a Bio-Rad Protean II system according to O'Farrell (1975). Protein samples were first separated isoelectrically in a 1:4 mixture of Bio-lyte pH3–10 and pH5–7 (Bio-Rad), followed by SDS-PAGE on a 12% (w/v) gel.

Enzyme Digestion Analysis of the Glycosyl Moiety of Phaseolin

To minimize protein degradation during extraction, globulins were extracted with acid salt buffer (0.5 M NaCl, 0.5 M Gly, 0.02% [w/v] NaN_3, 2 mM PMSF, pH 2.4). One gram of endosperm powder was homogenized in 5 mL of the buffer and vigorously shaken at room temperature for 1 h (Delaney and Bliss, 1991). Phaseolin was precipitated from the supernatant by addition of 4 volumes of water (Sun and Hall, 1975). Initially, 1.5 volumes of cold, distilled water were added, followed by an additional 2.5 volumes of water after a second centrifugation. The pH was adjusted to 5.3 to enhance the phaseolin precipitation. Pellets of phaseolin precipitates collected were lyophilized and dissolved in PBS.

Endo H and PNGase F were used according to the instructions of Boehringer Mannheim. Approximately 2 μg of phaseolin were denatured by boiling for 5 min in the presence of 0.02% (w/v) (Endo H) or 0.2% (PNGase F) SDS before digestion.

Immunolabeling Studies by Light Microscopy and EM

Immature rice seeds were harvested approximately 10 DAF. Endosperms were sliced transversely into several pieces, fixed in 3% (w/v) glutaraldehyde at 4°C overnight, and washed five times with 0.1 M sodium phosphate (pH 7.2). After sequential dehydration steps in ethanol, fixed endosperm preparations were washed twice for 20 min each and embedded overnight at 45°C in LR White resin (hard grade, London Resin Co. Ltd, London, UK). Embedded endosperm samples were cut with a microtome into 2-μm-thick sections for light microscope observation and into 50- to 70-nm sections for EM study.

Endosperm sections were hydrated for 30 min in 20 mM Tris-HCl (pH 8.2) containing 225 mM NaCl, 20 mM NaN_3, and 0.1% (w/v) BSA. Hydrated sections were first treated overnight with a primary antibody, either rabbit anti-phaseolin or mouse anti-glutelin antibody. Sections were reacted for 2 h with a secondary antibody, goat anti-rabbit IgG conjugated with FITC or 30-nm gold particles, or goat anti-mouse IgG conjugated with 15-nm gold particles (Amersham), respectively.

Thin sections for EM study were stained with 2% (w/v) uranyl acetate at 45°C for 30 min and with 2% (w/v) lead acetate at room temperature for 5 min and examined under a transmission electron microscope (Nippon Electric [Tokyo, Japan] JEM-1200EX). An Olympus IMT-2 light microscope was used with a G520 filter to eliminate wavelengths greater than 540 nm.

RESULTS

Six Phaseolin Fusion Genes Were Introduced into Transgenic Rice

Six different phaseolin fusion genes were constructed and used for rice protoplast transformation in this study. Figure 1 lists the abbreviated structures of the plasmid vectors tested. We used both the genomic and cDNA coding sequences of the β-phaseolin gene because of the uncertainty concerning the effective processing of dicotyledonous bean introns in monocotyledonous rice nuclei. Similarly, we chose to test the promoters of both bean and rice seed storage protein genes because there was no pertinent information available concerning the efficiency of bean promoter expression in rice at the time of gene construction.

A 782-bp phaseolin promoter was fused transcriptionally to the 1990-bp phaseolin genomic coding and 1100-bp downstream regions in pTRA321 or to the 1475-bp phaseolin cDNA coding and 1100-bp downstream regions in pTRA322. Thus, an authentic phaseolin gene from the native common bean is recreated in pTRA321. A 1.8-kb promoter fragment of the rice seed storage protein glutelin gene Gt1 was fused to either the phaseolin genomic coding sequence in pTRA323 or the phaseolin cDNA coding sequence in pTRA324. The 1.8-kb Gt1 fragments in pTRA323 and pTRA324 were replaced with the 5.1-kb Gt1 promoter in pTRA327 and pTRA328, respectively.

Approximately 200 hygromycin-resistant calli were obtained per million protoplasts after PEG-mediated transformation, nurse culture, and selection. Plantlets were regenerated from calli via somatic embryogenesis with up to 20% efficiency. The number of transgenic plants grown to maturity in the greenhouse was 11, 14, 32, and 27 for constructs 323, 324, 327, and 328, respectively. Approximately two-thirds of the plants set viable seeds, and phaseolin was detectable by ELISA in one-half of the fertile plants tested. Transgenic plants were numbered according to the vectors used for transformation.

The 5.1-kb Glutelin Promoter Directed the Highest Phaseolin Expression in Seeds

The total quantity of phaseolin accumulated in seeds of primary transgenic plants was estimated by ELISA using a polyclonal antibody against bean phaseolin. One end of the endosperm was sliced to take endosperm sections for ELISA assay and the other end, containing the embryo, was used to germinate F_1 seedlings. Endosperm samples were pooled from 10 seeds for each plant, since the introduced phenotype is expected to segregate among F_1 progeny. Phaseolin content in pooled seeds of each transgenic plant was expressed as a percentage of total endosperm protein and is represented by dots in Figure 2. A wide variation in phaseolin expression level was observed among the plants transformed with the same DNA construct. When four

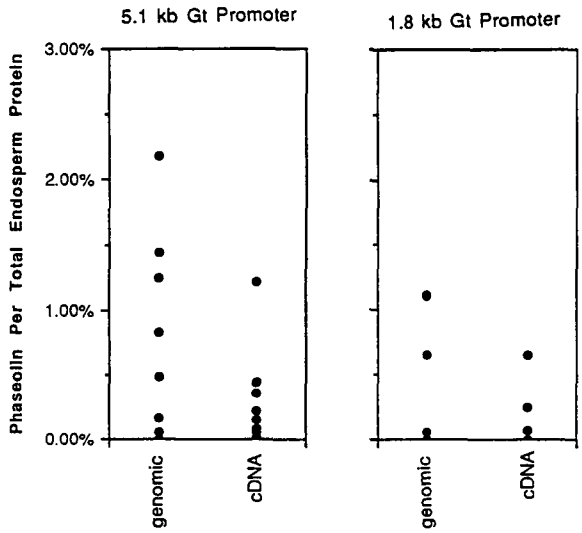

Figure 2. Phaseolin expression in endosperms compared among four groups of primary transgenic rice plants containing four different phaseolin fusion genes all under the control of the rice glutelin promoter. The phaseolin quantity was estimated by ELISA, and phaseolin contents are expressed as a percentage of total endosperm protein in 10 pooled seeds per plant. Each dot represents the phaseolin expression level in an independently transformed plant.

different phaseolin fusion genes all under the control of the rice glutelin promoters were compared, the 5.1-kb Gt1 promoter directed higher phaseolin expression than the 1.8-kb Gt1 promoter. The phaseolin genomic coding sequence, which contains five small introns, gave numerically higher, but not statistically significant, expression than the intron-free cDNA coding sequence under the control of either the 5.1- or the 1.8-kb Gt1 promoter. This result suggests that the existence of introns into the dicot bean coding sequences did not interfere with gene expression in monocot cells. No phaseolin was detectable in either embryos or endosperms of any seeds of rice plants containing the phaseolin coding sequences under the control of the native bean promoter, indicating that the dicot bean promoter worked poorly or did not function at all in the monocot rice nuclei. The highest phaseolin quantity in pooled seeds was 2.2% of total endosperm protein of plant 327–3. However, when individual seeds were analyzed for phaseolin level, the phaseolin content in the homozygous seed 327–3–12 reached 4.0% of the total endosperm protein (Fig. 3).

Progeny Segregation Analysis and Positive Gene Dosage Effect of the Phaseolin Trait

Segregation of the phaseolin trait was investigated by ELISA among individual seeds of the three selfed primary transgenic plants 323–5, 327–3, and 327–9. Figure 4 illustrates the result from such analysis of plant 327–9. The phaseolin-positive and -negative phenotypes segregated in a 3:1 ratio among 30 to 50 seeds tested in all three plants, indicating that the transferred gene was inserted into a single chromosomal locus or closely linked loci in a chromosome. Moreover, there appear to be three quantitatively distinct levels of phaseolin expression among the phaseolin-positive progeny, each level representing roughly one-fourth of the total seed population. This observation suggests a positive gene-dosage effect on the phaseolin expression, with phaseolin content reflecting the presence of one, two, and three copies of the phaseolin gene in the endosperm ($3n$). To test the genotype of F_1 plants, seeds of primary transgenic plants 323–5 and 327–3 were germinated and grown to mature plants, and segregation of the phaseolin trait was analyzed in individual F_1 seeds (Fig. 3). This analysis demonstrated that phaseolin contents in endosperms were higher in homozygous seeds (no segregation in F_1 seeds) than in heterozygous seeds (3:1 segregation in F_1 seeds). Furthermore, no segregation of the phaseolin trait was observed in the F_2 generation derived from 323–5–14 and 327–3–12, confirming that they are sta-

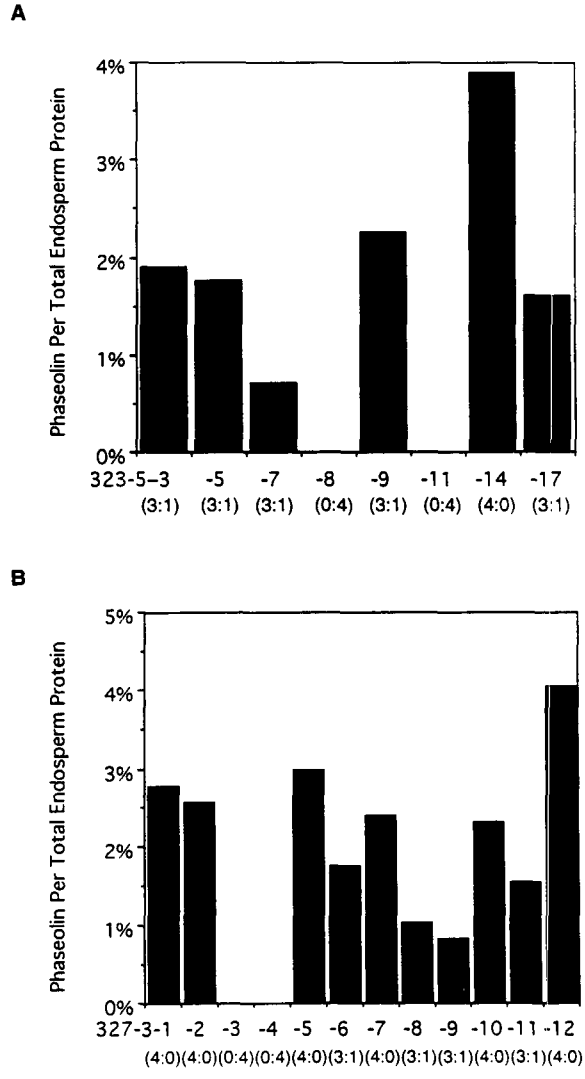

Figure 3. Positive gene-dosage effect and stable inheritance of the phaseolin trait. A, Phaseolin contents in endosperms of primary transgenic plant 323–5 and progeny segregation analysis (phaseolin positive:negative) in F_1 seeds. B, The same for primary transgenic plant 327–3.

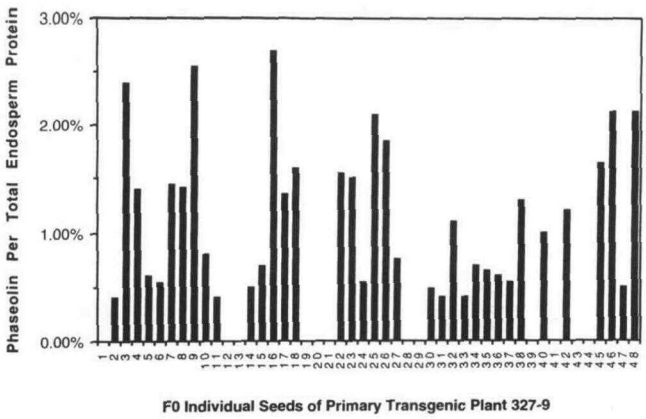

Figure 4. Progeny segregation analysis of the phaseolin trait in individual seeds of primary transgenic plant 327–9. The phaseolin quantity of each endosperm was estimated by ELISA, and phaseolin contents are expressed as a percentage of total endosperm protein.

ble, homozygous lines for the phaseolin trait. These results demonstrated that the introduced phaseolin gene was stably inherited as a single Mendelian trait through three successive generations and also indicated a positive gene-dosage effect on the expression level of phaseolin in transgenic rice endosperms.

Physical Structure of Transferred Phaseolin Genes in the Rice Genome

The physical structures of the introduced phaseolin gene in the genome of representative plants were probed by Southern blot hybridization analysis (Fig. 5). DNA was isolated from two F_1 plants, 323–5–14 and 327–3–10, that are homozygous for the phaseolin trait (Fig. 3) from offspring of the cross 323–5–14 × 327–3–10, and from two F1 plants, 327–3–3 and 327–3–4, with no detectable phaseolin expression. Rice DNA was digested with *Hin*dIII and probed with a ^{32}P-labeled 0.75-kb *Eco*RI/*Xba*I fragment derived from the 5′ half of the phaseolin coding sequence. The transformation vector pTRA323 contains two *Hin*dIII sites (Fig. 5B), and *Hin*dIII digestion should generate a 4.8-kb fragment of the intact phaseolin fusion gene comprising the 1.8-kb Gt1 promoter and the 3.0-kb genomic coding/downstream region (Fig. 5A, lanes 1 to 3). DNA from plant 323–5–14 had a 4.8-kb signal (Fig. 5A, lane 4), indicating the existence of the intact phaseolin fusion gene in the genome of that plant.

Transformation vector pTRA327 has a single *Hin*dIII site (Fig. 5B), and *Hin*dIII digestion should generate an 11-kb fragment from the vector (Fig. 5A, lanes 8–10). However, DNA from plant 327–3–10 had a 16-kb signal (Fig. 5A, lane 5), indicating that a second *Hin*dIII site is located 5 kb from the insertion site of the transferred DNA in the rice genome (Fig. 5B). The hybridization signal is that of less than one copy, and this is likely due to poor transfer onto the nitrocellulose membrane. Detection of a single band in the genome of plant 327–3–10 suggests that this homozygous F_1 plant carries a single copy of the transferred phaseolin gene per haploid genome (Fig. 5A, lane 5). DNA from progeny derived from a cross of two homozygous F_1 lines, 327–3–10 × 323–5–14, had the expected 4.8- and 16-kb signals, indicating that the phaseolin genes from both parents were stably transmitted to the genome of the progeny (Fig. 5A, lane 6). No signal was detectable in DNA of the phaseolin-negative plants 327–3–3 and 327–3–4, indicating that the transferred phaseolin gene was segregated out of these progeny plants (Fig. 5A, lane 7). These results demonstrate that the physical structure of the introduced phaseolin gene as well as the expression level of its trait were stably inherited through three successive generations of transgenic rice plants.

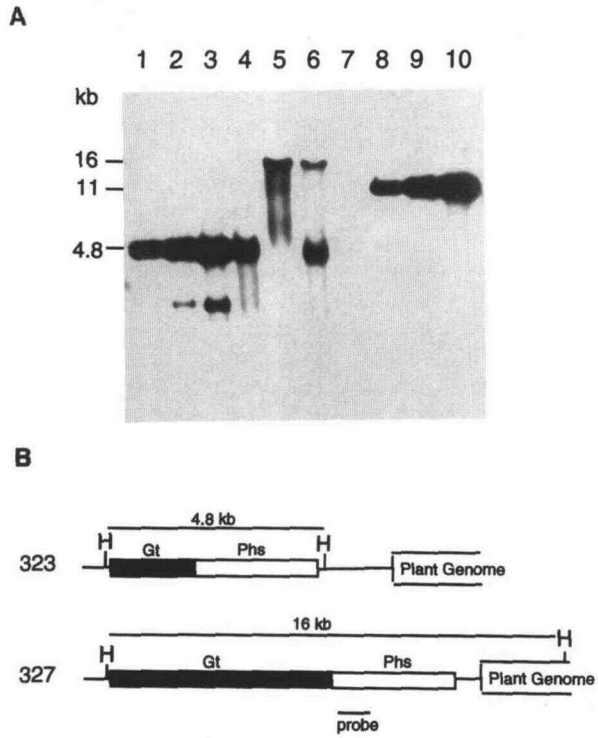

Figure 5. Southern blot hybridization analysis of the physical structure of the introduced phaseolin gene in the rice genome. A, Homozygous progeny derived from primary transgenic rice plants 323–5 and 327–3. Five micrograms of total genomic DNA were digested with *Hin*dIII and separated on a 0.7% (w/v) agarose gel. A 0.75-kb fragment isolated from the phaseolin genomic coding sequence was used as the probe as shown in B. Lanes 1 to 3, Copy number controls of transformation vector pTRA323, equivalent to 1, 5, and 10 copies per haploid rice genome; lane 4, DNA from homozygous plant 323–5–14; lane 5, DNA from homozygous plant 327–3–10; lane 6, DNA from offspring of cross 323–5–14 × 327–3–10; lane 7, DNAs from phaseolin-negative plants 327–3–3 and 327–3–4; lanes 8 to 10, copy number controls of transformation vector pTRA327, equivalent to 1, 5, and 10 copies per haploid rice genome. B, Deduced physical structure of the introduced phaseolin genes in the genome of plant 323–5–14 and 327–3–10. Gt, Gt1 1.8- or 5.1-kb promoter; Phs, phaseolin genomic coding sequence; H, *Hin*dIII. The sizes of *Hin*dIII fragments are marked.

Glycosylated Mature Phaseolin Was Synthesized in Transgenic Rice Endosperms

The biochemical properties of phaseolin synthesized in the endosperm of heterologous rice were compared with those of the authentic phaseolin derived from cotyledons of native common beans. Cultivar Tendergreen of the common bean carries six to eight genes per haploid genome for α- and β-type phaseolins, which are heterogeneous in molecular mass and pI (Fig. 6A). Two-dimensional gel electrophoresis resolved bean phaseolin into two distinguishable spots centered on pH 5.7 in IEF and into three molecular masses, 51, 48, and 45.5 kD, in SDS-PAGE (Fig. 6A). The smaller two polypeptides correspond to the glycosylated (48 kD) and nonglycosylated form (45.5 kD) of the β-phaseolin gene product (Fig. 7, lanes 3) (Slightom et al., 1983). Mature phaseolin contains two potential N-glycosylation sites, at positions Asn^{228} and Asn^{317}. The 48-kD band represents two isoforms of glycosylated β-phaseolin; one contains high-Man glycans at both positions, $Man_7(GlcNAc)_2$ at Asn^{228} and $Man_9(GlcNAc)_2$ at Asn^{317}, and the other contains a complex glycan, $Xyl\text{-}Man_3(GlcNAc)_2$, at Asn^{228} and no modification at Asn^{317} (Sturm et al., 1987).

Transgenic rice plant 327–3 carries the β-phaseolin genomic coding sequence under the control of the 5.1-kb Gt1 promoter (Fig. 1). Phaseolin synthesized in the endosperm of this plant shows biochemical properties similar to but distinct from those of its bean counterpart in two-dimensional gel electrophoresis and western blots. Rice β-phaseolin focused on an essentially identical spot at around pH 5.7 in IEF, but resolved into four distinguishable isoforms with slight differences in size (Fig. 6B). SDS-PAGE analysis demonstrated four immunologically reactive bands of apparent molecular mass 51 (A), 48 (B), 47 (C), and 45 (D) kD (Fig. 7, A and B, lanes 1). No difference was observed in sizes of β-phaseolin polypeptides produced from genomic and cDNA coding sequences (data not shown), suggesting that the existence of dicot bean introns in the transferred gene did not interfere with proper expression of the protein product in monocot rice.

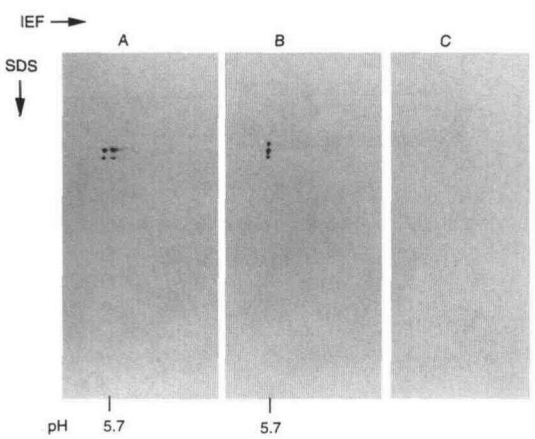

Figure 6. Two-dimensional gel electrophoresis analysis and subsequent western blot of phaseolin from transgenic rice endosperms. A, Two micrograms of purified phaseolin from native common bean. B, SDS-loading buffer extract from endosperm powder of primary transgenic plant 327–3; the extract should contain 2 μg of phaseolin as estimated by ELISA. C, Extract from endosperm of a nontransformed rice plant. Proteins were separated isoelectrically in a 1:4 mixture of Bio-lyte pH3–10 and pH5–7 (Bio-Rad) followed by SDS-PAGE on a 12% (w/v) gel.

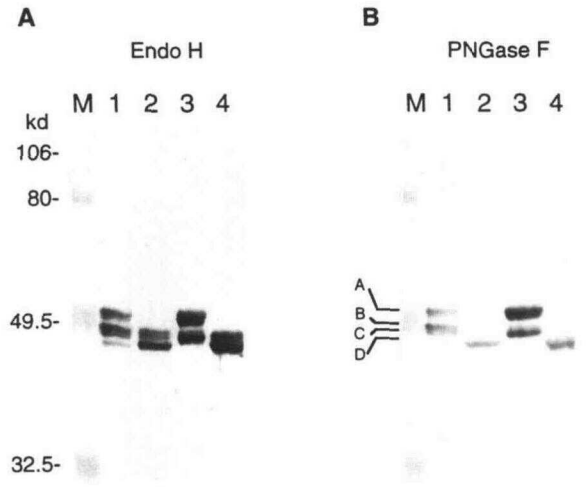

Figure 7. Enzyme digestion of glycosyl moieties of phaseolin synthesized in endosperm of transgenic plant 327–3. Proteins were separated by SDS-PAGE on a 12% (w/v) gel, and phaseolin polypeptides were visualized by western blot. A, Endo H digestion analysis. B, PNGase F digestion analysis. Lanes 1, No-enzyme control of protein extract from rice plant 327–3; lanes 2, enzyme digestion of protein extract from rice plant 327–3; lanes 3, no-enzyme control of purified bean phaseolin; lanes 4, enzyme digestion of purified bean phaseolin. M, Prestained Bio-Rad protein molecular mass markers.

Enzyme digestion of the N-glycosyl moieties showed that β-phaseolin synthesized in transgenic rice endosperm has oligosaccharide structures distinct from those in the native bean protein (Fig. 7) (Sengupta-Gopalan et al., 1985). Endo H and PNGase F were used to probe the chemical properties of oligosaccharide side chains attached to the Asn residues (Tarentino et al., 1989). Endo H hydrolyzes the glycosidic linkage between two terminal GlcNAc's, whereas PNGase F hydrolyzes the glycosylamide linkage between the terminal GlcNAc and the Asn amide nitrogen. The specificities of Endo H and PNGase F for N-glycan side chain structures differ: Endo H cleaves primarily high-Man structures, whereas PNGase F cleaves almost all N-glycan chains, including complex glycan structures. Band A contains a high-Man glycan structure at both positions Asn^{228} and Asn^{317}, since they are susceptible to Endo H digestion. However, because the molecular mass (51 kD) of band A is larger than that of the bean counterpart (48 kD), it should contain longer Man glycan chains. Bands B and C have complex glycans that are not susceptible to Endo H digestion (Fig. 7A, lane 2) but are susceptible to PNGase F digestion (Fig. 7B, lane 2). The difference in the glycosyl moieties between bands B and C is not known. After complete digestion with PNGase F, the upper three bands merged to a single band, D, indicating that the apparent size differences were caused by the differential glycosyla-

tion at the Asn residues. Band D had no apparent change in mobility in SDS-PAGE after PNGase F digestion; its apparent molecular mass (45 kD) is almost the same as the calculated molecular mass of the nonglycosylated mature form of β-phaseolin (45.2 kD) (Slightom et al., 1983). So band D should represent the nonglycosylated mature β-phaseolin polypeptide. A small amount of nonglycosylated β-phaseolin also exists in the native bean or when β-phaseolin is expressed in transgenic tobacco cotyledons (Sengupta-Gopalan et al., 1985). Rice β-phaseolin also bound to Con A-Sepharose to a similar extent as its bean counterpart, indicating that they are glyscosylated polypeptides (data not shown).

Subcellular Localization of β-Phaseolin in Transgenic Rice Seeds

Subcellular localization of β-phaseolin accumulated in transgenic rice endosperm was studied by immunocytochemical analysis under light microscopy and EM. The central rice endosperm possesses tightly packed starch grains and relatively few protein bodies, whereas the subaleurone cells have smaller starch grains and more abundant protein bodies. F_2 seeds of homozygous plant 323–5–14 were analyzed using immunofluorescence light microscopy (Fig. 8). Phaseolin was detected with rabbit anti-phaseolin antibody and goat anti-rabbit IgG conjugated with FITC. Figure 8B shows that the fluorescence signal from phaseolin was associated with the spherical bodies located at the peripheral protein-rich layer in starchy endosperms (i.e. subaleurone cells), suggesting that phaseolin was localized in the vacuolar protein bodies. The light microscopy study was further extended using double immunogold labeling under EM. Phaseolin was labeled with 30-nm gold particles (represented by large dots in Fig. 9) and glutelin was labeled with 15-nm gold particles (represented by small dots in Fig. 9). The analysis shows that heterologous phaseolin was located in the type-II protein bodies and not in the type-I protein bodies; this was also true of the resident glutelin (Fig. 9). However, the spatial localization of phaseolin in the type-II protein bodies is not as exclusive as that of glutelin; phaseolin signals occur in other subcellular spaces. Since we have used the signal peptide and downstream polypeptide sequence of mature β-phaseolin in this study, this result suggests that precise targeting of bean phaseolin to the type-II protein bodies was not readily accomplished using the heterologous bean signal in monocot rice cytoplasm.

DISCUSSION

Efficient Expression of Bean Phaseolin in Transgenic Rice Endosperms

High endosperm-specific expression of the bean seed storage protein phaseolin was demonstrated by placing the phaseolin coding and downstream region under the transcriptional control of promoters from the rice gene for the

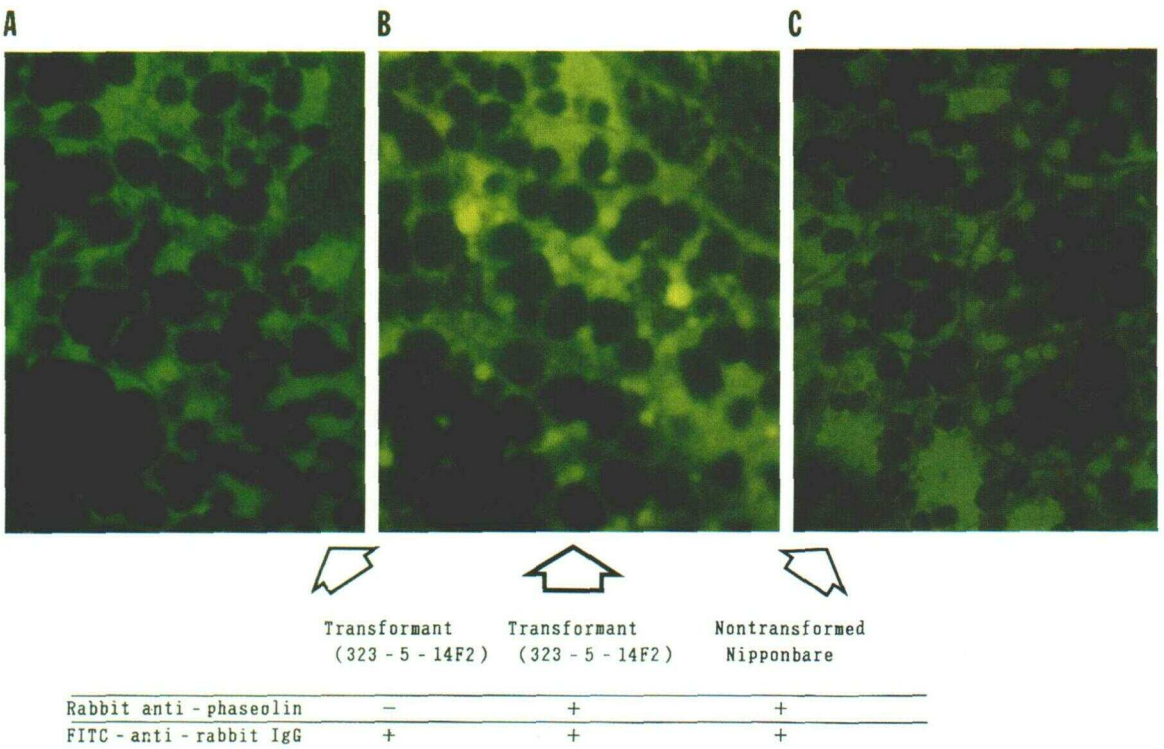

Figure 8. Immunochemical-fluorescence analysis of phaseolin localization in subaleurone cell layers of rice endosperm. A and B, Endosperms of F_2 plant 323–5–14 homozygous for the phaseolin trait. Phaseolin was first reacted with rabbit anti-phaseolin antibody (B, not A) and labeled with anti-rabbit IgG conjugated with FITC (A and B). C, Endosperm from nontransformed Nipponbare rice plants treated the same as those in B.

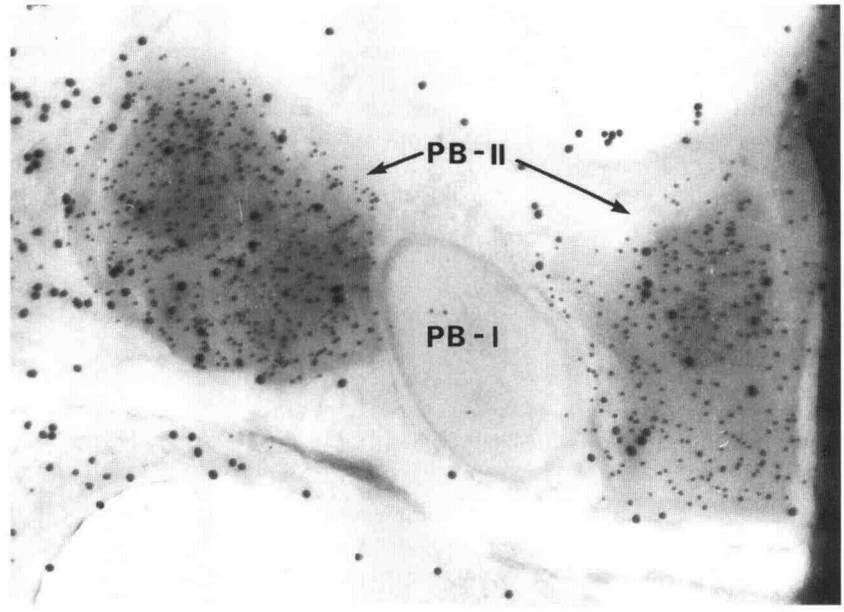

Figure 9. Electron microscopic observation of double immunogold labeling of phaseolin and glutelin accumulated in type-II protein bodies of endosperms of F2 plant 323–5–14. Large dots represent 30-nm gold particles for phaseolin and small dots represent 15-nm gold particles for glutelin. PB-I, Type-I vacuolar protein bodies; PB-II, type-II protein bodies.

seed storage protein glutelin (Okita et al., 1989). The 5.1-kb promoter directed higher expression than the 1.8-kb promoter, consistent with the previous results using the GUS reporter (Zheng et al., 1993). However, the difference between these two promoters was less than that previously observed, partly reflecting the nature of the two distinct reporter proteins, phaseolin versus GUS. The existence of introns in the coding region of the dicot bean gene did not interfere with the high-level protein expression in the monocot rice. This result contrasts with the previous report of inefficient intron processing of the monocot wheat pre-mRNA in the dicot tobacco (Keith and Chua, 1986). Studies of the requirements for efficient intron splicing have concluded that splicing in monocots is considerably more permissive than splicing in dicots (Goodall and Filipowicz, 1991). Splicing in dicots requires a high AT content in the intron, whereas monocot introns are GC rich. Gene expression in monocot cells was enhanced by the insertion of introns from both monocot and dicot genes, including those for maize alcohol dehydrogenase (Callis et al., 1987) and *Shrunken-1* (Vasil et al., 1989), rice actin (McElroy et al., 1990), and castor bean catalase (Tanaka et al., 1990). However, the existence of phaseolin introns in the dicot bean gene apparently did not increase the accumulation of protein products in rice endosperm.

Phaseolin synthesized in the monocot rice endosperms has a pI (pH 5.7) and molecular mass of the nonglycosylated form (45 kD) that are essentially identical with those of the authentic counterpart from cotyledons of native common beans. This suggests that the 24-residue signal peptide is cleaved co-translationally from the nascent translation product, generating the mature form of the β-phaseolin polypeptide. However, the structures of the N-linked oligosaccharide side chain at positions Asn^{228} and Asn^{317} of rice phaseolin were distinct from those of its bean counterpart. Three glycoforms, of 51, 48, and 47 kD, were generated in rice endosperms compared with two glycoforms of 48 kD in bean cotyledons and transgenic sunflower (Murai et al., 1983) and tobacco (Sengupta-Gopalan et al., 1985). The exact chemical nature of oligosaccharide side chains in these β-phaseolin glycoforms from rice endosperms remains to be elucidated. Even though the precise role that glycosylation plays in the mature phaseolin is not known, removal of glycosylation sites from phaseolin or phytohemagglutinin did not interfere with correct targeting of the nonglycosylated proteins into the vacuolar protein bodies, although it reduced protein accumulation in the protein bodies (Voelker et al., 1989; Busto et al., 1991).

The subcellular localization studies of phaseolin in rice endosperm demonstrated that phaseolin was located mainly in the type-II protein bodies. However, the spatial localization of phaseolin in the type-II protein bodies is not as exclusive as that of glutelin, suggesting that the signal peptide and mature polypeptide of phaseolin are not sufficient for targeting to the type-II protein bodies. Similarly, the maize seed storage protein zein was deposited to the subcellular structures similar to the type-I and -II protein bodies in transgenic tobacco seeds (Bagga et al., 1995). These observations suggest the critical role that the host protein sorting system plays in inter-organelle trafficking of passenger proteins.

Implications for Nutritional Improvement of Rice Grains

The physical structure of the introduced phaseolin genes as well as the high expression level were stably inherited through three successive generations. The ease with which the phaseolin trait is monitored by ELISA makes it possible to follow the agronomic stability of the phaseolin trait in many generations to come. The highest quantity of phaseolin was 4.0% of total endosperm protein in homozygous seeds of plants 323–5–14 and 327–12. We have taken two separate approaches to increasing the phaseolin content

and eventually the Lys content of rice grains. It should be possible to double the phaseolin content in endosperm by crossing the two homozygous lines and screening progeny that carry the phaseolin gene homozygous at two independent chromosomal loci. It should also be possible to increase the Lys content of phaseolin using structure-based protein design (Dyer et al., 1993) and to introduce the Lys-enhanced phaseolin gene into rice as was done for the Lys-enhanced zein (Ohtani et al., 1991).

Expression of phaseolin at the 4% level of total protein is theoretically sufficient to demonstrate an increased Lys content of total protein fraction based on the 6.0 mol% of Lys in β-phaseolin (Slightom et al., 1983). We have begun analysis of the composition of amino acids in the salt-soluble protein fraction from endosperms of homozygous plant 327–3-12. A preliminary result showed a significant increase in Lys content. The more comprehensive analysis of amino acid composition will be reported in future communications.

ACKNOWLEDGMENTS

The authors sincerely thank Drs. T.M. Okita for generously providing pGT1 genomic clone, K. Shimamoto for rice cv OC nurse cells, M.D. Burow for antibody, M.C. Rush for greenhouse space, Q.J. Xie for advice on rice crossing, and all members of the plant molecular biology laboratory for discussion and encouragement.

Received May 19, 1995; accepted August 15, 1995.

LITERATURE CITED

Bagga S, Adams H, Kemp JD, Sengupta-Gopalan C (1995) Accumulation of 15-kilodalton zein in novel protein bodies in transgenic tobacco. Plant Physiol **107:** 13–23

Bollini R, Chrispeels MJ (1978) Characterization and subcellular localization of vicilin and phytohemagglutinin, the two major reserve proteins of *Phaseolus vulgaris* L. Planta **142:** 291–298

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:** 248–254

Busto MM, Kalkan FA, VandenBosch KA, Hall TC (1991) Differential accumulation of four phaseolin glycoforms in transgenic tobacco. Plant Mol Biol **16:** 381–395

Callis J, Fromm M, Walbot V (1987) Introns increase gene expression in cultured maize cells. Genes Dev **1:** 1183–1200

Chrispeels MJ (1991) Sorting of proteins in the secretory system. Annu Rev Plant Physiol Plant Mol Biol **42:** 21–53

Cramer JH, Lea K, Slightom JL (1985) Expression of phaseolin cDNA genes in yeast under control of natural plant DNA sequence. Proc Natl Acad Sci USA **82:** 334–338

D'Amico L, Valsasina B, Daminati MG, Fabbrini MS, Nitti G, Bollini R, Ceriotti A, Vitale A (1992) Bean homologs of the mammalian glucose-regulated proteins: induction by tunicamycin and interaction with newly synthesized seed storage proteins in the endoplasmic reticulum. Plant J **2:** 443–455

Delaney DE, Bliss FA (1991) Selection for increased percentage phaseolin in common bean. 1. Comparison of selection for seed protein alleles and S_1 family recurrent selection. Theor Appl Genet **81:** 301–305

Dyer JM, Nelson JW, Murai N (1993) Strategies for selecting mutation sites for methionine enhancement in the bean seed storage protein phaseolin. J Prot Chem **12:** 545–560

Engvall E, Perlmann P (1972) Enzyme-linked immunosorbent assay, ELISA. III. Quantification of specific antibodies by enzyme-labeled anti-immunoglobulin in antigen-coated tubes. J Immunol **109:** 129–135

Ey PL, Ashman LK (1986) The use of alkaline phosphatase-conjugated anti-immunoglobulin with immunoblots for determining the specificity of monoclonal antibodies to protein mixtures. Methods Enzymol **121:** 497–509

Goodall GJ, Filipowicz W (1991) Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J **10:** 2635–2644

Hayashimoto A, Li Z, Murai N (1990) A polyethylene glycol-mediated protoplast transformation system for production of fertile transgenic rice plants. Plant Physiol **93:** 857–863

Keith B, Chua N-H (1986) Monocot and dicot pre-mRNAs are processed with different efficiencies in transgenic tobacco. EMBO J **5:** 2419–2425

Li X, Franceschi VR, Okita TW (1993a) Segregation of storage protein mRNAs on the rough endoplasmic reticulum membranes of rice endosperm cells. Cell **72:** 869–879

Li X, Wu Y, Zhang D-Z, Gillikin JW, Boston RS, Franceschi VR, Okita TW (1993b) Rice prolamine protein body biogenesis: a BiP-mediated process. Science **262:** 1054–1056

Li Z, Burow MD, Murai N (1990) High frequency generation of fertile transgenic rice plants after PEG-mediated transformation. Plant Mol Biol Rep **8:** 276–291

Loesch PJ Jr, Foley DC, Cox DF (1976) Comparative resistance of *opaque-2* and normal inbred lines of maize to ear-rotting pathogens. Crop Sci **16:** 841–842

McElroy D, Zhang W, Cao J, Wu R (1990) Isolation of an efficient actin promoter for use in rice transformation. Plant Cell **2:** 163–171

Mertz ET, Bates LS, Nelson OE (1964) Mutant gene that changes protein composition and increases lysine content of maize endosperm. Science **145:** 279–280

Murai N, Sutton DW, Murray MG, Slightom JL, Merlo DJ, Reichert NA, Sengupta-Gopalan C, Stock CA, Baker RF, Kemp JD, Hall TC (1983) Phaseolin gene from bean is expressed after transfer to sunflower *via* tumor-inducing plasmid vectors. Science **222:** 476–482

O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem **250:** 4007–4021

Ohtani T, Galili G, Wallace JC, Thompson GA, Larkins BA (1991) Normal and lysine-containing zeins are unstable in transgenic tobacco seeds. Plant Mol Biol **16:** 117–128

Okita TW, Hwang YS, Hnilo J, Kim WT, Aryan AP, Larson R, Krishnan HB (1989) Structure and expression of the rice glutelin multigene family. J Biol Chem **264:** 12573–12581

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol **5:** 69–76

Sen P, Murai N (1991) Oligolabeling DNA probes to high specific activity with sequenase. Plant Mol Biol Rep **9:** 127–130

Sengupta-Gopalan C, Reichert NA, Barker RF, Hall TC, Kemp JD (1985) Developmentally regulated expression of the bean β-phaseolin gene in tobacco seed. Proc Natl Acad Sci USA **82:** 3320–3324

Shimamoto K, Terada R, Izawa T, Fujimoto H (1989) Fertile transgenic rice plants regenerated from transformed protoplasts. Nature **338:** 274–276

Slightom JL, Sun SM, Hall TC (1983) Complete nucleotide sequence of a French bean storage protein gene: phaseolin. Proc Natl Acad Sci USA **80:** 1897–1901

Sturm A, Kuik AV, Vliegenthart JFG, Chrispeels MJ (1987) Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin. J Biol Chem **262:** 13392–13403

Sun SM, Hall TC (1975) Solubility characteristics of globulins from *Phaseolus* seeds in regard to their isolation and characterization. J Agric Food Chem **23:** 184–189

Tanaka A, Mita S, Ohta S, Kyozuka J, Shimamoto K, Nakamura K (1990) Enhancement of foreign gene expression by a dicot intron in rice but not in tobacco is correlated with an increased

level of mRNA and an efficient splicing of the intron. Nucleic Acids Res **18:** 6767–6770

Tarentino AL, Trimble RB, Plummer TH Jr (1989) Enzymatic approaches for studying the structure, synthesis, and processing of glycoproteins. Methods Cell Biol **32:** 111–139

Vasil V, Clancy M, Ferl RJ, Vasil IK, Hannah LC (1989) Increased gene expression by the first intron of maize *Shrunken-1* locus in grass species. Plant Physiol **91:** 1575–1579

Vitale A, Bielli A, Ceriotti A (1995) The binding protein associates with monomeric phaseolin. Plant Physiol **107:** 1411–1418

Voelker TA, Herman EM, Chrispeels MJ (1989) In vitro mutated phytohemagglutinin genes expressed in transgenic tobacco seeds: role of glycans in protein targeting and stability. Plant Cell **1:** 95–104

Yamagata H, Sugimoto T, Tanaka K, Kasai Z (1982) Biosynthesis of storage proteins in developing rice seeds. Plant Physiol **70:** 1094–1100

Zheng Z, Hayashimoto A, Li Z, Murai N (1991) Hygromycin resistance gene cassettes for vector construction and selection of transformed rice protoplasts. Plant Physiol **97:** 832–835

Zheng Z, Kawagoe Y, Xiao S, Li Z, Okita T, Hau TL, Lin A, Murai N (1993) 5′ distal and proximal *cis*-acting regulator elements are required for developmental control of a rice seed storage protein *glutelin* gene. Plant J **4:** 357–366