

Transcriptional and Post-Transcriptional Regulation of Soybean Seed Protein mRNA Levels



Linda Walling; Gary N. Drews; Robert B. Goldberg

Proceedings of the National Academy of Sciences of the United States of America, Vol. 83, No. 7 (Apr. 1, 1986), 2123-2127.

Stable URL:

<http://links.jstor.org/sici?sici=0027-8424%2819860401%2983%3A7%3C2123%3ATAPROS%3E2.0.CO%3B2-D>

Proceedings of the National Academy of Sciences of the United States of America is currently published by National Academy of Sciences.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/nas.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels

(soybean/gene regulation/transcription/seed proteins)

LINDA WALLING*, GARY N. DREWS, AND ROBERT B. GOLDBERG†

Department of Biology, University of California, Los Angeles, CA 90024

Communicated by James Bonner, November 20, 1985

ABSTRACT We investigated soybean seed protein gene transcription during development. We found that seed protein genes are transcriptionally activated and then repressed during embryogenesis and that these genes are either inactive or transcribed at low levels in the mature plant. We further observed that genes encoding mRNAs with vastly different prevalences are transcribed at similar rates. DNA gel blot studies showed that transcriptionally active and inactive seed protein genes have indistinguishable methylation patterns. We conclude that both transcriptional and posttranscriptional processes regulate seed protein mRNA levels in the absence of detectable DNA methylation changes.

Higher plants contain a diverse polypeptide set that accumulates during embryogenesis and is stored in the dormant seed (1, 2). In soybean, the storage proteins β -conglycinin (3) and glycinin (4) constitute 70% of the seed protein mass (5), are hydrolyzed during germination, and serve as a food source for the developing seedling. Soybean seeds also contain less-prevalent proteins such as lectin (6, 7), Kunitz and Bowman-Birk trypsin inhibitors (8, 9), and urease (10). The physiological relevance of low-abundance seed proteins is obscure because soybean lines that either lack or have reduced amounts of these proteins appear normal (11–13). Storage proteins only accumulate during embryogenesis (1, 2). On the other hand, low-prevalence seed proteins, or their relatives, may be present in mature plant organ systems at reduced levels (10, 14, 15).

We showed that approximately 15,000 genes are expressed during soybean embryogenesis and that the majority encode mRNAs present at <20 times per cell (16). Most diverse embryonic mRNA sequences persist throughout embryogenesis, are stored in the dry seed, and are present on leaf polysomes (16). In contrast, seed protein mRNAs undergo quantitative modulations during embryogenesis, constitute >50% of the embryo mRNA mass at a specific developmental period, and are either absent from or present as rare class messages in the leaf cytoplasm (6, 16, 17). Hybridization studies with steady-state embryo and leaf nuclear RNAs (nRNAs) suggested that seed protein genes are regulated at the transcriptional level (17).

In this study we addressed two questions. First, to what extent do transcriptional and posttranscriptional processes act to regulate seed protein mRNA levels? Second, do DNA methylation changes occur when seed protein genes are activated? Our results show that seed protein genes are transcriptionally induced and then repressed during embryogenesis. However, posttranscriptional processes also play a role in regulating seed protein mRNA levels. Our results further show that seed protein genes have similar methylation patterns, irrespective of their transcribed state.

The publication costs of this article were defrayed in part by charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Developmental Staging. Soybean embryos were staged as described (16).

Isolation of mRNA. Polysomal poly(A)⁺ mRNAs were isolated as described except that the EDTA-release step was omitted (16).

In Vitro Nuclear RNA Synthesis. Nuclei were isolated as described by Luthe and Quatrano (18, 19), except that all buffers were adjusted to pH 8.5. Runoff [³²P]nRNA synthesis was carried out for 20 min at 30°C in a buffer containing 0.5–1.0 nM [³²P]UTP, 100 mM (NH₄)₂SO₄, 30 mM Tris-HCl (pH 8.5), 7 mM MgCl₂, 500 μ M ATP, 500 μ M GTP, 500 μ M CTP, 3 μ M phosphocreatine, 0.025 μ g of creatine phosphokinase per ml, and 3 mM 2-mercaptoethanol. These conditions were optimal for extension of nascent RNA chains in all nuclei investigated. [³²P]UTP incorporation was reduced by 40% in the presence of 2 μ g of α -amanitin per ml; incorporation was linear for 20 min, transcripts ranged from 0.1 to 8.0 kilobases (kb) in size, transcription was asymmetric, and the kinetics of [³²P]UTP incorporation into nRNA were similar for each transcribed gene.

[³²P]nRNA Isolation. [³²P]nRNAs were isolated by the procedure of Groudine *et al.* (20) omitting the CCl₃COOH precipitation step. [³²P]nRNAs were pelleted through CsCl (21) and then extracted with cetyltrimethylammonium bromide (22).

Labeling and Isolation of Phage and Plasmid DNAs. Plasmid and phage DNAs were isolated and labeled as described (6, 17, 23, 24).

DNA-Excess Filter Hybridization. Plasmids were bound to nitrocellulose and hybridized with [³²P]nRNAs by published procedures (25, 26).

DNA Gel-Blot Hybridization. DNA fragments were blotted onto nitrocellulose and hybridized with labeled probes as described by Southern (27) and by Wahl *et al.* (28).

RNA Dot-Blot Hybridization. mRNAs were bound to nitrocellulose and then hybridized with a >10-fold mass excess of labeled plasmid DNA as described by Thomas (29). After hybridization, dots were cut out and counted in a scintillation counter.

RESULTS

Seed Protein Genes Are Differentially Regulated. Fig. 1A shows relative seed protein mRNA prevalences at different embryonic stages and in mature plant organ systems. Embryo mRNA prevalences at 70 days after flowering (DAF) are summarized in Table 1. Seed protein mRNA concentrations increased and decreased during embryogenesis and reached maximum values at \approx 70 DAF (Fig. 1A). Peak prevalences

Abbreviations: nRNA, nuclear RNA; kb, kilobases; DAF, days after flowering.

*Present address: Department of Botany and Plant Sciences, University of California, Riverside, CA 92521.

†To whom correspondence should be addressed.

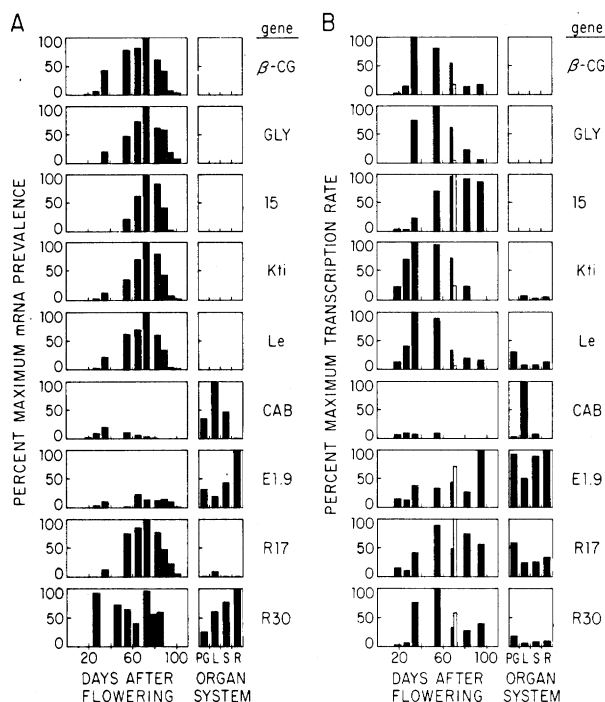


FIG. 1. Accumulation of seed protein mRNAs and relative gene transcription rates during the soybean life cycle. (A) Changes in seed protein mRNA prevalences during development. PG, L, S, and R refer to 14-day postgermination cotyledons, leaves, stems, and roots, respectively; β -CG, Gly, 15, Kti, Le, and CAB refer to β -conglycinin, glycinin, the 15-kDa protein, Kunitz trypsin inhibitor, and chlorophyll a/b plasmids, respectively. Data are plotted as a percentage of the maximum cpm hybridized. (B) Relative seed protein gene transcription rates during development. Data are plotted as a percentage of the maximum relative transcription rates (Tables 2 and 3). ■, Transcription in the entire embryo; □, transcription in the embryonic axis.

ranged from 10% of the mRNA mass (25,000 molecules per cell) for storage protein mRNAs to 0.75% of the mRNA mass (4,000 molecules per cell) for lectin mRNA (Table 1). Each seed protein mRNA class accumulated and then decayed with different kinetics (Fig. 1A). For example, β -conglycinin, glycinin, and 15-kDa protein messages became detectable at different embryonic stages, whereas quantitative changes of β -conglycinin and the 15-kDa protein mRNA levels occurred over 80-day and 40-day periods, respectively.

No detectable signals were observed for any seed protein mRNA class in mature plant organ systems (Fig. 1A). We estimated the sensitivity of our assay to be ≈ 50 molecules per cell. We have shown by other procedures, however, that root polysomes contain ≈ 0.5 molecule per cell of lectin mRNA (J. K. Okamoto and R.B.G., unpublished results) and that all mature plant organ systems contain about 5 molecules per cell of Kunitz trypsin inhibitor mRNA (D. Jofuku and R.B.G., unpublished results). Storage protein messages were not detected in mature plant polysomes at a level of <0.1 molecule per cell (J. J. Harada, R. L. Fischer, and R.B.G., unpublished results).

We also investigated several nonseed protein messages that constitute minor fractions of the embryo mRNA mass. Data summarized in Table 1 show that chlorophyll a/b binding protein mRNA is less prevalent than glycinin mRNA in 70-DAF embryos by a factor of ≈ 500 , whereas R-30 and β -conglycinin messages differ by a factor of 10,000 in concentration. Each nonseed protein mRNA is modulated quantitatively during embryogenesis (Fig. 1A). This is best exemplified by the chlorophyll a/b binding protein and R-17 mRNAs. Unlike seed protein messages, all nonseed protein mRNAs were detected in leaf, stem, and root polysomes.

Seed Protein Genes Are Inactive or Transcribed at Low Levels in Mature Plant Cells. Steady-state leaf mRNA contains <0.2 seed protein transcript per nucleus per family (17). This is lower than that observed in 75-DAF embryo mRNA by a factor of $\approx 10,000$, indicating that seed protein genes are either inactive or weakly transcribed or that primary transcripts rapidly turnover. To distinguish between these possibilities we isolated leaf, root, and stem nuclei and then used them to synthesize [32 P]nRNA *in vitro* (20). To quantitate transcription levels, each [32 P]nRNA was hybridized to filters containing seed protein and nonseed protein plasmid DNAs (Fig. 1B; Table 2) as well as to DNA gel blots containing restriction endonuclease-digested genomic clones (Fig. 2). These assays should measure the relative seed protein transcription rates (30, 31).

As shown in Fig. 1B and summarized in Table 2, glycinin and β -conglycinin transcripts were undetectable in stem and root [32 P]nRNAs; however, very low levels were observed in leaf [32 P]nRNA. Leaf storage protein gene transcription rates were lower than maximum embryonic rates by factors of 50–100 (Tables 2 and 3). Similarly, Kunitz trypsin inhibitor and lectin gene transcripts were lower in leaf, root, and stem [32 P]nRNAs than in embryo [32 P]nRNA by factors of 10–30 (Tables 2 and 3). In contrast, the relative transcription rates for nonseed protein genes were higher in leaf, root, and stem

Table 1. Comparison of embryo mRNA prevalences and relative transcription rates

Gene	% mRNA*			mRNA molecules per cell†		% [32 P]nRNA‡ $\times 10^3$		
	Embryo	Axis	E/A§	Embryo	Axis	Embryo	Axis	E/A§
β -Conglycinin	11	0.6	18	23,000	300	40	12	3
Glycinin	10	0.8	13	25,000	500	20	0.7	30
15-kDa protein	2	0.9	2	13,000	1500	17	17	1
Kunitz trypsin inhibitor	3	0.5	6	18,000	700	9	3	3
Lectin	0.75			4,000		6		
Chlorophyll a/b	0.02			100		1		
E1.9	0.001			4		2		
R-17	0.015			20		24		
R-30	0.001			1		49		

*Taken from the data of Goldberg *et al.* (6, 17) and unpublished experiments in our laboratory. Prevalences represent minimum estimates.

†Calculated according to Goldberg *et al.* (16, 17).

‡Taken from Table 3 for 70-DAF embryos.

§E/A refers to the ratio of embryo to axis mRNA prevalence or transcription rate. Embryo mRNA represents the cotyledon message population (17).

Table 2. Relative transcription rates in leaf, stem, and root

Gene	% [³² P]nRNA* × 10 ³		
	Leaf	Stem	Root
β-Conglycinin	0.8	ND	ND
Glycinin	0.5	ND	ND
15-kDa protein	0.4	0.4	2
Kunitz trypsin inhibitor	1	0.4	6
Lectin	1.5	1	2
Chlorophyll a/b	260	18	2
E1.9	3	5	5
R-17	12	12	17
R-30	7	9	11

Leaf, stem, and root [³²P]nRNAs were hybridized separately with excess filter-bound plasmid DNAs as outlined. Each hybridization reaction contained 3 × 10⁷ cpm of [³²P]nRNA, plasmid DNA filters, and a pBR322 background hybridization control. pBR322 filters generally contained only 5 cpm above counter background, or ≈1.5 × 10⁻⁵% of the input cpm. ND indicates that the filter-bound cpm were not statistically different from the pBR322 filter-bound cpm as determined by a *t* test (*P* = 0.01). We estimated that the lowest relative hybridization rate detectable by our methods was approximately 3 × 10⁻⁴% of the [³²P]nRNA.

$$* \% [^{32}\text{P}] \text{ nRNA} = \frac{(\text{cpm hybridized} - \text{pBR322 cpm})(R)(100)}{(\text{input cpm})(H)}$$

where R is the ratio of mRNA to plasmid insert lengths and H is the hybridization efficiency (0.2) which was estimated by hybridizing 25S [³²P]rRNA with a filter containing excess soybean 25S rDNA.

than those observed for seed protein genes. With the exception of *R17* and *R30*, these rates were equivalent to or higher than maximal embryonic transcriptional rates (Fig. 1B). For example, leaf [³²P]nRNA contained ≈500-fold more chlorophyll a/b binding protein gene transcripts than did glycinin gene transcripts. Furthermore, the leaf chlorophyll a/b binding protein transcription rate was ≈10-fold higher than that observed at a peak period of embryogenesis (Tables 2 and 3).

Hybridization patterns of leaf and embryo [³²P]nRNAs to DNA gel blots containing seed protein and nonseed protein genomic clones are shown in Fig. 2A and summarized in Fig. 2B. Each seed protein gene hybridized intensely with embryo [³²P]nRNA, whereas no detectable signals were observed

with leaf [³²P]nRNA. In contrast, the chlorophyll a/b binding protein genes reacted weakly with embryo [³²P]nRNA and strongly with leaf [³²P]nRNA. Fig. 2 also shows that regions flanking seed protein genes reacted with both embryo and leaf [³²P]nRNAs. Together, these data show that seed protein genes are transcriptionally inactive or weakly transcribed in mature plant organ systems.

Seed Protein Genes Are Transcriptionally Activated and Repressed During Embryogenesis. The low or undetectable seed protein gene transcriptional activities in mature plant cells suggested that seed protein genes are inactivated late in embryogenesis. To test this possibility, we measured the relative seed protein transcription rates in embryos at various developmental stages and in postgermination cotyledons. The results are shown in Fig. 1B and are summarized in Table 3. As was the case for seed protein messages (Fig. 1A), seed protein gene transcription increased significantly during early development and then diminished prior to dormancy. However, the timing and magnitude of these events were different for each gene family. For example, we detected transcription of β-conglycinin genes 15 days prior to that of the glycinin genes. Furthermore, transcription of the 15-kDa protein genes increased 2–4 weeks after that of other seed protein genes and remained relatively high during the final stage of embryogenesis (Fig. 1B). Chlorophyll a/b binding protein and *R-17* genes underwent analogous transcriptional fluctuations, correlating with quantitative changes in their mRNAs (Fig. 1). On the other hand, *E1.9* gene transcription increased late in development. This finding, and the relatively high 15-kDa protein gene transcription rate in 95-DAF embryos (Fig. 1B), suggests that decreased seed protein gene transcription rates are not entirely due to the generalized reduction in total transcription that occurs prior to dormancy (L.W. and R.B.G., unpublished results).

Seed protein genes were either inactive or transcribed at reduced rates in 14-day postgermination cotyledons (Table 3; Fig. 1B). In contrast, all nonseed protein genes were transcriptionally active in postgermination cotyledon cells, even though some (e.g., chlorophyll a/b binding protein) were inactive late in embryogeny. We conclude that seed protein genes are transcriptionally activated and repressed during embryogenesis and that the accumulation and decay of seed protein mRNAs are controlled in part by changes in seed protein gene transcription.

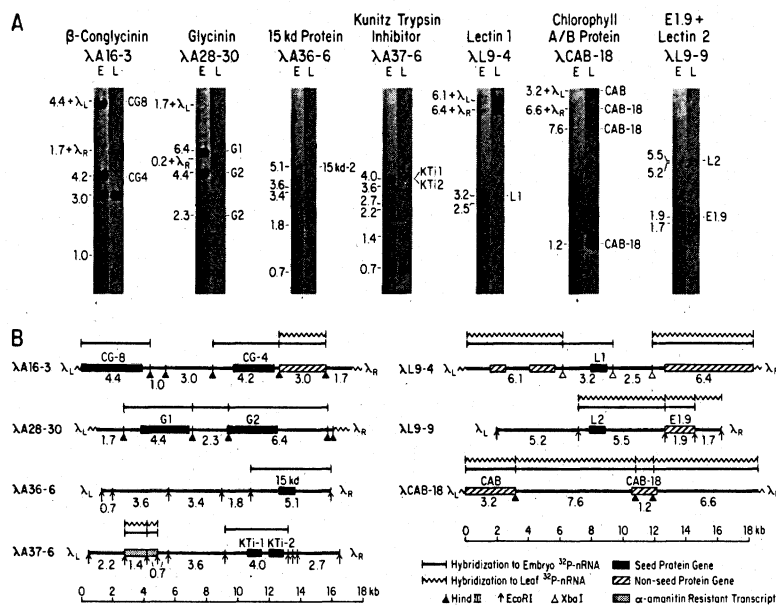


FIG. 2. Hybridization of embryo and leaf [³²P]nRNAs with seed protein genomic clones. (A) Autoradiograms of DNA gel blots. E and L refer to embryo and leaf [³²P]nRNAs, respectively. (B) Transcribed gene regions. CG-8, CG-4, G1, G2, Kti1, Kti2, L1, and L2 refer to different members of each gene family. Numbers represent DNA fragment lengths in kb. Transcripts from the α-amanitin-resistant region are not represented in embryo or leaf mRNAs and contain a sequence repeated >1000 times (D. Jofuku and R.B.G., unpublished data).

Table 3. Relative transcription rates during embryogenesis

Gene	% [³² P]nRNA* × 10 ³								PG [‡]
	DAF								
	20	26	35	55	70E [†]	70A [†]	84	95	
β-Conglycinin	2	10	70	60	40	12	10	12	ND
Glycinin	ND	ND	24	31	20	0.7	8	2	ND
15-kDa protein	0.6	0.5	4	13	17	17	16	15	ND
Kunitz trypsin inhibitor	3	9	12	12	9	3	3	ND	ND
Lectin	3	8	18	17	6	1	4	3	6
Chlorophyll a/b	13	16	20	21	1	1	2	ND	7
E1.9	0.8	0.7	2	2	2	4	1	5	5
R-17	ND	ND	21	43	24	50	36	30	28
R-30	4	9	110	140	49	90	43	60	26

Embryo [³²P]nRNAs were hybridized with filters containing excess plasmid DNAs as described. ND, hybridized cpm not significantly different from pBR322 controls (see Table 2).

*Calculated as outlined in Table 2.

†70E refers to [³²P]nRNA from 70-DAF whole embryos, whereas 70A refers to [³²P]nRNA from 70-DAF embryonic axis.

‡[³²P]nRNA from 14-day postgermination cotyledons.

Posttranscriptional Events Are Important in Regulating Seed Protein mRNA Levels. Several results suggest that factors other than transcription are important in establishing seed protein mRNA levels. First, Fig. 1 shows that, late in development, 15-kDa protein messages decline by a factor of >50 in prevalence while transcription rates are reduced by only 10%. Second, Table 1 shows that seed protein mRNAs are less prevalent in the embryonic axis as compared to the cotyledons but that transcription rates are not reduced proportionately. Finally, Table 1 shows that the relative transcription rates of seed protein and nonseed protein genes in 70-DAF embryos are similar, despite 100 to 10,000-fold differences in mRNA prevalences. We conclude that posttranscriptional processes also contribute to the establishment of seed protein mRNA levels.

Seed Protein Genes Have Similar Methylation Patterns Irrespective of the Transcribed State. We hybridized seed protein and chlorophyll a/b binding protein probes with gel blots containing developmentally distinct DNAs, which were digested with either *Hpa* II or *Msp* I to determine whether transcriptionally active and inactive seed protein genes had different methylation states (32). Most *Hpa* II/*Msp* I sites within and surrounding several seed protein genes were unmethylated (Fig. 3). In addition, each probe produced similar gel blot patterns with the DNAs tested. Because *Hpa* II and *Msp* I only measure a fraction of potentially methylatable cytosines, we hybridized the same probes with DNAs digested with *Ava* I, *Hha* I, *Pvu* I, *Xho* I, and *Eco*RII, which are also methylation-sensitive enzymes (32). Each DNA tested yielded the same hybridization pattern (data not shown). These findings indicate that seed protein gene activation and repression are not correlated with detectable methylation changes.

DISCUSSION

Seed Protein Genes Are Transcriptionally Regulated. We investigated the extent to which transcriptional processes regulate soybean seed protein gene expression. The gene families we studied represent only a small fraction of the genes expressed during soybean embryogenesis and encode highly prevalent messages. For comparison, we also studied several nonseed protein genes that are expressed in embryos and in

organ systems of the mature plant. To approach the issue of transcriptional control, we utilized [³²P]nRNAs synthesized in isolated nuclei by extending preinitiated chains in the presence of [³²P]UTP. In principle, this allowed us to distinguish between transcription and intranuclear turnover events.

We conclude from the data presented here that all seed protein gene families investigated are regulated in part at the transcriptional level. First, seed protein genes are either inactive or weakly transcribed in leaves, roots, stems, and postgermination cotyledons (Tables 2 and 3; Fig. 1B). This correlates well with seed protein mRNA prevalences in mature plant organs (Fig. 1A) and postgermination cotyledons (17) as well as with the representation of seed protein gene transcripts in steady-state leaf nRNA (17). Second, seed protein genes are transcriptionally activated early in embryogenesis (Fig. 1B), and transcription diminishes and/or becomes undetectable prior to seed dormancy (Fig. 1B; Table 3). Finally, there is a general correlation between the timing of seed protein gene transcriptional activity and the quantitative fluctuations in seed protein mRNA prevalences (Fig. 1). We conclude that seed protein genes are only transcribed during periods of the life cycle when mRNAs are produced. That is, they do not represent a class of genes whose transcripts are present constitutively in heterologous nRNAs (33, 34).

By using a similar approach, other plant genes have been shown to be transcriptionally regulated (35–40). The strength of our conclusions, and those of others, is based on the assumption that run-off transcription accurately reflects *in vivo* events. Several results indicate that this assumption is

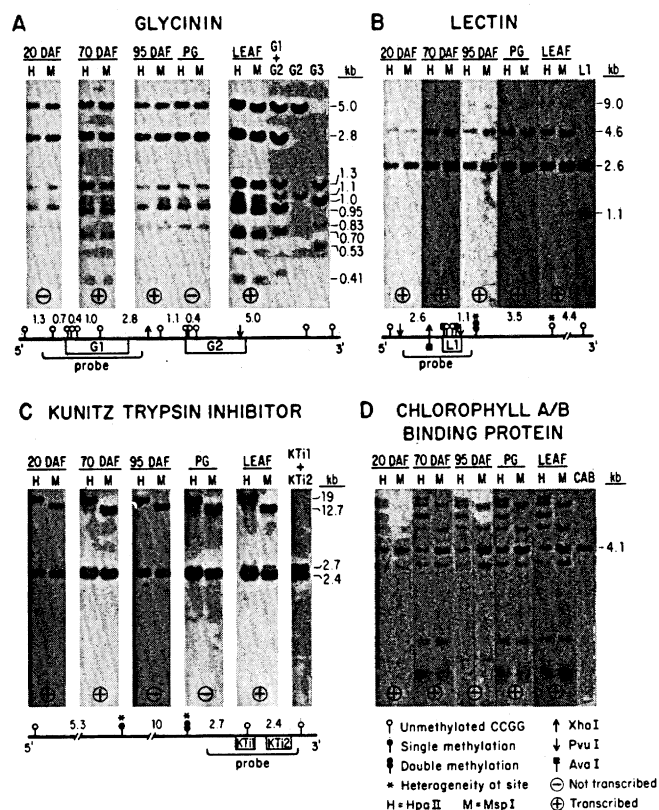


FIG. 3. DNA gel blots of *Hpa* II- and *Msp* I-digested soybean DNAs. DNA was extracted from 20-, 70-, and 95-DAF embryos as well as from 14-day postgermination cotyledons (PG) and leaves. Each DNA was digested with *Hpa* II (H) and *Msp* I (M), subjected to electrophoresis, blotted, and hybridized with the relevant probe. The regions represented by each probe are bracketed. Mapping data for the chlorophyll a/b region are unavailable. Lanes: G1, G2, G3, L1, Kti1, and Kti2 are members of each gene family; G1 + G2, G2, G3, Kti1 + Kti2, L1, and CAB are single-copy DNA reconstructions.

valid. First, pulse-chase experiments showed that labeled transcripts are stable for at least 60 min in isolated nuclei (data not shown). Second, [³²P]nRNAs reacted with specific regions within long genomic clones (Fig. 2) and were complementary to transcribed DNA strands (data not shown). Third, no hybridization was observed when [³²P]nRNAs were synthesized in the presence of low levels of α -amanitin (data not shown). Finally, results obtained with [³²P]nRNA were synthesized by chain extension in isolated animal nuclei agree very closely with those obtained with *in vivo* labeled nRNA (30, 31).

Posttranscriptional Processes Are Important in Regulating Seed Protein Gene Expression. A second conclusion drawn from our data is that seed protein mRNA levels cannot be controlled exclusively by transcriptional events. This is indicated by the findings presented in Table 1, which show that seed protein and nonseed protein embryo mRNA prevalences differ by several orders of magnitude despite similar gene transcription rates. Although our measurements reflect averages of several related genes within each family, recent measurements with gene-specific probes support this conclusion. For example, we demonstrated that β -conglycinin α , α' , and β subunit mRNAs accumulate at different times in embryogenesis; however, each family member is transcriptionally activated and repressed during the same developmental periods (J. J. Harada and R.B.G., unpublished results). We conclude that post-transcriptional events such as cytoplasmic entry rates and/or differential mRNA stabilities (41, 42) are important in regulating seed protein gene expression.

Seed Protein Genes Are not Coordinately Regulated. Each seed protein gene family is regulated independently at both the mRNA and gene levels (Fig. 1). Moreover, each seed protein gene family is differentially represented in mature plant organs (Fig. 1; Table 2). Although seed protein gene families are not coordinately regulated in the formal sense, they all share similar features. That is, they are expressed at high levels during embryogenesis (Table 1), their mRNAs accumulate and decay in a precise developmental timetable (Fig. 1A), and they are inactive or weakly expressed in mature plant organ systems (Fig. 1).

Transcriptionally Active and Inactive Seed Protein Genes Have Similar Methylation Patterns. Previously we showed that selective gene amplification and DNA rearrangements do not play a role in regulating seed protein gene expression (17, 23). Here we demonstrate that transcriptionally active and inactive seed protein genes have similar methylation patterns (Fig. 3) and that seed protein genes are undermethylated in relation to average soybean DNA regions (43). Recently, we showed that soybean lectin (J. K. Okamuro and R.B.G., unpublished results), Kunitz trypsin inhibitor (D. Jofuku and R.B.G., unpublished results), and β -conglycinin (S. Barker and R.B.G., unpublished results) genes retain their developmental-specific expression programs in transformed tobacco cells. Similar observations have been made for *Phaseolus* storage protein genes (44). The tobacco gene transfer system should allow us to identify the DNA sequences and cellular factors required for the regulation of seed protein gene expression.

We dedicate this paper to Professor James Bonner who first suggested that legume storage protein genes are regulated transcriptionally (45). This research was supported by the National Science Foundation and the U.S. Department of Agriculture grants to R.B.G.

1. Larkins, B. A. (1981) in *The Biochemistry of Plants*, eds. Stumpf, P. K. & Conn, E. C. (Academic, New York), Vol. 6, pp. 449-489.
2. Higgins, T. J. V. (1984) *Annu. Rev. Plant Physiol.* **35**, 191-221.
3. Thanh, V. H. & Shibasaki, K. (1978) *J. Agric. Food Chem.* **26**, 692-695.

4. Nielsen, N. C. (1984) *Philos. Trans. R. Soc. London Ser. B* **304**, 287-296.
5. Hill, J. E. & Breidenbach, R. W. (1974) *Plant Physiol.* **53**, 742-746.
6. Goldberg, R. B., Hoschek, G. & Vodkin, L. O. (1983) *Cell* **33**, 465-475.
7. Vodkin, L. O., Rhodes, P. R. & Goldberg, R. B. (1983) *Cell* **34**, 1023-1031.
8. Hildebrand, D. F. & Hymowitz, T. (1980) *Crop Sci.* **20**, 818-819.
9. Hammond, R. A., Foard, D. E. & Larkins, B. A. (1984) *J. Biol. Chem.* **259**, 9883-9890.
10. Polacco, J. C. & Winkler, R. G. (1984) *Plant Physiol.* **74**, 800-803.
11. Pull, S. P., Pueppke, S. G. & Hymowitz, T. (1982) *Science* **220**, 1277-1279.
12. Orf, J. H. & Hymowitz, T. (1979) *J. Am. Oil Chem. Soc.* **56**, 722-726.
13. Polacco, J. C., Thomas, A. L. & Bledsoe, P. J. (1982) *Plant Physiol.* **69**, 1233-1240.
14. Gade, W., Schmidt, E. L. & Wold, F. (1983) *Planta* **158**, 108-110.
15. Brown, W. E., Takio, K., Titani, K. & Ryan, C. A. (1985) *Biochemistry* **24**, 2105-2108.
16. Goldberg, R. B., Hoschek, G., Tam, S. H., Ditta, G. S. & Breidenbach, R. W. (1981) *Dev. Biol.* **83**, 201-217.
17. Goldberg, R. B., Hoschek, G., Ditta, G. S. & Breidenbach, R. W. (1981) *Dev. Biol.* **83**, 218-231.
18. Luthe, D. S. & Quatrano, R. S. (1980) *Plant Physiol.* **65**, 305-308.
19. Luthe, D. S. & Quatrano, R. S. (1980) *Plant Physiol.* **65**, 309-313.
20. Groudine, M., Peretz, M. & Weintraub, H. (1981) *Mol. Cell. Biol.* **1**, 281-288.
21. Goldberg, R. B., Hoschek, G., Kamalay, J. C. & Timberlake, W. E. (1978) *Cell* **14**, 123-131.
22. Langridge, J., Langridge, P. & Bergquist, P. L. (1980) *Anal. Biochem.* **103**, 264-271.
23. Fischer, R. L. & Goldberg, R. B. (1982) *Cell* **29**, 651-660.
24. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237-251.
25. Gillespie, D. & Spiegelman, S. (1965) *J. Mol. Biol.* **12**, 829-842.
26. Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M. & Darnell, J. E. (1981) *Cell* **23**, 731-739.
27. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517.
28. Wahl, G. M., Stern, M. & Stark, G. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3687-3693.
29. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201-5205.
30. Hofer, E. & Darnell, J. E. (1981) *Cell* **23**, 585-593.
31. Hofer, E., Hofer-Warbinek, R. & Darnell, J. E. (1982) *Cell* **29**, 887-893.
32. McClelland, M. (1983) *Nucleic Acids Res.* **11**, 169-173.
33. Kamalay, J. C. & Goldberg, R. B. (1980) *Cell* **19**, 935-946.
34. Kamalay, J. C. & Goldberg, R. B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2801-2805.
35. Evans, I. M., Gatehouse, J. A., Croy, R. D. & Boulter, D. (1984) *Planta* **160**, 559-568.
36. Beach, L. R., Spencer, D., Randell, P. J. & Higgins, T. J. V. (1985) *Nucleic Acids Res.* **13**, 999-1013.
37. Chappell, J. & Hahlbrock, K. (1984) *Nature (London)* **311**, 76-78.
38. Silverthorne, J. & Tobin, E. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1112-1116.
39. Hagen, G., Kleinischmidt, G. & Guilfoyle, T. (1984) *Planta* **162**, 147-153.
40. Eckes, P., Schell, J. & Willmitzer, L. (1985) *Mol. Gen. Genet.* **199**, 216-224.
41. Cabrera, C. V., Lee, J. J., Ellison, J. W., Britten, R. J. & Davidson, E. H. (1984) *J. Mol. Biol.* **174**, 85-111.
42. Tobin, A. J. (1979) *Dev. Biol.* **68**, 47-58.
43. Adams, R. L. P. & Burdon, R. H. (1982) *CRC Crit. Rev. Biochem.* **13**, 349-384.
44. Sengupta-Gopalan, C., Reichert, N., Barker, R. F., Hall, T. C. & Kemp, J. D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3320-3324.
45. Bonner, J., Huang, R. C. & Gilden, R. V. (1963) *Proc. Natl. Acad. Sci. USA* **50**, 893-900.