Plant Molecular Biology **14**: 909–925, 1990.

Comparison of the expression of two highly homologous members of the soybean ribulose-1,5-bisphosphate carboxylase small subunit gene family

B.W. Shirley,[1,3] D.P. Ham,[1,4] J.F. Senecoff,[1] S.L. Berry-Lowe,[1,5] L.L. Zurfluh,[2] D.M. Shah[2] and R.B. Meagher[1]*

[1]*Department of Genetics, University of Georgia, Athens, GA 30602, USA (*author for correspondence)*; [2]*Division of Biological Sciences, Monsanto Company, 700 Chesterfield Parkway, St. Louis, MO 63198, USA*; [3]*Current address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA*; [4]*Current address: Department of Poultry Science, University of Georgia, Athens, GA 30602, USA*; [5]*Current address: Department of Botany, University of Georgia, Athens, GA 30602, USA*

Received 22 June 1989; accepted in revised form 29 January 1990

Key words: cDNA library, light regulation, polyadenylation, post-transcriptional processing, ribulose-1,5-bisphosphate carboxylase small subunit, RNase H, S1 nuclease, soybean

Abstract

Two soybean ribulose-1,5-bisphosphate carboxylase small subunit (SSU) genes, SRS1 and SRS4, are highly homologous over a region that includes 4 kb of 5′ and 1 kb of 3′ flanking sequences. The expression of these genes was compared using synthetic oligonucleotide probes. Analysis of a soybean leaf cDNA library indicates that SRS1 and SRS4 are the most highly expressed members of the soybean SSU gene family. Similar changes were observed in the RNA levels for these genes in response to white light, far-red light and darkness, although SRS1 was expressed at a four-fold higher level in total RNA than SRS4 under all conditions. However, nuclear run-on assays indicate that SRS1 is transcribed at a lower rate than SRS4, which suggests that SRS1 RNA is more stable. S1 nuclease analysis and oligonucleotide directed RNase H cleavage indicate that transcripts from both genes are polyadenylated within two principle regions separated by 35 nt. Sequence analysis of 16 independent cDNA clones identified seven different polyadenylation sites, and six of these sites lie within these two regions. Although SRS1 RNA was poorly recovered during poly(A)$^+$ fractionation, RNase H cleavage experiments showed that transcripts from SRS1 and SRS4 had similar poly (A) tail lengths ranging from 0 to 220 nt. In addition, and despite differences in the untranslated leader sequences, SRS1 and SRS4 RNAs are assembled into polysomes with equal efficiencies. The overall similarity in expression patterns for these two genes further illustrates the coordinate evolution of individual members of a SSU gene family and is consistent with the proposal that gene conversion homogenizes both the coding and regulatory regions of these genes.

Introduction

The small subunit (SSU) of the chloroplast enzyme, ribulose-1,5-bisphosphate carboxylase (Rubisco), which catalyzes competing reactions in photosynthetic carbon fixation and photorespiration, has served as a model for studies on the regulation of gene expression. SSU poly-

peptides are encoded by small nuclear gene families in higher plants and green algae. Expression of these genes is regulated by light and involves both transcriptional and post-transcriptional processes [62, 65]. Although the members of SSU gene families are highly homologous in nucleotide and amino acid sequence, differences have been found in the organ and tissue-specificity or levels of expression of individual SSU genes in a number of plants including maize [55], pea [27, 59], petunia [19, 63] and tomato [60].

Two SSU genes have been characterized from soybean [5, 29]. These genes share extensive homology at the nucleotide level, ranging from 95% in the coding regions to approximately 80% in introns and proximal flanking regions. The region of homology covers 6 kb and is too large to be accounted for by conservation of regulatory sequences. Beyond this region there is no detectable sequence homology. Recent quantitative molecular evolution studies on SSU gene families from a variety of plant species demonstrate that SSU gene family members in all species examined interact frequently by gene conversion, homogenizing SSU coding sequences [45, 48]. Thus the homology between the two soybean sequences may be the result of recent gene conversion events covering the 6 kb region. In this case any differences which might evolve in the regulation of expression of these SSU gene family members would also be homogenized in this process. The advantage to the organism is that multiple SSU genes interacting frequently by gene conversion are able to evolve novel sequence combinations more rapidly than could a single SSU gene [23]. In addition, this results in all the gene family members encoding a homogeneous mature SSU polypeptide [45].

In this study the expression of the two soybean SSU genes was compared using gene-specific oligonucleotide probes. Similar changes in RNA levels were observed for both genes in response to white light, far-red light and darkness. The complex polyadenylation patterns, poly(A) tail lengths and efficiency of loading into polysomes are similar for the two transcripts. However, discrepancies between the transcription rates and steady state RNA levels of these genes suggest that transcripts from the SRS1 gene are more stable.

Materials and methods

Plants and RNA isolation

Soybean plants (*Glycine max* cv. Wayne or Pella '86) were grown in the greenhouse. Light treatments were as described in Shirley and Meagher [57]. Total RNA was prepared from leaves at various stages according to previously published methods [33, 58, 72]. Poly(A)$^+$ RNA was prepared by oligo(dT) cellulose chromatography [72]. Polysomes were isolated from soybean leaves and RNA was purified from the sucrose gradient fractions as described in Shirley and Meagher [57]. All plant material was harvested at 12.00 noon.

Gene-specific oligonucleotides

Oligonucleotide probes for SRS1 and SRS4 were synthesized on an Applied Biosystems DNA synthesizer and purified on polyacrylamide gels or by HPLC. The 18-mer oligonucleotides were from homologous sites in the 3′ untranslated region of the mRNA (nt 789–807, Fig. 5). Antisense probes, complementary to the sequence of the mRNA, were used to screen the cDNA library and to measure steady state RNA levels in total and poly(A)$^+$ RNA samples. Sense probes having the same sequence as the mRNA were used to sequence the cDNA clones by primer extension through the polyadenylation sites.

The gene specificity of the SRS1 and SRS4 antisense oligonucleotide probes was determined according to various criteria. Both probes hybridized to one band of the correct size for SSU on northern blots of soybean total RNA. The SRS1 oligonucleotide hybridized to bands predicted for SRS1 on Southern blots of soybean genomic DNA digested with *Eco* RI or *Hin*d III. The oligo-

nucleotide was also specific for SRS1 when used as a primer for RNA sequencing (J.F. Senecoff and R.B. Meagher, unpublished data). On the other hand, the SRS4 oligonucleotide hybridized to bands of the correct predicted size for SRS4 in *Eco* RI and *Hin*d III digests of genomic DNA, but also to an additional SSU fragment on Southern blots under non-stringent conditions. Thus these two genes may comprise a subfamily separate from SRS1. However, the SRS4 oligonucleotide was specific for SRS4 in primer extension sequencing of total RNA from soybean leaves (J.F. Senecoff and R.B. Meagher, unpublished data), detected only full length products when a SRS4 DNA fragment was used in S1 protection analysis of total and poly(A)$^+$ RNA (described in Results), and detected only SRS4 clones (based on sequence data from the 3′ end) in a leaf cDNA library (see Results). Thus it appears that the second gene in the SRS4 gene subfamily is expressed at low levels relative to SRS1 and SRS4, if at all, in soybean leaves. It was therefore assumed that hybridizations with the SRS4 oligonucleotide probe reflected expression of only the SRS4 gene in the experiments described in this study.

Soybean leaf cDNA library

Double-stranded cDNA was synthesized from the poly(A)$^+$ RNA of 40-day-old plants and was cloned into the λgt10 vector essentially as described by Huynh *et al.* [39]. Approximately 2×10^7 pfu were obtained per μg of cDNA (4.5×10^6 pfu per μg of poly(A)$^+$ RNA). Phage containing the SSU cDNA sequences were selected through three rounds of plaque hybridization using a nick-translated insert from a pea Rubisco cDNA clone, pSS15 [14]. Plaques were transferred to nitrocellulose filters and the DNA was denatured and fixed to the filters. The filters were baked at 80 °C for 2 h and were prehybridized for 4–5 h at 37 °C in 35% formamide, 5× SSC, 5× Denhardt's solution [20], 0.1% SDS, 100 μg/ml yeast tRNA and 10 μg/ml poly(A) RNA. The filters were hybridized in fresh solution overnight at 37 °C using $0.5–1.0 \times 10^6$ cpm of nick-translated probe per filter. After hybridization, the filters were washed once for 30 min at room temperature and three times for 30 min at 50 °C. All washes were carried out in 3× SSC and 0.1% SDS. The conditions used in plaque hybridizations for estimating the occurrence of specific sequences in the cDNA library and in dot blot hybridizations for probing the λgt10 SSU clones with the oligonucleotide probes were as follows: filters were prehybridized for 4–5 h at 48 °C in 6× SSC, 10× Denhardt's solution, 100 μg/ml yeast tRNA, 0.5% SDS. The filters were hybridized in fresh solution overnight at 48 °C. Usually $1–2 \times 10^6$ cpm of the ^{32}P-labelled oligonucleotide probes were used per ml of hybridization solution. The filters were washed twice in 6× SSC, 0.5% NP40 at room temperature for 20–30 min, and finally once in 6× SSC for 5 min at 54 °C. DNA was isolated from individual plaques by standard procedures [43].

Additional DNA was subsequently prepared for sequence analysis. Small amounts of phage DNA were used to transform *E. coli* directly by the method of Hanahan [35]. The transformed cells were plated as for a λ library according to Maniatis *et al.* [43] and the resulting plaques were transferred into 1 ml SM buffer and vortexed with 50 μl chloroform and stored at 4 °C. For each clone, approximately 10–15 μg of λDNA was prepared by the method of Grossberger [31] with the modification that the bacterial culture was infected with 100 μl of the plaque suspension rather than with an entire plaque.

In order to confirm that the SRS1 and SRS4 clones were derived by priming of independent SSU RNA molecules and did not arise from amplification of a single cDNA during cloning, the *Eco* RI digestion patterns of several SRS1 and SRS4 cDNA clones were compared on non-denaturing polyacrylamide gels. One *Eco* RI site was present in the same location within the second exon of the genes. Thus differences in sizes of the fragments reflected differences in the location of the polyadenylation site and the site within the poly(A) tail which was primed by

poly(dT) during cDNA synthesis. All ten of the SRS4 clones and at least five of six SRS1 clones examined contained 3′ *Eco* RI fragments of distinctly different sizes (data not shown), and must therefore have been derived from individual mRNA molecules.

DNA, RNA dot blots and transcription assays

Dot blots of the λgt10 SSC cDNA clones were prepared on nitrocellulose. Dot blots of SSU DNA and soybean RNA were prepared on Biotrans filters as described in Shirley and Meagher [57]. Transcription assays were performed as initially described for soybean seedlings [6]. The technical guidelines for quantifying both the transcription rates and steady-state assays of SSU RNA in total and poly(A)$^+$ RNA samples on dot blots have recently been outlined [57]. Most experiments were repeated several times and none fewer than twice with completely independent RNA preparations.

SI nuclease protection assays

The 0.8 kb *Eco* RI fragment from SRS1 [5] and the 1.5 kb *Hin*d III fragment from SRS4 [29], contain sequences from the 3′ halves of the genes. These were hybridized to total or poly(A)$^+$ RNA at 50 °C and treated with S1 nuclease as described in Condit and Meagher [16]. The products were separated on 8% polyacrylamide denaturing gels and transferred electrophoretically [26] to Biotrans filters (I.C.N.). The gene-specific antisense oligonucleotides (see above) were used as probes for the protected RNA fragments.

Oligonucleotide directed RNase H cleavage assays

RNase H cleavage assays were performed as described in Carrazana *et al.* [15]. A synthetic oligonucleotide, 1043N (5′TGGGGGCTTGTAGGCGATGAA3′), was used to direct the sequence specific RNase H cleavage of the RNA. This oligonucleotide is complementary to a unique sequence 21 nt long, positioned 3–23 nt upstream from the translational stop codon within the 3rd exon of both the SRS1 and SRS4 transcripts [29]. A 10 μg sample of total RNA from 8-day-old light grown soybean seedlings was incubated with 50 pmol of the 1043N oligonucleotide alone or with both 1043N and 500 pmols oligo-dT (12–18). Following RNase H cleavage, the products were separated by electrophoresis in 6% polyacrylamide-8.3 M urea gels and electrophoretically transferred to a nylon membrane. The membrane was prehybridized in 2 × SSC, 0.2% SDS, 0.25% dry milk for at least one hour followed by overnight hybridization with 30×10^6 cpm end-labelled oligonucleotide probe. The gene specific oligonucleotide probes described above were used to differentiate the cleavage products for SRS1 and SRS4. For SRS4, gene specificity was achieved by performing hybridization at 56 °C in Blotto, washing twice in 2 × SSC, 0.2% SDS for 10 min at 23 °C and once in 2 × SSC, 0.2% SDS for 1 min at 56 °C. Gene specificity was achieved for the SRS1 oligonucleotide by performing both the hybridization and final wash at 62 °C.

Subcloning and sequencing

The cDNA inserts were subcloned into the plasmid vector pTZ18U (Pharmacia). Double-stranded DNA from plasmid minipreps was used directly for sequencing by the method of Senecoff *et al.* [54]. Gene-specific sense oligonucleotides which anneal to the nonsense strand of the SRS1 and SRS4 genes were used as primers in the sequencing reactions, thus generating sequence data for the 3′ ends of the cDNA clones.

Results

SRS1 and SRS4 RNA levels respond similarly to light and darkness

Previous studies have shown that the steady-state level of SSU RNA in soybean seedlings is

approximately eight-fold higher in light grown leaves than in etiolated leaves relative to other RNAs [29, 57]. In addition, SSU RNA levels change rapidly in response to white light, far-red light and darkness. However, it was not known whether individual members of the soybean SSU gene family exhibited similar levels and patterns of expression. Therefore, the expression of the two closely related soybean SSU genes, SRS1 and SRS4, was compared using gene-specific oligonucleotide probes. RNA was extracted from leaves of etiolated seedlings shifted into the light for different amounts of time. RNA was also isolated from leaves of light-grown seedlings placed in darkness or treated with 15 min of far-red light and shifted to darkness. The levels of SRS1 and SRS4 RNA were measured on dot blots containing two fold dilution series of each RNA sample. DNA standards were used to normalize between filters hybridized with different oligonucleotide probes and to confirm that gene-specific hybridization conditions were used in each experiment. Hybridization of duplicate filters with a probe for soybean 18S rRNA [29] was used to confirm that equivalent amounts of total RNA were present in each sample.

The results of this experiment are shown in Fig. 1. The levels of SRS1 and SRS4 RNA showed similar changes in response to light and followed the patterns previously described for the soybean SSU gene family as a whole [57]. SRS1 and SRS4 RNA levels increased slowly in response to light to levels 4–8-fold higher than in etiolated leaves. In leaves of light-grown plants the levels of both RNAs decreased 4–8-fold in response to darkness. Although far-red light has a marked effect on SSU transcription, SRS1 and SRS4 RNA levels were not affected by this treatment relative to treatment with darkness alone (data not shown), consistent with previous observations for the SSU gene family [56]. However, a difference was found in the relative levels of SRS1 and SRS4 expression. Normalization of the hybridization signal to the DNA standards showed that at least four-fold more SRS1 than SRS4 RNA was present in total RNA under all conditions examined.

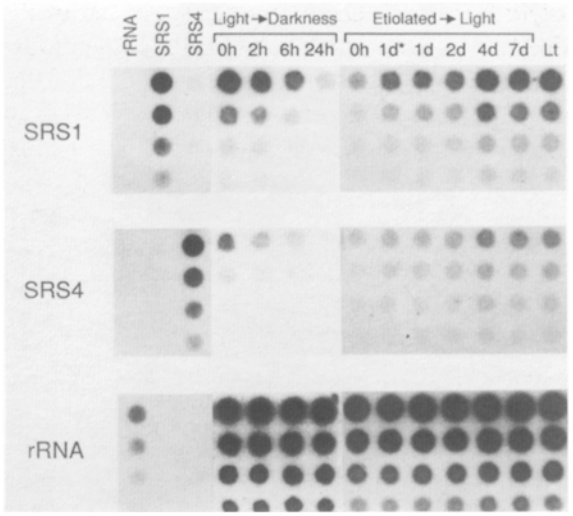

Fig. 1. Comparison of light regulation of SRS1 and SRS4 steady state RNA levels. DNA and RNA samples were applied to nylon filters in two-fold dilution series. Filters were hybridized with oligonucleotide probes specific for SRS1, SRS4 and 18S rRNA and washed as described in Materials and methods. The first four samples on each blot contain plasmid DNA encoding: rRNA (pSR1.1, 0.01 pmol in first dot), SRS1 (pSRS0.8, 0.033 pmol in first dot) and SRS4 (pSRS4-1.5, 0.033 pmol in the first dot). The remaining samples contain RNA (10 μg in the first dot) from leaves of light-grown seedlings treated with darkness or etiolated seedlings treated with white light, as indicated. The asterisk indicates RNA obtained from plants placed in continuous light rather than a 12 h light/dark cycle.

Analysis of a soybean leaf cDNA library

The expression of SRS1, SRS4 and other potential members of the soybean SSU gene family was examined by analysis of a cDNA library. The library was constructed by oligo(dT) priming of poly(A)$^+$ RNA from soybean leaves. The gene-specific probes were used to compare the numbers of SRS1 and SRS4 clones in the library to the number of clones hybridizing to a general SSU probe. The library was plated at a density of approximately 2000 pfu per plate and a replica filter imprints of the phage were made from each plate. One filter from each pair was hybridized to a nick-translated heterologous pea SSU cDNA insert [14] at low stringency to estimate the total number of SSU sequences in the library (Fig. 2A,

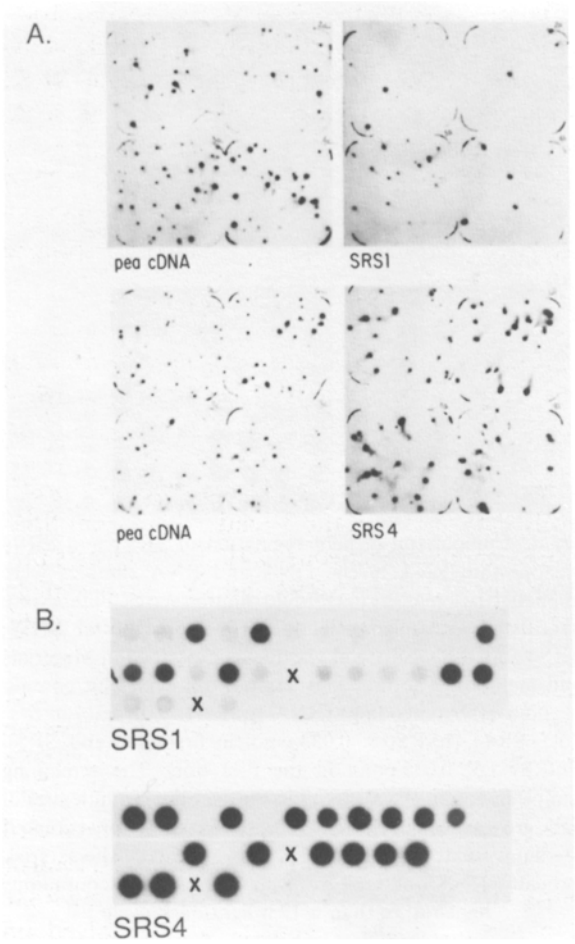

Fig. 2. Analysis of a soybean leaf cDNA library. A. Autoradiograms of nitrocellulose filter imprints of DNA from a soybean cDNA library made from mature leaf poly(A)$^+$ RNA probed with a general SSU pea cDNA clone at low stringency (left) and a replicate filter probed with the oligonucleotide specific for SRS1 (right, top pair) or SRS4 (right, bottom pair) at high stringency (as described in Materials and methods). B. Twenty-eight cDNA clones were selected based on hybridization to a pea SSU clone, spotted onto nitrocellulose and hybridized with the SRS1 or SRS4 gene-specific oligonucleotide. Clones not hybridizing to either probe are marked with an X. Filters were hybridized and washed as described in Materials and methods.

left). The consistently large number of synonymous nucleotide substitutions between the pea and soybean SSU genes suggests that a pea SSU probe should hybridize equivalently to all soybean SSU genes [45]. This probe was shown to hybridize to six to ten genes on Southern blots of soybean genomic DNA and appears to be no more specific for the SRS1 and SRS4 genes than for any other soybean SSU genes [5]. Based on hybridization with this probe it was estimated that approximately 1.5% of the phage in the library contained SSU sequences. Replicate filters were hybridized to the gene-specific oligonucleotides to determine the frequency of SRS1 and SRS4 cDNA clones in the library (Fig. 2A, right). Of the sequences hybridizing to the pea probe, approximately 20% contained the SRS1 sequence whereas approximately 55% contained the SRS4 sequence.

In order to verify these results, DNA was purified from 28 SSU cDNA clones selected with the pea probe and used to prepare DNA dot blots. Duplicate blots were hybridized to the SRS1 and SRS4 oligonucleotide probes. Figure 2B shows that eight of the 28 clones (28%) contained the SRS1 3′ mRNA sequence while 18 of the 28 clones (64%) contained the SRS4 3′ mRNA sequence. Only two of the 28 SSU cDNA clones did not contain the SRS1 or SRS4 3′ sequence.

The results of this analysis suggest that SRS1 and SRS4 encode the majority of the SSU mRNA in soybean leaves. This is based on the assumption that the cDNA sequences in the library were representative of the poly(A)$^+$ mRNA in leaf cells. However, analysis of the cDNA library suggested that SRS4 was expressed at a higher level than SRS1, even though SRS1 RNA levels were significantly higher in total RNA from soybean leaves (Fig. 1). Thus, although SRS1 and SRS4 appear to be the most highly expressed members of the SSU gene family in soybean leaves, it is possible that other SSU transcripts were underrepresented in the cDNA library, as well.

SRS1 RNA is under-represented in poly(A)$^+$ RNA relative to SRS4 RNA

To further examine the discrepancy between cDNA frequency and total RNA levels for SRS1 and SRS4, dot blots were used to compare the levels of SRS1 and SRS4 RNA in total and poly(A)$^+$ RNA preparations. DNA standards were again used to normalize hybridization sig-

nals between filters and to confirm the gene-specificity of the hybridization conditions. Figure 3 and other independent experiments show that levels of SRS1 RNA were approximately two- to four-fold higher than SRS4 RNA in poly(A)$^+$ RNA, but were four to eight fold higher in total RNA from 10-day-old soybean seedlings. Identical results were obtained with RNA from 60-day-old plants (data not shown). This result indicates that SRS1 RNA is poorly recovered relative to SRS4 RNA during poly(A)$^+$ RNA fractionation.

The underrepresentation of SRS1 in poly(A)$^+$ preparations only partially explains the low number of SRS1 clones recovered in the cDNA library. Even though the levels of SRS1 were approximately 2× that of SRS4 RNA in poly(A)$^+$ RNA samples, twice as many SRS4 clones as SRS1 clones were identified in the cDNA library. Thus SRS1 RNA was further underrepresented in the oligo-(dT)-primed cDNA library when compared to the recovery of poly(A)$^+$ RNA.

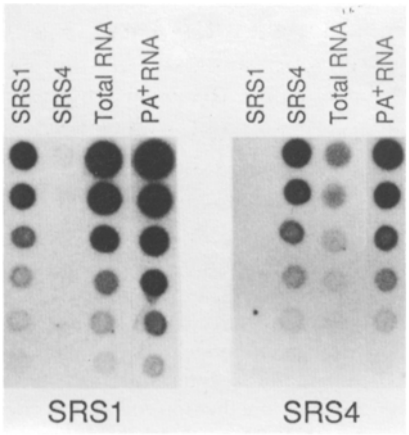

Fig. 3. Comparison of SRS1 and SRS4 RNA levels in total and poly(A)$^+$ RNA. DNA and RNA samples were applied to nylon filters in two fold dilution series. The first two lanes of each filter contained plasmid DNA encoding 3′ sequences for SRS1 (pSRS0.8) or SRS4 (pSRS4-1.5), 0.0025 pmol in the first dot of each sample. The last two lanes of each filter contained total RNA (10 μg in the first dot) and poly(A)$^+$ RNA (PA$^+$, 1 μg in the first dot) from 10-day-old light-grown soybean leaves. The filters were hybridized with the gene-specific oligonucleotide probes for SRS1 and SRS4 and washed as described in Grandbastien *et al.* [29].

These experiments suggest that SRS1 transcripts hybridize less efficiently to oligo (dT) than do SRS4 RNAs. One possibility is that SRS1 transcripts have shorter poly(A) tails than do SRS4 transcripts. Alternatively, differences in the secondary structure of SRS1 and SRS4 RNA could affect the recovery of transcripts from both oligo(dT) chromatography and oligo(dT)-primed cDNA synthesis.

Gene-specific S1 protection reveals two principle regions of 3′ processing for both SRS1 and SRS4 RNAs

The results of S1 nuclease analysis have previously indicated that the 3′ ends of soybean SSU RNAs are heterogeneous [29]. This suggests that transcripts from the SRS1 and SRS4 genes might exhibit differences in 3′ processing and polyadenylation. Preliminary experiments showed that in standard protection assays neither S1 nor mung bean nucleases differentiated between homologous and heterologous SRS1 and SRS4 duplexes. Therefore, gene-specific S1 nuclease protection assays were performed. Unlabeled DNA fragments (Fig. 4A) hybridized with total or poly(A)$^+$ RNA were treated with S1 nuclease and the protected fragments were resolved on sequencing gels and transferred to nylon filters. The filters were hybridized with the antisense oligonucleotide probes to identify protected RNA fragments while at the same time discriminating between SRS1 and SRS4. Based on DNA standards, the sizes of the RNA protection products cluster in two groups of approximately 360 nt and 390 nt in both total and poly(A)$^+$ RNA preparations (Fig. 4B). These experiments indicate that SRS1 and SRS4 transcripts have multiple but related 3′ processing sites clustered about 160 nt and 190 nt from the termination codon. For each gene the distribution of RNA protection products was approximately the same in total RNA and in poly(A)$^+$ RNA, although there were fewer of the long S1 products in total RNA probed with the SRS1-specific oligonucleotide. Similar results were obtained using mung bean nuclease (data not shown).

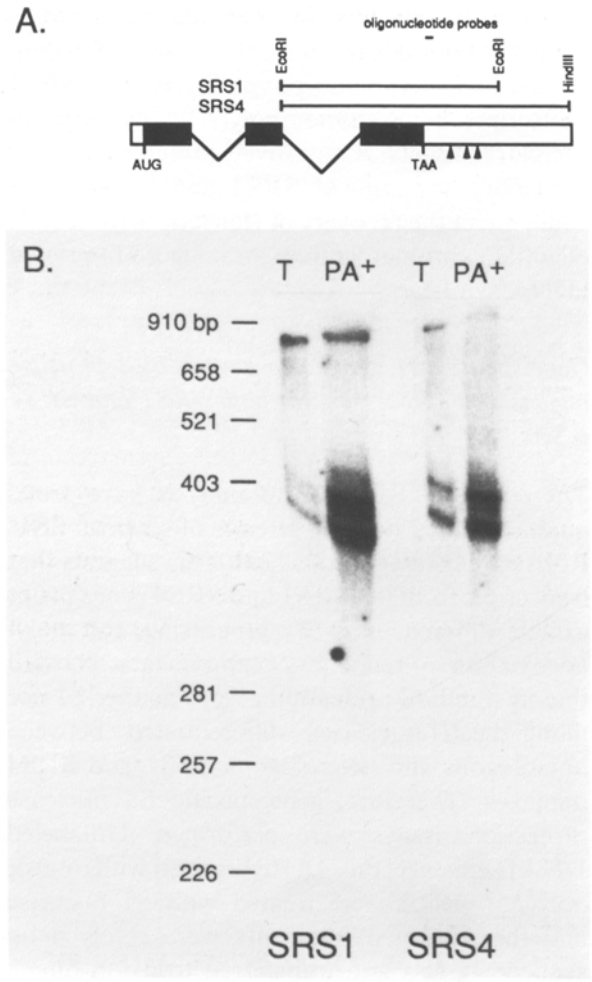

Fig. 4. S1 analysis of SRS1 and SRS4 in total and poly(A)$^+$ RNA. A. The DNA fragments used in S1 nuclease protection are indicated above the structure of the SRS1 and SRS4 genes. Boxes indicate the positions of the three exons. B. DNA fragments containing the 3′ half of the SRS1 or SRS4 genes were hybridized with either 25 μg of total RNA (T) or 2 μg of poly(A)$^+$ RNA (PA$^+$) from leaves of light grown soybean seedlings. Hybrids were treated with S1 nuclease as described in Grandbastien *et al.* [29]. Blots of S1 products were hybridized with the SRS1 or SRS4 oligonucleotide probes.

Oligonucleotide directed RNase H cleavage identifies two major regions of polyadenylation and similar poly(A) tail lengths for both transcripts

In order to confirm the S1 nuclease data and compare the lengths for of the poly(A) tails of SRS1 and SRS4 transcripts, oligonucleotide directed RNase H cleavage products of the transcripts were analyzed. The strategy is outlined in Fig. 5A. A hybrid is formed between both SRS1 and SRS4 transcripts in total seedling RNA and an oligonucleotide complementary to 21 nucleotides just prior to the stop codon in the third exon. These molecules are treated with RNAse H, resolved on acrylamide gels and blotted to nylon filters. The gene specific oligonucleotide probes for SRS1 (Fig. 5B) and SRS4 (Fig. 5C) each detected a broad range of cleavage products due to different poly(A) tail lengths. These range from approximately 190 to 380 nt (lane c), which represents primarily the 3′ untranslated sequences of

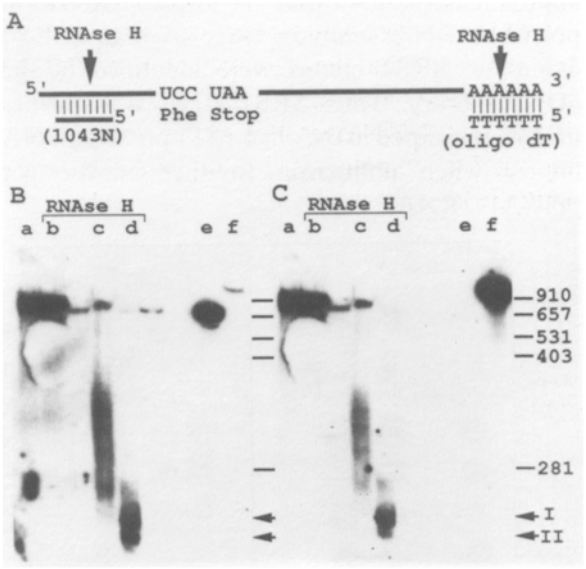

Fig. 5. Oligonucleotide directed RNase H cleavage. A. Strategy. The oligonucleotide, 1043N, complementary to both the SRS1 and SRS4 transcripts at a position nine nucleotides upstream from the stop codon, and oligo(dT) are annealed to total RNA. RNA/DNA hybrids are cleaved with RNase H and blotted to nylon filter. B. RNAs detected on autoradiograph of filter imprint probes with the SRS1 gene-specific oligonucleotide. C. RNAs detected on autoradiograph of filter imprint probed with the SRS4 gene-specific oligonucleotide. The lanes in 5B and 5C contain: a. 10 μg total seedling leaf RNA; b. RNA + RNase H; C; RNA + oligo(1043N) + RNase H; d. RNA + oligo(1043N) + oligo-(dT) + RNase H; e. 0.1 μg of the 0.8 kb *Eco* RI plasmid DNA insert from SRS1; f. 0.1 μg of the 1.4 kb SRS4 *Hin*d III plasmid DNA insert. Details are described in Materials and methods.

the transcripts and variable poly(A) tail lengths. When oligo(dT) is included in an identical reaction in order that the poly(A) tails are also removed by RNase H, two clusters of products at approximately 160 nt and 190 nt in length are detected for both SRS1 and SRS4 (Fig. 5B,C, lane d). These products define the sites at which poly(A) addition occurs. A single filter was stripped and reprobed such that the results for the two oligonucleotides could be compared directly. These data suggest that the majority of poly(A) tails from both transcripts vary from approximately 0 nt to 220 nt in length. This result also demonstrates that there are two major regions of polyadenylation for both transcripts approximately 160 nt and 190 nt from the termination codon consistent with the results of S1 nuclease protection assays. Appropriate controls included demonstrating that the oligonucleotide probes are gene specific (Fig. 5, lanes e,f) and that RNase H directed cleavage does not occur when oligonucleotides complementary to the RNA are omitted (lane b).

Seven polyadenylation sites were found in 16 independent cDNA clones

In order to define more precisely the polyadenylation sites in these genes, the 3′ ends of a number of SRS1 and SRS4 cDNA clones were sequenced. The results are illustrated in Fig. 6. Seven different polyadenylation sites were found in ten independent SRS4 cDNA clones and these were concentrated in two regions. Five of the polyadenylation sites were in a region between 156 and 168 nt downstream of the translation stop codon (Fig. 6, region I); another four were in a region between 191 and 205 nt downstream of the stop codon (Fig. 6, region II). Only one clone contained polyadenylation site outside of these two regions, 106 nt downstream of the stop codon. The two clusters of polyadenylation sites in the cDNA clones correspond to the regions of RNA processing and polyadenylation detected in the S1 nuclease protection and oligonucleotide-directed RNase H cleavage experiments.

However, the six independent SRS1 clones

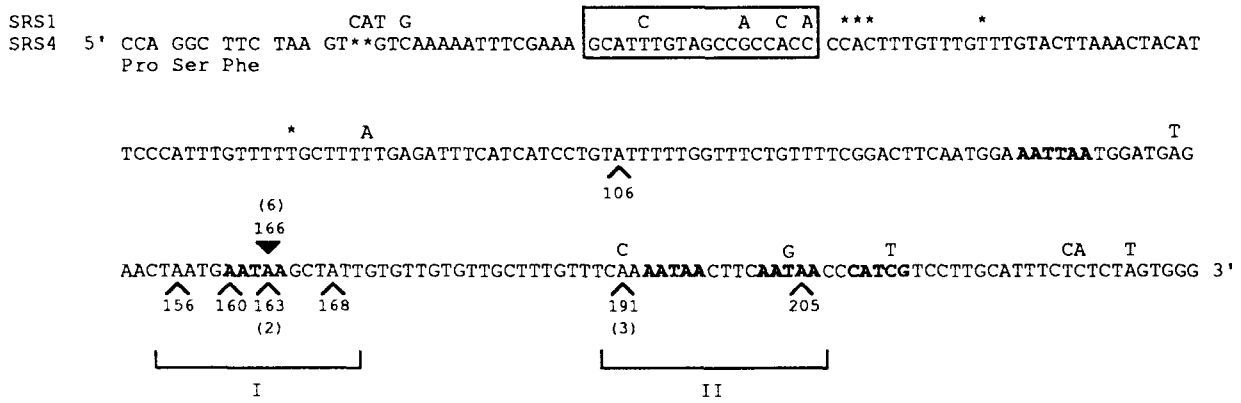

Fig. 6. Sequence of SRS1 and SRS4 showing polyadenylation sites. The sequence of SRS4 is shown, with the nucleotide substitutions for SRS1 indicated above. Deletions of nucleotides in either sequence with respect to the other are shown by an asterisk (*). The last three codons and the stop codon of exon 3 are shown in groups of three nucleotides, with the predicted amino acid sequence given below. The SRS4 sequence is numbered beginning with the first nucleotide after the stop codon (deletions and insertions of nucleotides, make the numbering for SRS1 slightly different). The boxed nucleotides indicate the gene specific primers used for sequencing cDNA clones. The sequences in bold face type are homologous to known mammalian polyadenylation signals. The arrows indicate the polyadenylation sites found for SRS1 (▼) and SRS4 (△). The positions of these sites are numbered relative to the stop codon. Numbers in parentheses indicate the number of cDNAs sequenced that contained poly(A) tracts starting at that position. The two regions of polyadenylation found in SRS4 are labelled I and II.

that were sequenced all contained polyadenylation sites at one position, 166 nt downstream of the stop codon. This site corresponds to the homologous sequence position (163 nt) at which polyadenylation occurred for two of the SRS4 clones and is located in one of the regions identified by S1 nuclease protection and RNase H cleavage analysis (Fig. 6). It appears likely that the differences in the distribution of polyadenylation sites in the SRS1 and SRS4 cDNA clones is related to the under-representation of SRS1 RNAs in the cDNA library and may reflect differences in the 2° structure of the SRS1 transcripts with different 3′ processing sites.

SRS1 and SRS4 are both transcribed at high levels in soybean leaves

The observation that SRS1 RNAs were present at higher levels in total RNA samples than SRS4 RNAs suggested that the SRS1 gene was transcribed at a higher rate or that its RNA was more stable. Transcription rates for the two genes were measured using nuclear run-on assays as described in Berry-Lowe and Meagher [6, 57]. Dot blots containing SRS1 and SRS4 DNA were probed with RNA labelled with [α-^{32}P]UTP in transcriptionally active nuclei isolated from leaves of soybean seedlings. The results of this analysis are shown in Fig. 7. This experiment indicates that SRS4 is transcribed at twice the rate of SRS1, even though levels of SRS1 RNA were four-fold higher in total RNA from soybean leaves. The two-fold higher transcription rate for SRS4 is only an approximate value since some cross-hybridization occurs between full-length SRS1 and SRS4 RNA and the DNA sequences used as probes in this study. Attempts to use gene-specific synthetic oligonucleotides bound to filters in these transcription studies have not been successful. However, this preliminary result suggests that SRS1 RNA is more stable than SRS4 RNA since SRS1 RNA was present at a higher steady state level despite an apparently lower transcription rate.

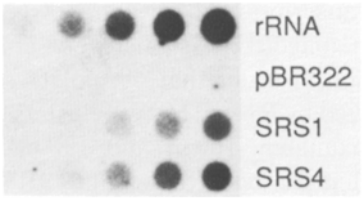

Fig. 7. Transcription of the SRS1 and SRS4 genes in isolated nuclei from soybean seedling leaves. Nitrocellulose filters containing two-fold dilution series of plasmid DNAs were hybridized within run-on transcripts labelled with [α-^{32}P]UTP in nuclei isolated from leaves of light-grown seedlings. The rRNA control contained 1 μg of the soybean 18S rRNA clone pSR1.2B3 [24] in the first dot. 5 μg of the remaining DNA samples, pBR322, λSRS1 [5] and λSRS4 [29] were present in the first dots. Hybridization was in 3 × SSC/50% formamide at 56 °C as described in Berry-Lowe and Meagher [6]. Filters were washed 4 × 15 min at 56 °C in 0.2 × SSC, 0.1% SDS.

SRS1 and SRS4 mRNAs are assembled with equal efficiencies into polyribosomes

Recent studies suggest that SSU RNAs are among the most efficiently translated RNAs in soybean leaves [57]. However, transcripts from the different members of the SSU gene family might be translated with different efficiencies and this could effect differences in RNA stability. The untranslated leaders of SRS1 and SRS4 transcripts are identical except for several nucleotides preceding the start codon. Therefore, polyribosomes from light and dark grown soybean leaves were fractionated on sucrose gradients [57] and the distribution of SRS1 and SRS4 RNAs was determined. RNA was purified from (A) the region at the top of the gradient, presumably containing ribonucleoprotein particles and free RNA, (B) the region of the gradient containing ribosomal subunits and monosomes and (C) the region containing polyribosomes (see Materials and methods) and was analyzed on dot blots (Fig. 8). In the polyribosomal RNA fraction, 4–8-fold more SRS1 and SRS4 RNA was present per μg of RNA loaded than found in the other two fractions. This is identical to what was previously found for the soybean SSU gene family as a whole [57]. In addition, the levels of SRS1 RNA are

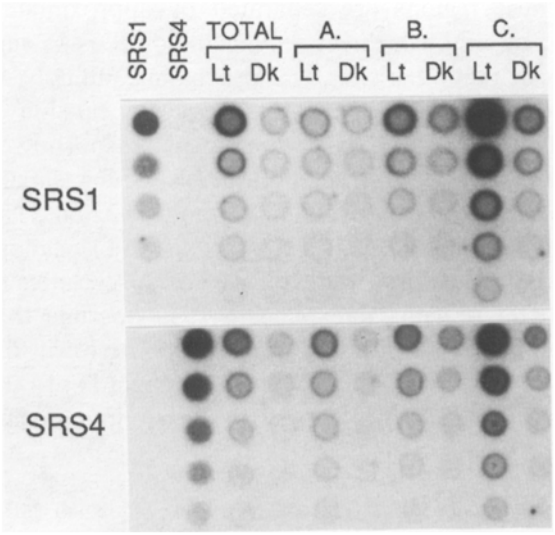

Fig. 8. Prevalence of SRS1 and SRS4 mRNAs in polyribosomes from light (Lt) and dark (Dk) grown soybean seedlings. Dot blots contained two-fold dilution series of SRS1 and SRS4 plasmid DNAs (0.5 fmol in the first dot), total RNA and RNA from three regions of the polysome gradients: A) RNPs and free RNA at the top of the gradient, B) ribosomal subunits and monosomes, C) polysomes (each lane contains 2 μg RNA in the first dot). Filters were hybridized with the SRS1 and SRS4 oligonucleotide probes as described in Grandbastien *et al.* [29].

higher than SRS4 RNA in all three fractions, reflecting the relative levels of these RNAs in total RNA samples. Thus it appears that SRS1 and SRS4 transcripts are assembled into polysomes with approximately equal efficiencies, despite differences in leader sequences and apparent differences in stability.

Discussion

In this study the expression of two members of the soybean SSU gene family was compared. They appear to be the most highly expressed SSU genes in the leaves of soybean seedlings. The patterns of expression of SRS1 and SRS4 RNA were identical in response to treatment with white light, far-red light and darkness. Complex but similar patterns of polyadenylation and comparable poly(A) tail lengths were found for both transcripts. The transcripts were also assembled into polyribosomes with equivalent efficiencies. However, differences were found in the transcription rates, steady-state RNA levels and apparent RNA stabilities of SRS1 and SRS4. In addition, transcripts from SRS1 were inefficiently recovered during poly(A)$^+$ fractionation, which could reflect differences in secondary structure.

RNA stability

Discrepancies between transcription rates and steady-state RNA levels suggest that SRS1 RNA is more stable than SRS4 RNA. In a recent study the 3′ sequences of an *Arabidopsis* SSU gene resulted in significantly higher levels of expression of a reporter gene in transgenic plants than the 3′ sequences of several other diverse genes [40]. Differential processing or stability was suggested as a possible explanation for this result. A number of different mechanisms have been proposed to explain the rate at which RNA is degraded. One hypothesis is that poly(A)$^+$ sequences play a role in protection of the 3′ terminus of mRNA against exonucleolytic attack [11, 12, 66]. Poly(A) tails may be continuously exposed to exonuclease activities and thus more stable mRNAs would acquire shorter tails as they age [10]. However, our data indicate that SRS1 and SRS4 have similar poly(A) tail lengths suggesting that the degradation of poly(A) tails does not control the differential stability of these RNAs. In a second postulate the sequences adjacent to the poly(A) addition site can control the rate at which the poly(A) tail is degraded [4, 39, 66]. A model was proposed in which RNA sequences adjacent to the poly(A) tails modulate the poly(A)-protein interaction, resulting in the differential stability of the poly(A) tails of individual mRNAs. Differential polyadenylation could expose or remove sites which activate degradation. However, S1 nuclease protection and RNase H cleavage experiments demonstrate a similar distribution of polyadenylation sites for SRS1 and SRS4 RNAs. A third possibility is that SSU RNA turnover is linked to translation as it is in a number of animal systems. The degradation of histone [30, 52],

β-tubulin [70] and c-fos [67] mRNAs in mammalian systems have all been shown to be dependent on the association of the mRNAs with ribosomes and in several cases on active translation of the mRNA. However, no significant differences were observed in the assembly of SRS1 and SRS4 mRNA into polyribosomes in soybean seedling leaves. It appears unlikely that the difference in the stabilities of SRS1 and SRS4 RNA is related to differential translation of these transcripts. Thus, the mechanisms that control the overall stabilities of SSU RNAs and contribute to the differential stabilities of SRS1 and SRS4 RNA are not yet known.

The similarity in the distribution of SRS1 and SRS4 in polysomes is also surprising in light of sequence differences in the mRNA leaders of the two genes. One of the most striking differences is the presence of an A at the −3 position relative to the AUG initiation codon in SRS1 and a C at this position in SRS4. According to the observations of Kozak in mammalian systems [42], this should significantly decrease the translation efficiency of SRS4 relative to SRS1. However, transcripts from both genes appear to be very efficiently assembled into polyribosomes. Thus, sequences that direct ribosome loading and translation initiation appear to be different in plants than in animals.

SRS1 and SRS4 mRNA share similar patterns of polyadenylation and poly(A) tail lengths

The poor recovery of SRS1 poly(A)$^+$ RNA relative to SRS4 RNA on oligo dT cellulose and the low frequency of SRS1 clones relative to SRS4 clones in a soybean leaf cDNA library suggested that SRS1 mRNA is retained less efficiently by oligo(dT) cellulose than SRS4 mRNA. This observation suggested that SRS1 mRNA might have shorter poly(A) tails than SRS4 mRNA. However, S1 protection analysis and oligonucleotide directed RNase H cleavage demonstrated that both SRS1 and SRS4 transcripts were polyadenylated at multiple sites concentrated in two regions of the RNA molecules (I, II, Fig. 6). These regions are separated by approximately 35 nt although direct sequencing of cDNAs suggests that other sites of 3′ processing and poly(A) addition occur at a lower frequency outside of these regions. These experiments also showed that transcripts from both genes exhibit similar poly(A) tail lengths ranging from 0 to 220 nt. The low affinity of SRS1 RNA for oligo(dT) cellulose and the low frequency of SRS1 cDNA relative to SRS4 remains unexplained. It is possible that subtle conformational differences between the RNAs sterically block access of oligo(dT) to the poly(A) tails of the SRS1 transcripts during chromatography.

Sequences involved in 3′ processing and polyadenylation

The locations of polyadenylation sites in the SRS1 and SRS4 cDNAs provide clues regarding the signals that are involved in the 3′ processing of SSU RNAs. The 3′ non-translated regions of SRS1 and SRS4 contain numerous sequences homologous to the consensus polyadenylation signals in animals [5, 29]. One example is the sequence element, AAUAAA, which is a signal for polyadenylation in animals and is usually found 10 to 40 nt upstream of the polyadenylation site [50]. Sequences homologous to AAUAAA (indicated in bold in Fig. 5) are present immediately upstream of the SRS1 polyadenylation site and the two clusters of polyadenylation in SRS4. However, no AAUAAA-like sequence is found upstream of the most 5′ proximal polyadenylation site in SRS4. In addition, no polyadenylation sites were found downstream of the third AAUAAA element in the 10 cDNA clones examined. The sequence CAYUG is another sequence element often found either immediately upstream or downstream of the 3′ processing site in animals [2]. The CAYUG sequences is 57 nt downstream of the polyadenylation site in SRS1 (CATTG, Fig. 6), but is disrupted by a nucleotide change from T to C in SRS4 (CATCG, Fig. 6). No other CAYUG element is found in the 3′ end of either gene. Another element, the G/T cluster, is found

approximately 30 nt downstream of the AAUAAA element in many animal genes [8]. G/T-rich regions are present downstream of all of the polyadenylation sites in both SRS1 and SRS4. Interestingly, all of the SRS1 and SRS4 polyadenylation sites occurred at A residues, which agrees with the general pattern for polyadenylation in animals [8].

Plant transcripts have complex patterns of polyadenylation

Even though polyadenylation appears to be ubiquitous among eukaryotes and probably involves highly conserved mechanisms, the analysis of the polyadenylation of SRS1 and SRS4 RNAs provides further evidence for differences in the polyadenylation processes of plants and animals. Examples have been found of multiple polyadenylation sites at the 3′ ends of animal genes and in some cases multiple polyadenylation signals are used within a gene to regulate synthesis of alternative protein products. However, polyadenylation generally occurs at a single precisely defined site in animal RNAs. On the other hand, in plant systems multiple polyadenylation sites may be quite common [18, 38, 46]. A survey of published plant genes which appear to contain multiple polyadenylation sites based on cDNA sequence analysis or S1 nuclease protection assays is presented in Table 1. In many other studies only one cDNA clone was sequenced in order to define the polyadenylation site for a particular plant gene. Complex patterns of polyadenylation observed in S1 protection analyses are often attributed to stuttering of the enzyme or to the presence of closely related gene family members. However, substantial evidence exists that multiple polyadenylation sites are used in many plant genes and that, in general, polyadenylation in plants may be less precise than in animal systems.

Evidence for differences in the polyadenylation processes in plants and animals has also come from two studies on the processing of animal genes in plant cells. Hunt *et al.* [37] showed that polyadenylation signals from the human growth hormone gene and two animal virus genes are processed inefficiently or incorrectly in transgenic tobacco cells. In addition, Martinez-Zapater *et al.* [44] recently showed that when a gene encoding the *Drosophila* P-element transposase was expressed in tobacco plants, the 3′ end of the most abundant transcript was located in a region of high A + T content in the transposase coding sequence. Thus, significant differences must exist between plants and animals in the signals and/or mechanisms that control 3′ cleavage and polyadenylation.

Although 3′ processing and polyadenylation have been extensively studied in animal systems, the signals that control these processes in plants are not yet well defined. Functional tests have demonstrated the role of the conserved sequences. AAUAAA and CAYUG, and GT-rich regions in the 3′ processing and polyadenylation of mRNAs in animal systems [8]. In addition, a 64 kDa protein has been found to interact specifically with viral AAUAAA sequences [68] and regions of the highly conserved small nuclear RNA U4 have been found to be complementary to the AAUAAA and CAYUG sequences [3]. Many plant genes possess sequences homologous to these control signals, often in appropriate locations relative to known polyadenylation sites (Table 1) but their significance is unknown. In many cases these sequences have been assumed to be involved in 3′ processing and polyadenylation. However, the 3′ untranslated regions of a number of plant genes do not contain these sequences [37]. In addition, an AAUAAA-like sequence is not present upstream of the first polyadenylation site in SRS4.

Only a few functional tests have been performed to identify the sequences required for 3′ processing and polyadenylation in plants. Analysis of a pea SSU deletion mutant in transgenic plants showed that sequences between − 235 and − 6 (relative to the major polyadenylation site) are required for normal polyadenylation [36]. Deletion of these sequences results in polyadenylation at at least seven new sites located from 20 to 300 nt downstream of the normal sites. Deletion

Table 1. Plant genes with multiple polyadenylation sites.

Species	Gene	Gene designation	A_2UA_3-like sequences	Number of Poly (A) sites	Method	Ref.
A. tumefaciens[1]	ocs[2]	wt	4	2	S1	[21]
A. tumefaciens	ocs[2]	rGV1(del)[3]	4	4	S1	[21]
A. tumefaciens	nos[4]	–	1	3	S1	[7]
castor bean	ricin	–	2	2	cDNAs (2)	[34]
French bean	endochitinase	–	1	2	cDNAs (4)	[13]
maize	*bronze*	–	2	3	cDNAs (6)	[18]
maize	GST[5]	GST-I	1	4	cDNAs (5)	[32]
maize	GST	GST-III	2	4	cDNAs (5)	[32]
maize	PEPC[6]	–	3	4	cDNAs (5)	[69]
maize	*Adh*[7]	*Adh 1-S*	2	4	S1	[53]
maize	*Adh*	*Adh 1-F*	2	7	S1	[53]
maize	*waxy*	*Adh 1-F*	3	3	cDNAs (3)	[41]
parsley	4CL[8]	Pc4CL-1/2	0	10	cDNAs (17)	[22]
pea	seed lectin	*LecA*	1	2	cDNAs (2)	[28]
petunia	cab[10]	cab91R	2	2	cDNAs (2)	[18]
petunia	cab	cab29	2	2	cDNAs (2)	[18]
petunia	Rubisco SSU[11]	SSU301	3	3	cDNAs (12)	[18]
petunia	Rubisco SSU	SSU511	3	3	cDNAs (13)	[18]
petunia	Rubisco SSU	SSU211	2	2	cDNAs (2)	[18]
potato	not determined	SF-LS1	2	2	cDNAs (2)	[25]
rice	glutelin	type I	2	2	cDNAs (2)	[61]
rice	glutelin	type II	2	2	cDNAs (2)	[61]
soybean	lectin	Le1	1	2	S1	[64]
soybean	hsp[12]	Gmhsp17.5-E	3	4	S1	[17]
TMV[13]	PRP[14]	PR-1a	3	2	cDNAs (2)	[47]
TMV[13]	PRP	PR-1c	1	2	cDNAs (2)	[47]
tobacco	ATPS βSU[15]	atp2-1	1	4	S1	[9]
tobacco	Rubisco SSU	rbcS-8B	1	2	S1	[49]
tobacco	stylar glycoprotein	S_2	2	2	cDNAs (2)	[1]
wheat	a/b gliadin	–	2	2	cDNAs (3)	[51]
wheat	lipoxygenase 3	–	4	2	cDNAs (2)	[71]

[1] *Agrobacterium tumefaciens*; [2] octopine synthase; [3] 3′ deletion mutant of octopine synthase gene; [4] nopaline synthase; [5] glutathione S-transferase; [6] phospho*enol*pyruvate carboxylase; [7] alcohol dehydrogenase; [8] 4-coumarate : CoA ligase; [9] the authors do not identify which of the 17 cDNAs are Pc4CL-1 or Pc4CL-2; [10] Chlorophyll a/b-binding protein; [11] ribulose-1,5-bisphosphate carboxylase small subunit; [12] heat shock protein; [13] tobacco mosaic virus; [14] pathogenesis-related protein; [15] Beta subunit of the mitochondrial ATP synthase.

of an additional 57 nt causes loss of the cryptic polyadenylation sites as well [36]. A more recent study suggests that regions as large as 100 nt on either side of the major polyadenylation sites in a pea SSU gene are required for normal processing [38]. These results suggest that distant flanking sequences play a role in directing the complex polyadenylation patterns observed in plants. However, the specific sequences required for 3′ processing and polyadenylation of plant RNAs have not yet been identified.

Conclusion

In this study the expression of two members of the soybean Rubisco SSU gene family, SRS1 and SRS4, was compared. These genes share exten-

sive homology, even in distant flanking sequences, thus the similarity in their expression in response to light is perhaps not surprising. These similarities may result from periodic gene conversion events among SSU gene family members [45] and a selection for organisms with coordinately regulated SSU gene family members. However, an apparent difference in the stabilities of the transcripts from these genes was inferred from our data. Although RNA stability has been linked to polyadenylation and translation in other systems, no significant differences were found between SRS1 and SRS4 in complex patterns of polyadenylation, in poly(A) tail lengths or in efficiency of assembly into polyribosomes. Future studies will focus more specifically on the sequences and mechanisms that direct post-transcriptional control of these genes.

Acknowledgements

We would like to acknowledge Jim Zobel for his assistance in preparing λgt10 cloned cDNA and Deborah Thompson for giving useful technical advice and revisions of the manuscript. Portions of this work serve as partial fulfillment of a M.S. degree in Genetics for D.P.H. and a Ph.D. degree in Genetics for S.B.-L. and B.W.S. This project was supported by a grant from the Department of Energy. S.B.-L. and B.W.S. were supported by NIH Predoctoral Training Grants and J.F.S. by an NIH postdoctoral fellowship.

References

1. Anderson MA, Cornish EC, Mau S-L, Williams EG, Hoggart R, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penschow JD, Hiall HD, Tregear GW, Coghlan JP, Crawford RJ, Clarke AE: Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature 321: 38–44 (1986).
2. Benoist C, O'Hare K, Breathnach R, Chambon P: The ovalbumin gene – sequence of putative control regions. Nucleic Acids Res 8: 127–142 (1980).
3. Berget SM: Are U4 small nuclear ribonucleoproteins involved in polyadenylation? Nature 309: 179–182 (1984).
4. Bergmann IE, Brawerman G: Control of the breakdown of the polyadenylate sequence in mammalian polyribosomes: role of poly(adenylic acid)-protein interactions. Biochemistry 16: 259–264 (1977).
5. Berry-Lowe SL, McKnight TD, Shah DM, Meagher RB: The nucleotide sequence, expression, and evolution of one member of a multigene family encoding the small subunit of ribulose-1,5-bisphosphate carboxylase in soybean. J Mol Appl Genet 1: 483–498 (1982).
6. Berry-Lowe SL, Meagher RB: Transcriptional regulation of a gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase in soybean tissue is linked to the phytochrome response. Mol Cell Biol 5: 1910–1917 (1985).
7. Bevan M, Barnes WM, Chilton M-D: Structure and transcription of the nopaline synthase gene region of T-DNA. Nucleic Acids Res 11: 369–385 (1983).
8. Birnstiel ML, Busslinger M, Strub K: Transcription termination and 3′ processing: the end is in site! Cell 41: 349–359 (1985).
9. Boutry M, Chua N-H: A nuclear gene encoding the beta subunit of the mitochondrial ATP synthase in *Nicotiana plumbaginifolia*. EMBO J 4: 2159–2165 (1985).
10. Brawerman G: Characteristics and significance of the polyadenylate sequence in mammalian messenger RNA. Prog Nucl Acid Res Mol Biol 17: 117–148 (1976).
11. Brawerman G: The role of the poly(A) sequence in mammalian messenger RNA. CRC Crit Rev Biochem 10: 1–38 (1981).
12. Brewer G, Ross J: Poly(A) shortening and degradation of the 3′ A + U-rich sequences of human c-myc mRNA in a cell-free system. Mol Cell Biol 8: 1697–1708 (1988).
13. Broglie KE, Gaynor JJ, Broglie RM: Ethylene-regulated gene expression: molecular cloning of the genes encoding endochitinase from *Phaseolus vulgaris*. Proc Natl Acad Sci USA 83: 6820–6824 (1986).
14. Broglie R, Coruzzi G, Lamppa G, Keith B, Chua N-H: Structural analysis of nuclear gene coding for the precursor to the small subunit of wheat ribulose-1,5-bisphosphate carboxylase. Bio/technology 1: 55-61 (1983).
15. Carrazana EJ, Pasieka KB: The vasopressin mRNA poly(A) tract is unusually long and increases during stimulation of vassopressin gene expression *in vivo*. Mol Cell Biol 8: 2267–2274 (1988).
16. Condit CM, Meagher RB: Expression of a gene encoding a glycine-rich protein in petunia. Mol Cell Biol 7: 4273–4279 (1987).
17. Czarnecka E, Gurley WB, Nagao RT, Mosquere LA, Key JL: DNA sequence and transcript mapping of a soybean gene encoding a small heat shock protein. Proc Natl Acad Sci USA 82: 3726–3730 (1985).
18. Dean G, Tamaki S, Dunsmuir P, Favreau M, Katayama C, Dooner H, Bedbrook J: mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vivo*. Nucleic Acids Res 5: 2229–2240 (1986).
19. Dean C, van den Elzen P, Tamaki S, Dunsmuir P, Bedbrook J: Differential expression of the eight genes of

the *Petunia* ribulose bisphosphate carboxylase small subunit multi-gene family. EMBO J 4: 3055–3061 (1985).
20. Denhardt DT: A membrane-filter technique for the detection of complementary DNA. Biochem Biophys Res Comm 23: 641–646 (1966).
21. Dhaese P, De Greve H, Gielen J, Seurinck J, Van Montagu M, Schell J: Identification of sequences involved in the polyadenylation of higher plant nuclear transcripts using *Agrobacterium* T-DNA genes as models. EMBO J 2: 419–426 (1983).
22. Douglas C, Hoffmann H, Schulz W, Hahlbrock K: Structure and elicitor or u.v.-light-stimulated expression of two 4-coumarate : CoA ligase genes in parsley. EMBO J 6: 1189–1195 (1987).
23. Dover CA: Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Theor Appl Genet 2: 159–165 (1986).
24. Eckenrode BK, Arnold J, Meagher RB: Comparison of the nucleotide sequence of soybean 18S rRNA with the sequences of other small-subunit rRNAs. J Mol Evol 21: 259–269 (1985).
25. Eckes P, Rosahl S, Schell J, Willmitzer L: Isolation and characterization of a light-inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. Mol Gen Genet 205: 14–22 (1986).
26. Fitzwater T, Chang X-Y, Elble R, Polisky B: Conditional high copy ColE1 mutants: resistance to RNA1 inhibition *in vivo* and *in vitro*. EMBO J 7: 3289–3297 (1988).
27. Fluhr R, Moses P, Morelli G, Coruzzi G, Chua N-H: Expression dynamics of the pea rbcS multigene family and organ distribution of transcripts. EMBO J 5: 2063–2071 (1986).
28. Gatehouse JA, Bown D, Evans IM, Gatehouse LN, Jobes D, Preston P, Croy RRD: Sequence of the seed lectin gene from pea (*Pisum sativum* L.). Nucleic Acids Res 15: 7642 (1987).
29. Grandbastien MA, Berry-Lowe S, Shirley BW, Meagher RB: Two soybean ribulose-1,5-bisphosphate carboxylase small subunit genes share extensive homology even in distant flanking sequences. Plant Mol Biol 7: 451–465 (1986).
30. Graves RA, Pandey NP, Chodchoy N, Marzluff WF: Translation is required for regulation of histone mRNA degradation. Cell 48: 615–626 (1987).
31. Grossberger D: Minipreps of DNA from bacteriophage lambda. Nucleic Acids Res 15: 6737 (1987).
32. Grove G, Zarlengo RP, Timmerman KP, Li N-Q, Tam MF, Tu C-PD: Characterization and heterospecific expression of cDNA clones of genes in the maize GSH S-transferase multigene family. Nucleic Acids Res 16: 425–437 (1988).
33. Hall TC, Ma Y, Buchbinder BU, Pyne JW, Sun SM, Bliss FA: Messenger RNA for G1 protein of French bean seeds: Cell-free translation and product characterization. Proc Natl Acad Sci USA 75: 3196–3200 (1978).
34. Halling KC, Halling AC, Murray EE, Ladin BF, Houston LL, Weaver RF: Genomic cloning and characterization of a ricin gene from *Ricinus communis*. Nucleic Acids Res 13: 8019–8033 (1985).
35. Hanahan D: Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166: 557–580 (1983).
36. Hunt AG: Identification and characterization of cryptic polyadenylation sites in the 3′ region of a pea ribulose-1,5-bisphosphate carboxylase small subunit gene. DNA 7: 329–336 (1988).
37. Hunt AG, Chu NM, Odell JT, Nagy F, Chua N-H: Plant cells do not properly recognize animal gene polyadenylation signals. Plant Mol Biol 8: 23–35 (1987).
38. Hunt AG, MacDonald MH: Deletion analysis of the polyadenylation signal of a pea ribulose-1,5-bisphosphate carboxylase small subunit gene. Plant Mol Biol 13: 125–138 (1989).
39. Huynh TV, Young RA, Davis RW: Constructing and screening cDNA libraries in λgt10 and λgt11. In: Glover DM(eds) DNA Cloning: A Practical Approach, vol. I, pp. 49–78. IRL press, Oxford, Washington, DC (1985).
40. Ingelbrecht ILW, Herman LMF, Dekeyser RA, Van Montagu MC, Depicker AG: Different 3′ end regions strongly influence the level of gene expression in plant cells. Plant Cell 1: 671–680 (1989).
41. Klösgen RB, Gierl A, Schwarz-Sommer Z, Saedler H: Molecular analysis of the *waxy* locus of *Zea mays*. Mol Gen Genet 203: 237–244 (1986).
42. Kozak M: Point mutations close to the AUG initiator codon affect the efficiency of translation of rat preproinsulin *in vivo*. Nature 308: 241–246 (1984).
43. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
44. Martinez-Zapater JM, Finkelstein R, Somerville CR: *Drosophila* P-element transcripts are incorrectly processed in tobacco. Plant Mol Biol 11: 601–607 (1988).
45. Meagher RB, Berry-Lowe S, Rice K: Molecular evolution of the small subunit of ribulose bisphosphate carboxylase: nucleotide substitution and gene conversion. Genetics 123: 845–863 (1989).
46. Messing J, Geraghty D, Heidecker G, Hu N-T, Kridly J, Rubenstein I: Plant gene structure. In: Hollaender A (eds) Genetic Engineering of Plants, pp. 211–227. Plenum Press, New York (1983).
47. Pfitzner UM, Goodman HM: Isolation and characterization of cDNA clones encoding pathogenesis-related proteins from tobacco mosaic virus infected tobacco plants. Nucleic Acids Res 15: 4449–4465 (1987).
48. Pichersky E, Bernatsky R, Tanksley SD, Cashmore AR: Evidence for selection as a mechanism in the concerted evolution of *Lycopersicon esculentum* (tomato) genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. Proc Natl Acad Sci USA 83: 3880–3884 (1986).
49. Poulsen C, Fluhr R, Kauffman JM, Boutry M, Chua

N-H: Characterization of an *rbc* S gene from *Nicotiana plumbaginifolia* and expression of an *rbc* S-CAT chimeric gene in homologous and heterologous nuclear background. Mol Gen Genet 205: 193–200 (1986).
50. Proudfoot NJ, Brownlee GG: 3′ non-coding region sequences in eukaryotic messenger RNA. Nature 263: 211–214 (1976).
51. Rafalski JA, Scheets K, Metzier M, Peterson DM, Hedgcoth C, Soll DG: Developmentally regulated plant genes: the nucleotide sequence of a wheat gliadin genomic clone. EMBO J 3: 1409–1415 (1984).
52. Ross J, Kobs G, Brewer G, Peltz SW: Properties of the exonuclease activity that degrades H4 histone mRNA. J Biol Chem 19: 9374–9381 (1987).
53. Sachs MM, Dennis ES, Gerlach WL, Peacock WJ: Two alleles of maize alcohol dehydrogenase I have 3′ structural and poly(A) addition polymorphisms. Genetics 113: 449–467 (1986).
54. Senecoff JF, Rossmeissl PJ, Cox MM: DNA recognition by the FLP recombinase of the yeast 2 μ plasmid. A mutational analysis of the FLP binding site. J Mol Biol 201: 405–421 (1988).
55. Sheen J-Y, Bogorad L: Expression of the ribulose-1,5-bisphosphate carboxylase large subunit gene and three small subunit genes in two cell types of maize leaves. EMBO J 5: 3417–3422 (1986).
56. Shirley BW, Berry-Lowe SL, Rogers SG, Flick JS, Horsch R, Fraley RT, Meagher RB: 5′ proximal sequences of a soybean ribulose-1,5-bisphosphate carboxylase small subunit gene direct light and phytochrome controlled transcription. Nucleic Acids Res 15: 6501–6514 (1987).
57. Shirley BW, Meagher RB: The role of RNA turnover in the regulation of plant gene expression: ribulose-1,5-bisphosphate carboxylase small subunit in soybean. Submitted to Nucleic Acids Res.
58. Shure M, Wessler S, Federoff N: Molecular identification and isolation of the *Waxy* locus in maize. Cell 35: 225–233 (1983).
59. Simpson J, Van Montagu M, Herrera-Estrella L: Photosynthesis-associated gene families: differences in response to tissue-specific and environmental factors. Science 233: 34–38 (1986).
60. Sugita M, Manzara T, Pichersky E, Cashmore A, Gruissem W: Genomic organization, sequence analysis and expression of all five genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from tomato. Mol Gen Genet 209: 247–256 (1987).
61. Takaiwa F, Kikuchi S, Oono K: A rice glutelin gene family. A major type of glutelin mRNAs can be divided into two classes. Mol Gen Genet 208: 15–22 (1987).
62. Tobin EM, Silverthorne L: Light regulation of gene expression in higher plants. Ann Rev Plant Physiol 36: 569–593 (1985).
63. Tumer NE, Clark WG, Tabor GJ, Hironaka CM, Fraley RT, Shah DM: The genes encoding the small subunit of ribulose-1,5-bisphosphate carboxylase are expressed differentially in petunia leaves. Nucleic Acids Res 8: 3325–3342 (1986).
64. Vodkin LO, Rhodes PR, Goldberg RB: A lectin gene insertion has the structural features of a transposable element. Cell 34: 1023–1031 (1983).
65. Watson JC: Photoregulation of gene expression in plants. In: Kung SD, Arntzen CJ (eds) Plant Biotechnology, pp. 161–205. Butterworths, Boston (1989).
66. Wilson MC, Sawicki SG White PA: A correlation between the rate of poly(A) shortening and half-life messenger RNA in adenovirus transformed cells. J Mol Biol 126: 23–36 (1978).
67. Wilson T, Treisman R: Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3′ AU-rich sequences. Nature 336: 396–399 (1988).
68. Wilusz J, Shenk T: A 64 kd nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. Cell 52: 221–228 (1988).
69. Yanagisawa S, Izui K, Yamaguchi Y, Shigesada K, Katsuki H: Further analysis of cDNA clones for maize phosphoenolpyruvate carboxylase involve in C_4 photosynthesis. Nucleotide sequence of an entire open reading frame and evidence for polyadenylation of mRNA at multiple sites *in vivo*. FEBS Lett 229: 107–110(1988).
70. Yen TJ, Gay DA, Pachter JS, Cleveland DW: Autoregulated changes in stability of polyribosome-bound β-tubulin mRNAs are specified by the first 13 translated nucleotides. Mol Cell Biol 8: 1224–1235 (1988).
71. Yenofsky RL, Fine M, Liu C: Isolation and characterization of a soybean (*Glycine max*) lipoxygenase-3 gene. Mol Gen Genet 211: 215–222 (1988).
72. Zurfluh LL, Guilfoyle TJ: Auxin-induced changes in the population of translatable messenger RNA in elongating sections of soybean hypocotyl. Plant Physiol 69: 332–337 (1982).