AtSig5 Is an Essential Nucleus-Encoded Arabidopsis σ-Like Factor[1]

Junlan Yao, Sanchita Roy-Chowdhury, and Lori A. Allison*

Department of Biochemistry, University of Nebraska-Lincoln, N250 Beadle Center, Lincoln, Nebraska 68588–0664

Transcription of chloroplast genes is subject to control by nucleus-encoded proteins. The chloroplast-encoded RNA polymerase (PEP) is a eubacterial-type RNA polymerase that is presumed to assemble with nucleus-encoded σ-factors mediating promoter recognition. Recently, families of σ-factor genes have been identified in several plants including Arabidopsis. One of these genes, Arabidopsis *SIG5*, encodes a σ-factor, AtSig5, which is phylogenetically distinct from the other family members. To investigate the role of this plant σ-factor, two different insertional alleles of the *SIG5* gene were identified and characterized. Heterozygous mutant plants showed no visible leaf phenotype, but exhibited siliques containing aborted embryos and unfertilized ovules. Our inability to recover plants homozygous for a *SIG5* gene disruption indicates that *SIG5* is an essential gene. *SIG5* transcripts accumulate in flower tissues, consistent with a role for AtSig5 protein in reproduction. Therefore, *SIG5* encodes an essential member of the Arabidopsis σ-factor family that plays a role in plant reproduction in addition to its previously proposed role in leaf chloroplast gene expression.

Transcription of plant mitochondrial and plastid genomes relies on nucleus-encoded RNA polymerases resembling those of the T3 and T7 bacteriophage, as well as promoter selectivity factors for these enzymes (for review, see Hess and Börner, 1999; Liere and Maliga, 2001). In addition, plastids, but not mitochondria, require an organelle-encoded RNA polymerase (PEP) for transcription of many genes on the organelle genome (Allison and Maliga, 1996; Hajdukiewicz et al., 1997; Serino and Maliga, 1998). PEP is similar in structure to *Escherichia coli* RNA polymerase and, like the eubacterial enzyme, is thought to assemble with σ-factors to achieve promoter-specific transcription initiation. Although the subunits of the PEP catalytic core are plastid-encoded, the putative promoter-specificity σ-factors for PEP are encoded in the plant nucleus. Therefore, even the organelle-encoded transcription machinery is subject to nuclear control.

Recently, σ-factor gene families containing as many as six members have been identified in several plant species including Arabidopsis and maize (*Zea mays*; for review, see Allison, 2000). Within a species, the sigma-like proteins share approximately 35% overall amino acid sequence identity and contain the conserved domains found in the principal σ-factors of all eubacteria (Helmann and Chamberlin, 1988). Genetic analyses of the eubacterial σ-factors, as well as more recent structural studies, have supported the importance of these conserved domains in such eubacterial transcription functions as promoter recognition, interaction with the RNA polymerase core enzyme, and transcription initiation (Gross et al., 1998; Burgess and Anthony, 2001; Campbell et al., 2002; Murakami et al., 2002). The conservation of primary structure among plant and eubacterial σ-factors is sufficiently high to permit recombinant plant σ-factors to function with *E. coli* core RNA polymerase in vitro (Kestermann et al., 1998; Hakimi et al., 2000; Beardslee et al., 2002). Therefore, it is assumed (although not yet experimentally demonstrated) that these nucleus-encoded plant σ-factors assemble in plastids with the PEP RNA polymerase to effect promoter recognition and transcription initiation. Consistent with this model, many plant σ-factors contain transit peptides capable of targeting the proteins into chloroplasts (Isono et al., 1997; Kanamaru et al., 1999; Lahiri et al., 1999; Fujiwara et al., 2000; Lahiri and Allison, 2000). Interestingly, the maize σ-factor, ZmSig2B, accumulates not only in chloroplasts but also in mitochondria, and contains an NH_2-terminal sequence able to target fused green fluorescent protein (GFP) into mitochondria in transient expression assays (Beardslee et al., 2002). A specific role for this σ-factor in mitochondria remains to be elucidated.

The presence of multiple σ-factor genes in a single plant species suggests several possible models for σ-factor function, none of which is mutually exclusive. σ-Factors exhibiting developmental- or tissue-specific expression profiles may mediate the differential plastid gene promoter activity observed during plant development (Satoh et al., 1999; Lahiri and Allison, 2000). Different σ-factors expressed in the same plant tissues may recognize distinct subsets of

[1] This work was supported by the National Research Initiative Competitive Grants Program/U.S. Department of Agriculture (grant nos. 97–35301–4514 and 00–35301–9016 to L.A.A.).

* Corresponding author; e-mail lallison@creighton.edu; fax 402–472–7842.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.102.017913.

plastid gene promoters. It is also possible that plant σ-factors are functionally redundant. Dissection of individual σ-factor function through reverse genetics approaches is feasible in Arabidopsis and maize and has begun to yield results, indicating that at least one plant σ-factor does recognize a specific group of plastid promoters. Disruption of the Arabidopsis *SIG2* gene by a T-DNA insertion in exon 6 (encoding conserved domain 3), resulted in homozygous mutant plants exhibiting pale green leaves and aberrant chloroplast development (Shirano et al., 2000). In these plants, expression of a subset of tRNAs was reduced, indicating that the AtSig2 protein recognizes certain plastid tRNA gene promoters (Kanamaru et al., 2001). Because low levels of transcripts were still detected for the affected tRNAs, the lack of AtSig2 in the knockout plants appeared to be partially compensated by activity from one or more of the other σ-factors.

Recently, a potential promoter target was proposed for another Arabidopsis σ-factor, AtSig5, encoded by the *SIG5* gene. *SIG5* is unique among the Arabidopsis *SIG* gene family in that it is the only member whose expression is induced in leaves by blue light (Tsunoyama et al., 2002). This observation suggests that AtSig5 in leaf chloroplasts may specifically recognize the blue-light responsive promoter of the plastid *psbD* gene, whose activation is required for continued expression of *psbD* and its product, subunit D2 of photosystem II, in high-light conditions (Thum et al., 2001). The *SIG5* gene is unique in other ways. In phylogenetic analyses, *SIG5* does not cluster closely with the other plant σ-factors but appears to form its own branch (Allison, 2000; Fujiwara et al., 2000). Moreover, the *SIG5* gene does not exhibit any of the conserved intron sites that are shared by the other five Arabidopsis σ-factor genes (Fujiwara et al., 2000).

On the basis of these observations, we reasoned that AtSig5 is a unique σ-factor whose function in the plant may be revealed by examining insertional mutants in the *SIG5* gene. Here, we report analysis of two independently derived T-DNA insertional *sig5* mutants. Our results suggest that AtSig5 plays a role in plant reproduction in addition to its previously proposed role in green leaf chloroplast transcription.

RESULTS

Identification of Arabidopsis Lines with T-DNA Insertions in SIG5

To study the function of AtSig5 in Arabidopsis, we screened two collections of insertional mutants for lines with T-DNAs located within the *SIG5* gene. One such T-DNA-tagged line was recovered from each collection. The *sig5-1* allele was identified in the WS ecotype by PCR-based screens of DNA pools derived from a collection of 60,480 Arabidopsis lines transformed with a derivative of the T-DNA vector (Krysan et al., 1999). In the recovered line, the T-DNA insertion was located after nucleotide number 198 of *sig5* exon 5 (Fig. 1A). This insertion would generate an AtSig5 protein truncated in conserved region 3 and completely missing conserved region 4. Because these conserved domains are essential for bacterial σ-factor activity (Gross et al., 1998; Murakami et al., 2002), it is presumed that a truncated AtSig5 protein expressed from *sig5-1* would be nonfunctional. The second T-DNA insertional allele, *sig5-2*, was detected

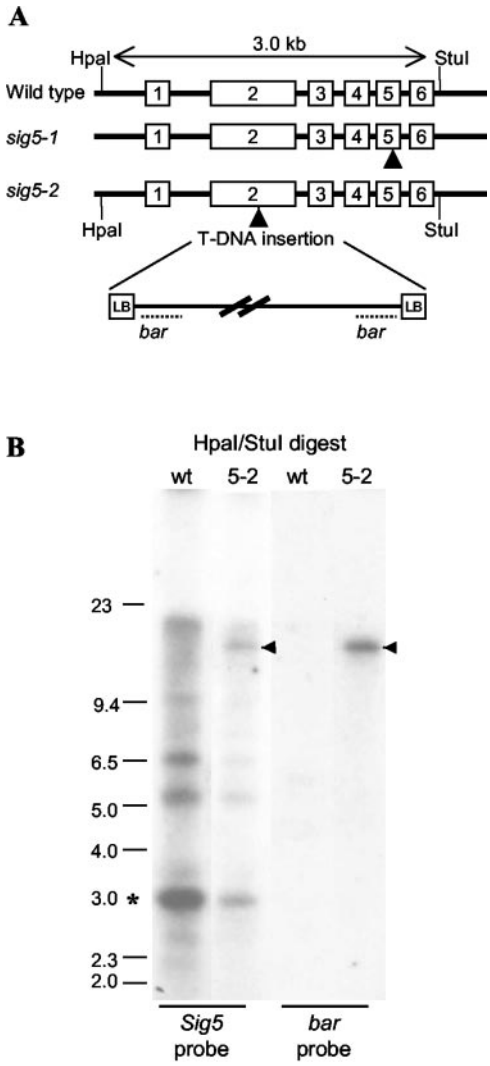

Figure 1. Characterization of T-DNA insertions in two *SIG5* knockout lines, *sig5-1* and *sig5-2*. A, Diagram illustrating the exon-intron structure of the *SIG5* gene (exons are depicted as numbered boxes) and the position of T-DNA insertions (arrowheads) in *sig5-1* and *sig5-2*. Recognition sites for restriction enzymes *Hpa*I and *Stu*I and the sizes of fragments generated by digestions of wild-type DNA are indicated. B, DNA gel blot of wild-type and heterozygous *sig5-2* plant DNA digested with *Hpa*I and *Stu*I and probed with either a *SIG5* (left) or T-DNA (bar, right) probe. The asterisk marks the wild-type *SIG5* band, whereas the arrowhead indicates the hybridizing band corresponding to the *sig5-2* insertional allele. Additional bands hybridizing to the *SIG5* probe are most likely derived from related σ-factor gene family members.

in the Columbia ecotype by screening a database of left border flanking sequences recovered by thermal asymmetric interlaced PCR from a collection of approximately 100,000 individual T-DNA mutagenized Arabidopsis plants (McElver et al., 2001). In the identified line, the insertion follows nucleotide number 381 of the second exon of the *sig5* gene (Fig. 1A) and would generate an AtSig5 protein missing all of the conserved domains essential for bacterial σ-factor activity. Thus, any protein expressed from *sig5-2* would be predicted not to function as a σ-factor. DNA gel blots revealed more than one T-DNA insertion site in the line containing *sig5-1* (data not shown). DNA gel-blot analyses on *sig5-2* DNA, digested with restriction endonucleases that did not cut within the T-DNA, indicated a single T-DNA insertion site in the *sig5-2* line (Fig. 1B, arrowhead). The large size of the restriction fragment hybridizing to the T-DNA-derived *bar* gene probe revealed that multiple T-DNAs had been inserted at that site. This was confirmed by PCR analysis (data not shown), although the exact arrangement and number of the T-DNAs was not determined. Because the *sig5-2* line contained T-DNAs at a single insertion site, most of our analysis was done with the progeny of this line. We note that the phenotypes observed with the *sig5-2* line were also seen in progeny from the *sig5-1* line.

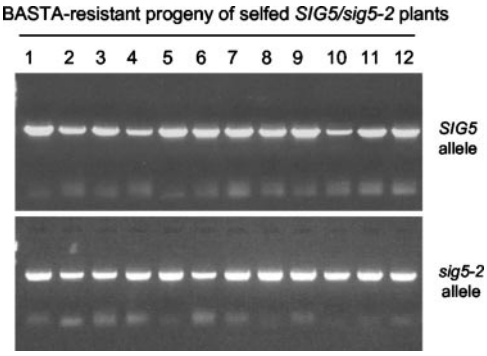

Figure 2. Homozygous *sig5-2* lines could not be recovered from the selfed progeny of heterozygous *sig5-2* plants. DNA prepared from 12 BASTA-resistant progeny was subjected to PCR analysis using primers flanking the T-DNA insertion site (top panel). The presence of a wild-type allele in progeny plants was indicated by the amplification of a 2.5-kb PCR product (*SIG5*). The large PCR product that would have resulted from the mutant insertional allele was not readily amplified. The same DNA samples were subjected to PCR with a *SIG5* primer adjacent to the T-DNA insertion site and a T-DNA left border primer (bottom panel). The amplification of a product of the predicted size for the insertional allele (*sig5-2*) indicated the presence of the mutant *sig5-2* allele in every BASTA-resistant progeny line recovered.

Analysis of *sig5-1* and *sig5-2* Lines

Disruption of the gene for a plastid-localized σ-factor could potentially generate plants with defects in chloroplast gene expression and development, leading to a visible pale green leaf phenotype (Shirano et al., 2000). Because heterozygous mutant *sig5-1* and *sig5-2* lines had no visible leaf phenotype, we attempted to uncover lines homozygous for either allele. However, in neither case were any homozygous progeny recovered from selfed heterozygous mutant plants. The T2 and T3 progeny from *sig5-2* selfed heterozygotes were two-thirds BASTA-resistant, and one-third sensitive, consistent with the lack of viable homozygous progeny. Sixty T-DNA-containing progeny (BASTA-resistant) of selfed *sig5-2* heterozygotes were analyzed by PCR for the presence or absence of a wild-type *SIG5* allele and for the presence of the *sig5-2* allele. In each case, both a wild-type and a mutant allele were recovered demonstrating the presence of only heterozygous progeny (PCR results for a subset of the 60 progeny lines are shown in Fig. 2).

The lack of homozygous progeny indicated that *SIG5* might be essential for normal seed or embryo development. Therefore, siliques from selfed heterozygous *sig5-1* and *sig5-2* lines were examined for defective seeds. In both of the mutant lines, small wrinkled seeds and normal-looking seeds were detected in mature siliques (Fig. 3A). In addition, some immature siliques of both mutant lines contained gaps in the seed rows, indicative of multiple unfertilized ovules present between normal-looking developing seeds (Fig. 3B, arrowheads). This was in contrast to the immature siliques of wild-type plants, which were filled with round green seeds (Fig. 3C). A comparison of the appearance of unfertilized ovules and aborted embryos in mutant versus wild-type plants demonstrated a significantly higher frequency of defective siliques in the mutant plants compared with the wild type (Table I). The semisterility detected in *sig5* mutant plants was not found in all siliques of the heterozygous *sig5-2* lines. Moreover, within a silique exhibiting defective seeds or unfertilized ovules, the percentage of embryo or fertilization defects was not consistently 25%, as would be expected for a recessive lethal gene defect, but was higher or lower, depending on the silique. In each mutant heterozygous plant, some effected siliques contained up to 50% unfertilized ovules and could be identified without dissection because they were visibly shorter than other siliques on the plant (data not shown).

These observations indicate that the semisterility of *sig5* mutant plants is leaky. In this respect, the *sig5* lines resemble mutants in the Arabidopsis *PROLIFERA* gene. A *prolifera* gene trap line exhibited homozygous embryo lethality, with heterozygous siliques containing a few unfertilized ovules as well as seeds with aborted embryos (Springer et al., 1995). It is proposed that the *prolifera* silique phenotype is variable due to a varying requirement for maternal stores of the *PROLIFERA* RNA or protein during different stages of female gametogenesis and zygote development (Springer et al., 2000). The similarity in

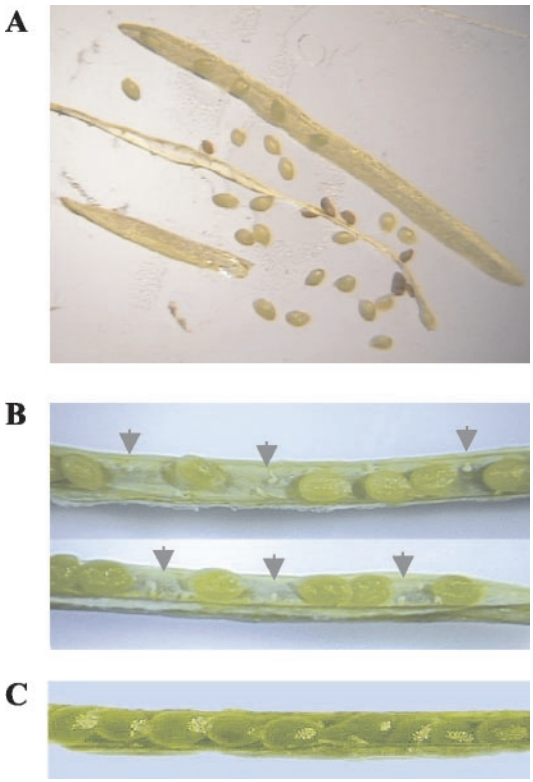

Figure 3. Heterozygous *sig5-2* lines display a seed phenotype. A, Seeds from a mature silique on a self-fertilized heterozygous *sig5-2* plant include brown wrinkled defective seeds and normal round seeds. B, Immature siliques from selfed heterozygous *sig5-2* plants contain gaps in the seed rows, indicative of unfertilized ovules (arrowheads). C, A typical silique from a wild-type plant shows round green seeds with no gaps in the seed row.

phenotype could reflect a similar function for the *SIG5* gene product. Alternatively, the leaky nature of the *sig5* mutant phenotype could be due to partial complementation of the disrupted gene by other members of the Arabidopsis σ-factor gene family.

AtSig5 Accumulates in Reproductive Tissues of Wild-Type Arabidopsis

The silique phenotype of the *sig5-1* and *sig5-2* lines indicated that AtSig5 may be required for gametogenesis and/or early embryogenesis. If this is the case, the *SIG5* gene should be expressed in reproductive tissues in wild-type plants. Previously published analysis of *SIG5* expression detected *SIG5* transcripts in whole 12-d-old seedlings that had been ground up for RNA extraction (Fujiwara et al., 2000) and in rosette leaves of 4-week-old plants (Tsunoyama et al., 2002). Expression of *SIG5* in rosette leaves agrees with the previously proposed role for AtSig5 in chloroplast gene expression. To determine whether AtSig5 also accumulates in reproductive tissues, we did immunoblots on both vegetative and reproductive plant organs using an antibody raised against the NH_2-terminal region of AtSig5. Because the NH_2 terminus of AtSig5 is not conserved with the other Arabidopsis σ-factors (Fujiwara et al., 2000), this antibody is specific for the AtSig5 protein. Immunoblot analysis detected a band of the M_r predicted for AtSig5 (approximately 58 kD) in rosette leaves, stems, immature siliques, and flowers, but not in roots (Fig. 4, right panel, arrow). A lower M_r protein also detected by the antibody in tissues with higher expression levels of AtSig5 (stem and flower samples in Fig. 4) may be a degradation product of the AtSig5 protein. The detection of equivalent levels of a nonspecific band in each lane by preimmune serum (Fig. 4, left panel, asterisk) indicates that equal quantities of protein were loaded for each tissue sample. On the basis of the assumption of equal loading, it is evident that accumulation of AtSig5 protein is even higher in flowers than in leaves (Fig. 4, compare lanes L and F). Therefore, AtSig5 is expressed in reproductive as well as vegetative tissues, consistent with a role in plant reproduction.

SIG5 Transcripts Are Differentially Processed in Leaf and Flower Tissue

It has previously been noted (Fujiwara et al., 2000) that *SIG5* has two potential in-frame AUG translational start codons, one (M1) located within intron 1, and the other (M2) located 75 bp downstream within exon 2 (diagrammed in Fig. 5A; sequence shown by Fujiwara et al. [2000]). Exon 1 is a non-coding exon whose function is currently unknown. Characterization of *SIG5* cDNAs in leaf tissues by 5′-RACE (data not shown) and reverse transcriptase (RT)-PCR (Fig. 5B) indicated that M2 was the initiating Met for AtSig5 in leaves because *SIG5* leaf transcripts do not include intron 1 and therefore do not contain the M1 AUG codon (also diagrammed in Fig. 5A). However, in flower tissue, several products were amplified by RT-PCR from total RNA: a small PCR product similar in size to that amplified from leaf RNA (Fig. 5B, black arrow, compare lanes 3 and 5) and two lower-abundance larger fragments (Fig. 5B, lane 3, white arrowhead). Cloning and sequencing of the small flower RT-PCR product confirmed that it corresponded to the leaf-type RT-PCR product as expected based on its similar apparent size to the leaf product. Interestingly, one of the higher M_r products found in flower RNA but not leaf RNA was similar in size to the genomic PCR product (Fig. 5B, white arrowhead, compare control genomic PCR in lane 1 to lane 3), indicating that it may be an unspliced transcript. Several control reactions were done to demonstrate the absence of contaminating genomic DNA in the RNA preparations (Fig. 5B, lanes 2, 4, 7, and 8). Therefore, this additional RT-PCR product was derived from *SIG5* transcripts present in the flower RNA population. Cloning and sequencing of this RT-PCR product confirmed that it was derived from a splice variant of the *SIG5* transcript, in which intron

Table I. *Frequency of defective siliques in selfed wild-type plants compared with selfed heterozygous sig5–2 plants*

Siliques were opened with a razor blade, and contents were examined under a dissecting microscope. An average of 15 siliques were scored per plant. Note that percentages do not total 100% for each plant genotype because some siliques exhibited both aborted embryos and unfertilized ovules.

Plant Genotype	Siliques with 100% Normal Embryos	Siliques with Aborted Embryos	Siliques with Unfertilized Ovules	Total Siliques
sig5–2/+	119 (53.6%)	34 (15.3%)	77 (34.7%)	222
+/+	179 (78.9%)	15 (6.6%)	33 (14.5%)	227

1 was not spliced out. Because intron 1 contains the M1 Met in frame with the downstream M2 (diagrammed in Fig. 5A), these data suggested that a subset of the AtSig5 protein in flower tissue could be composed of a longer AtSig5 variant, initiated at M1. Because this protein would be longer by only 25 amino acids, it would not be detectably different in size on immunoblots from leaf AtSig5 protein.

AtSig5 NH_2 Termini Initiating at M1 or M2 Are Targeted to Different Organelles in Transient Assays

Computer-based predictions of the intracellular location of AtSig5 differed depending on which NH_2-terminal Met, M1 or M2, was used. The PSORT algorithm (Nakai and Kanehisa, 1992) predicted a location in the chloroplast stroma for leaf AtSig5 initiated at the M2 AUG codon. In contrast, localization in the mitochondrial matrix was predicted for the flower variant of AtSig5 should it initiate at the M1 AUG codon. To determine the targeting properties of NH_2-terminal peptides initiating at M1 or at M2, we fused two derivatives of these sequences to the coding sequence of GFP and determined the subcellular location of the fused GFP in transient expression assays (Fig. 6). Rosette leaves from Arabidopsis were bombarded with the GFP fusion constructs and

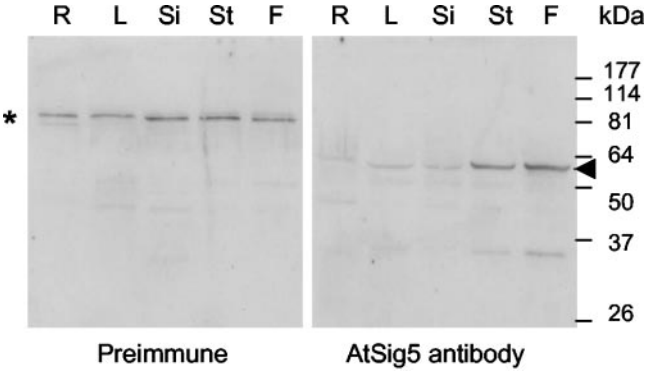

Figure 4. SIG5 is expressed in reproductive and vegetative tissues. Immunoblot of total soluble protein prepared from roots (R), rosette leaves (L), siliques (Si), stems (St), and flowers (F) of 4-week-old wild-type Arabidopsis plants. Immunoblots were probed with preimmune serum (left panel) or anti-AtSig5 antibodies (right panel). A nonspecific cross-reacting band in the preimmune blot (asterisk) serves as a loading control. The AtSig5 cross-reacting band is indicated by an arrow. Migration of M_r markers is shown.

then examined by laser scanning confocal microscopy to determine the subcellular location of GFP fluorescence in cells transiently expressing the fused GFP from a linked plant promoter and 35S enhancer sequences.

As controls to demonstrate chloroplast and mitochondrial targeting patterns, GFP was fused to the chloroplast transit peptide sequence of Rubisco small subunit or to the mitochondrial targeting sequence of the F_0F_1 ATPase γ-subunit (Chang et al., 1999; Beardslee et al., 2002). Strong GFP fluorescence (Fig. 6, left panel, green signal) colocalizing with chloroplast autofluorescence (Fig. 6, right panel, red signal) was detected in leaf cells bombarded with the chloroplast control construct (Fig. 6A). Merging of the two fluorescent signals (Fig. 6, center panel) confirmed that the GFP fluorescence completely overlapped with the chloroplast autofluorescence, as evidenced by the yellow merged signal. In contrast, the GFP fluorescent signal was detected in punctate mitochondria in leaf cells bombarded with the mitochondrial control GFP fusion construct (Fig. 6B). As expected, this mitochondrial signal did not overlap with the chlorophyll autofluorescence. To examine the targeting properties of AtSig5 initiated at M2 we fused the M2 AUG codon followed by 105 downstream AtSig5 codons in frame with GFP. This fusion protein was exclusively targeted to chloroplasts when transiently expressed in leaf cells, as demonstrated by the overlap of the GFP signal with the chlorophyll autofluorescence signal (Fig. 6C). Therefore, proteins translated from *SIG5* transcripts present in leaf tissues are likely localized to chloroplasts, in agreement with the proposed function of AtSig5 as a plastid RNA polymerase specificity factor. To determine whether translation initiation from M1 changed the targeting properties of AtSig5, we fused the 130 amino acids downstream of M1 to GFP. To ensure that the targeting properties we observed were due solely to protein initiated at M1, we mutagenized the M2 translational initiation codon in this construct to AUC, encoding Ile. Transient expression assays demonstrated that the new fusion protein had a GFP fluorescence pattern suggesting dual-targeting, because it exhibited not only the chloroplast localization pattern but also a pattern of punctuate fluorescence, consistent with mitochondrial localization (Fig. 6D). Therefore, AtSig5 protein in flowers could poten-

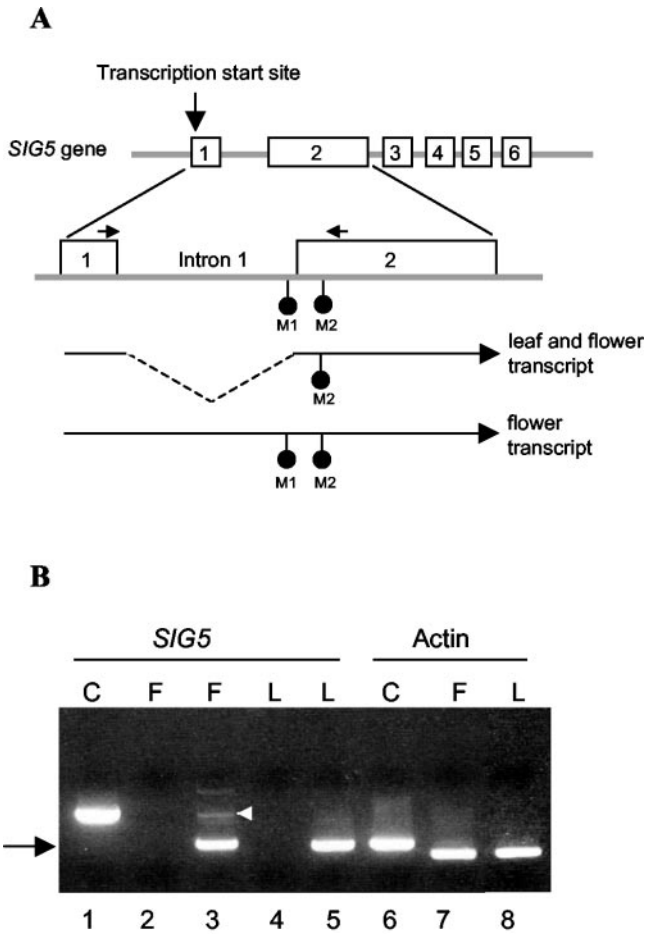

Figure 5. Partially unspliced *SIG5* transcripts accumulate in flower tissue but not in leaf tissue. A, Diagram of the *SIG5* gene indicating exons (numbered boxes). The exon1-intron1-exon2 region is enlarged to indicate the two in-frame Met codons (M1 and M2; ●). Transcripts with or without intron 1 are indicated below. The annealing positions of the intron-spanning primers used in RT-PCR are indicated by horizontal arrows. B, Total RNA prepared from leaf (L, lane 5) or flower (F, lane 3) was subjected to RT-PCR analysis using *SIG5* primers that span intron 1. Both leaf and flower RNA contain spliced transcripts (black arrow). To control for DNA contamination, the reactions were repeated with no reverse transcription step (lane 2 for flower RNA and lane 4 for leaf RNA). Lane 1 (C) contains *SIG5* product amplified from genomic DNA to demonstrate the size of the unspliced *SIG5* PCR product. A product of this size is amplified from flower but not leaf RNA (lane 3, white arrowhead). Lanes 6 through 8 show RT-PCR reactions with actin gene primers to control for the amount of RNA in each sample. Lane 6 is a control PCR using intron-spanning actin gene primers on genomic DNA. The lack of this larger product in lanes 7 and 8 reconfirms the absence of genomic DNA in the flower and leaf RNA preparations.

tially be targeted not only to chloroplasts but also to mitochondria.

DISCUSSION

We report here that disruption of the Arabidopsis *SIG5* gene has a negative impact on plant reproduction and results in the recovery of unfertilized ovules and aborted embryos at a higher frequency than is seen with wild-type self-fertilized plants. Homozygous knockout plants could not be recovered from either of two *sig5* insertional mutant lines, *sig5-1* or *sig5-2*. Preliminary crossing data (not shown), indicate that when *sig5-2* is used as a female parent in crosses to wild-type plants, a higher frequency of siliques with unfertilized ovules is recovered than when the mutant line is used as the male parent. The seed phenotype of *sig5-2* is leaky because individual siliques may contain both unfertilized ovules and aborted embryos. The inability to recover homozygous *sig5-1* or *sig5-2* lines indicates that *SIG5* is an essential gene (as defined by McElver et al. [2001]).

In several respects, the *sig5* knockout plants resemble Arabidopsis lines containing a disruption in the *PROLIFERA* gene. Mutant lines heterozygous for a *prolifera* disruption allele were semisterile, generating siliques with a few unfertilized ovules and aborted embryos (Springer et al., 1995). Use of heterozygous *prolifera* disruption lines as the female parents in crosses to wild type generated a higher proportion of siliques with unfertilized ovules than was found in the reciprocal cross (Springer et al., 2000). Homozygous *prolifera* mutants could not be recovered from selfed heterozygous lines. It has been proposed that the *PROLIFERA* RNA is required maternally for early Arabidopsis development (Springer et al., 2000). The similar overall phenotype of the *sig5* disruption lines indicates that the *SIG5* gene product may also be required for an early step in plant reproduction.

The leaky nature of the *sig5* phenotype may reflect partial compensation for the missing *SIG5* function by another Arabidopsis σ-factor. For example, previous analyses of Arabidopsis *SIG1* and *SIG2* promoter fusions to the *uidA* reporter gene in stably transformed lines revealed GUS expression in siliques, although not in seeds (Kanamaru et al., 1999). GUS staining was seen throughout mature siliques, and at the top and base of smaller or immature siliques. We note however, that a T-DNA insertion in the *SIG2* gene did not generate the same phenotype as the *sig5* mutant alleles reported here, and plants homozygous for the *sig2* disruption could be recovered (Shirano et al., 2000). Previously published examinations of the other Arabidopsis σ-factor genes focused primarily on expression in leaf and root tissues, and on the effects of light in inducing *SIG* gene expression. Therefore, the potential accumulation of the other Arabidopsis σ-factors AtSig3, AtSig4, and AtSig6 in reproductive tissues remains to be examined.

Because disruption of the *SIG5* gene results in a reproductive defect, it was of interest to determine whether AtSig5 is a chloroplast-targeted protein. There is ample precedent for nuclear genes encoding chloroplast-targeted proteins that are essential for embryogenesis or seed development (for review, see McElver et al., 2001). Alternatively, mitochondrial-targeted or peroxisomal-targeted proteins are also

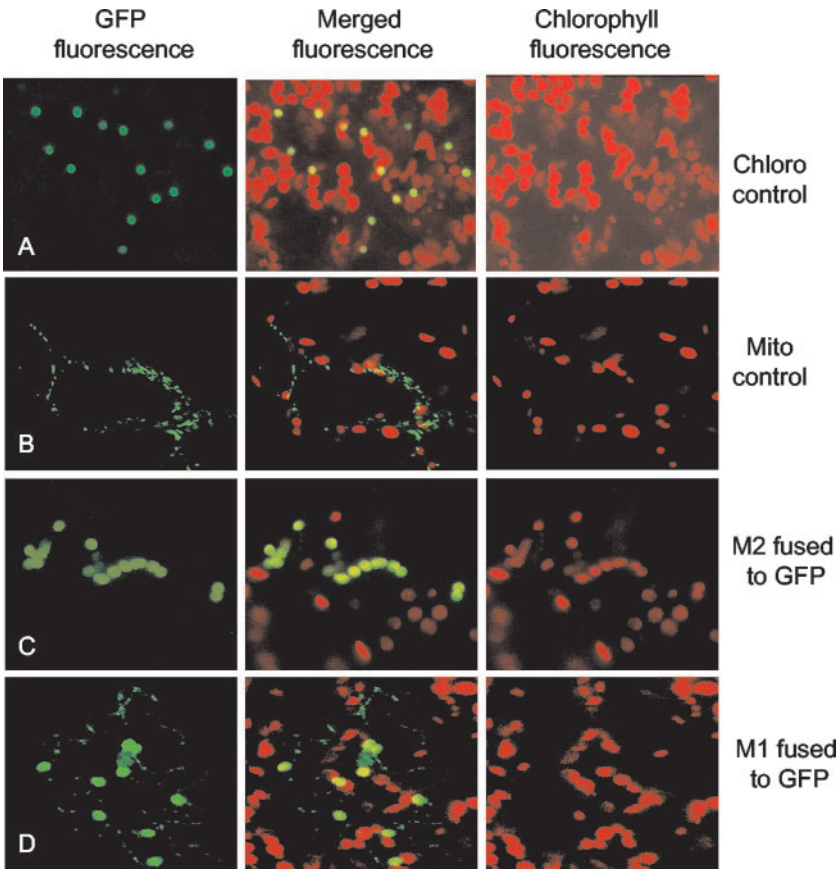

Figure 6. Differential targeting of AtSig5-GFP fusion proteins translated from the M1 or M2 Met. NH_2-terminal AtSig5 sequences initiated from the M1 (row D) or M2 (row C) Met were fused to GFP and were transiently expressed in Arabidopsis leaf cells. Expression of control constructs with GFP targeted to chloroplasts (row A) or mitochondria (row B) is shown. In each part of the figure, the left panel represents the GFP fluorescent signal, the right panel represents the chlorophyll autofluorescent signal of chloroplasts, and the center panel is the GFP signal merged with the chlorophyll autofluorescence.

known to function in aspects of plant reproduction (Tsugeki et al., 1996; Skinner et al., 2001; Hu et al., 2002). Because the amino acid sequence of AtSig5 bears a strong resemblance to that of eubacterial σ-factors, it had been proposed (but not experimentally demonstrated) that AtSig5 is a chloroplast-localized protein that assembles with the eubacterial-like PEP RNA polymerase in plastids. However, the recent unexpected discovery of a maize σ-factor targeted to mitochondria as well as to plastids has emphasized how important it is to generate experimental evidence for plant σ-factor subcellular localization (Beardslee et al., 2002). We used GFP fusion proteins to demonstrate that, in transient expression assays in rosette leaves, an AtSig5 protein initiated at the upstream in-frame M1 Met could be targeted to both mitochondria and plastids, whereas the shorter protein initiated at M2 was exclusively plastid targeted. Flower RNA contains a small, but detectable, amount of *SIG5* transcripts containing the M1 codon within an unspliced intron, whereas in leaf transcripts, this coding sequence is spliced out. A similar mechanism of alternative splicing giving rise to transcripts with alternative targeting potentials has been found for an Arabidopsis arginyl tRNA synthetase gene (Small et al., 1998). Alternative splicing has also been demonstrated for the Arabidopsis *SIG4* gene (Fujiwara et al., 2000). In this case, the alternate transcript contains a 14-base upstream shift of the intron 2 splice donor site and would result in a protein missing all conserved domains C-terminal to conserved region 2.4. The alternatively spliced transcript accumulates in all tissues examined, however its functional significance is unknown (Fujiwara et al., 2000).

The level of the M1-unspliced *SIG5* transcript variant that we detected in flower tissues was so low that it would not be expected to produce very much protein. Therefore, the physiological significance of a potentially mitochondrial-targeted AtSig5 remains to be determined, especially because there is no evidence for a eubacterial-type RNA polymerase in plant mitochondria. However, previous studies on a maize σ-factor, ZmSig2B, revealed that this protein accumulates in leaf mitochondria as well as in chloroplasts and is copurified through several steps with the mitochondrial RNA polymerase (Beardslee et al., 2002). Therefore, preliminary evidence in two plant species supports a mitochondrial location for at least one member of the σ-factor family. Additional experiments to confirm the in vivo presence of AtSig5 in flower mitochondria will help to establish whether or not differential splicing is a physiological mechanism used to control the targeting properties of this Arabidopsis σ-factor.

The data presented here reveal that AtSig5 participates directly or indirectly in the generation of fer-

tilized ovules and/or production of viable plant embryos. However, this does not rule out an additional role for the σ-factor in leaf chloroplast transcription. Our transient expression data indicate that the AtSig5 protein expressed in leaf cells is localized to chloroplasts. Studies on accumulation of *SIG5* transcripts in rosette leaves in response to different fluences and wavelengths of light revealed that *SIG5* expression is specifically induced by blue light and led to the proposal that AtSig5 recognizes the *psbD/psbC* blue-light-regulated promoter (Tsunoyama et al., 2002). Therefore, we would anticipate that, in plants lacking *SIG5* expression, *psbD* transcript levels should be dramatically reduced. Because it has not proven possible to recover homozygous *sig5-1* or *sig5-2* lines, we were not able to test this prediction. Preliminary data on transgenic Arabidopsis lines engineered to overexpress a self-complementary *SIG5* hairpin transcript indicate that RNAi plants exhibiting the same silique phenotype as the *sig5* knockout lines can be recovered (J. Yao and L.A. Allison, unpublished data). Therefore, analysis of chloroplast transcripts in plants with reduced *SIG5* transcript levels may provide clues as to AtSig5 function in leaves. Another strategy to uncover potential leaf phenotypes of plants homozygous for a *SIG5* gene disruption would be to complement the *sig5-2* lines with a *SIG5* cDNA under control of a seed-specific promoter. This approach proved useful in analysis of an insertional mutant allele of *EMB506*, an Arabidopsis gene that encodes a chloroplast-targeted protein with functions both in leaf tissues and in embryogenesis (Despres et al., 2001).

MATERIALS AND METHODS

Upon request, all novel materials described in this publication will be made available in a timely manner for noncommercial research purposes, subject to the requisite permission from any third-party owners of all or parts of the material. Obtaining any permissions will be the responsibility of the requestor.

Identification of *sig5-1* and *sig5-2*

Isolation of plants with the *sig5* T-DNA insertion allele designated as *sig5-1* was done by screening the T-DNA insertion collection generated by the Arabidopsis knockout facility at the University of Wisconsin Biotechnology Center (Madison). Two *SIG5*-specific primers, UNL257 (5′-TCACATTGTTATTAATCTGTGGACACCAT-3′) and UNL 258 (5′-CCTTGACTTCATGTTTCCTAACCATTTCT-3′), were submitted to the Wisconsin facility for first-round PCR screening with 30 DNA super pools. A positive product was identified in the amplification reaction from super pool number 4 with the *SIG5*-specific primer UNL258 and T-DNA left border primer JL-202 (5′-CATTTTATAATA ACGCTGCGGACATCTAC-3′). The product was sequenced by the Genomics Core Research Facility of University of Nebraska (Lincoln) and confirmed the presence of a T-DNA insertion in the *SIG5* gene-coding sequence. Second-round PCR screens performed at the Wisconsin facility identified a positive subpool. Seeds from plants making up the subpool were obtained from the Arabidopsis Biological Resource Center (Columbus, OH) and were analyzed further to finally identify *sig5-1* heterozygous plants by the same PCR-screening procedure.

The second T-DNA insertional allele, *sig5-2*, was identified by screening an Arabidopsis (Columbia ecotype) T-DNA insertion collection generated by the Syngenta company and available for screening on a collaborative basis from the Torrey Mesa Research Institute (San Diego). The full-length *SIG5* sequence was used to query the SAIL (formerly called GARLIC) database of Arabidopsis sequences flanking T-DNA left borders. One sequence from the database (Garlic_1232_H11) aligned with *SIG5* sequence starting at nucleotide number 381 of exon 2. Seeds from this knockout line were obtained from the company and renamed *sig5-2* plants. The presence of the insertional allele was confirmed by sequencing the PCR product from *sig5-2* genomic DNA generated by a *SIG5*-specific primer, UNL257, and a T-DNA left border primer (5′-TAGCATCTGAATTTCATAACCAATCTCGATACAC-3′).

Production of AtSig5 Polyclonal Antibodies

The cDNA sequences encoding amino acids 140 to 258 of the AtSig5 NH_2 terminus starting from the M2 AUG codon were amplified by PCR, using primers with restriction sites for subcloning of the products. Amplification primers were UNL207 (5′-ttggatccAGTGGTGAGGAGAAGAAAGTG-3′) and UNL179 (5′-ttctcgag TATTTCAGCTTCCCTAGGC-3′) where gene-specific sequences are indicated by uppercase letters. The amplified fragment was subcloned into vector pET28A (Novagen, Madison, WI) using the engineered restriction sites, *Bam*HI and *Xho*I, to place the coding regions in frame with the vector NH_2-terminal hexa-His sequence. Insert sequences were confirmed by sequencing. Recombinant protein expression was induced in BL21 (DE3) cells (Novagen) by addition of isopropylthio-β-galactoside to a final concentration of 1 mM. His-tagged proteins from the cell lysate were bound to Talon resin (BD Biosciences Clontech, Palo Alto, CA) in denaturing buffer and purified according to the manufacturer's protocols. Rabbit polyclonal antisera were prepared against purified recombinant protein by Bethyl Research Labs (Montgomery, TX).

Immunoblot Analysis

Wild-type Arabidopsis plants (ecotype Columbia) were grown at 23°C in cycling light (16 h of light/8 h of dark) for 35 d. Tissues were harvested and immediately ground under liquid nitrogen. Tissue powder were extracted using protein extraction buffer (0.5 M Suc, 0.1% [w/v] ascorbic acid, 0.1% [w/v] Cys-HCl, and 0.1 M Tris-HCl). Crude extracts were centrifuged to pellet insoluble debris. Supernatant protein extract was assayed for protein concentration using the Bio-Rad protein assay system (Bio-Rad, Hercules, CA). Protein extracts containing 5 μg of total protein were separated with 10% (w/v) SDS-PAGE on minigels and transferred to Immobilon-P polyvinylidene difluoride membranes (Millipore, Bedford, MA) with a Bio-Rad mini-transblot apparatus in Tris-Gly transfer buffer. After transfer, membranes were blocked overnight in 5% (w/v) nonfat dry milk and then incubated for 1 h in primary antibody (IgG fraction of rabbit anti-AtSig5) diluted 1:2,000. The secondary antibody incubation was performed for 1 h with donkey anti-rabbit horseradish peroxidase (Amersham Biosciences, Piscataway, NJ) diluted 1:50,000. Secondary antibody detection was performed with an ECL-Plus kit using manufacturer-supplied protocols (Amersham Biosciences).

Determination of *SIG5* Transcript 5′ Ends with RACE

For total RNA isolation, leaves were harvested from 4-week-old wild-type Arabidopsis (Columbia ecotype) plants and kept in liquid nitrogen. One hundred milligrams of leaf tissue was ground thoroughly under liquid nitrogen and was used to isolate total RNA with the RNeasy Plant Mini Kit from Qiagen (Valencia, CA). RNA recovery and quality was analyzed by agarose gel electrophoresis. Ten micrograms of the RNA sample was used to determine *SIG5* transcript 5′ ends with the FirstChoice RLM-RACE kit from Ambion (Austin, TX). Nested PCR primers designed according to the manufacturer's protocol were: UNL143 (5′-CAACATCATCCTCGACTAATG-3′) for outer PCR, and UNL144 (5′-CTGGATCCGGTTGCAGGACTAAGC-3′) for inner PCR. The product of the inner PCR reaction was sequenced for 5′-end determination.

Subcellular Localization of Transiently Expressed AtSig5-GFP Fusion Protein

Sequences encoding the NH_2-terminal 131 amino acids of AtSig5 starting from M1 and 106 amino acids starting from M2 were PCR-amplified and

subcloned into the *Bam*HI site of GFP expression vector 35S-C4PPDK-sGFP(S65T) (Chiu et al., 1996). Primers used were: UNL339 (5′-GTGGATCCATGAGATATGTTTCTGCTTG-3′) and UNL147 (5′-CGGGATCCGCTTTCAAGTCTTGCAGCAGCTT-3′) for the M1 construct and UNL284 (5′-TTGGATCCATGGGAGTTGTGTCTATTTC-3′) and UNL147 for the M2 construct. The inserts were sequenced to verify the insert orientation and the lack of PCR-based sequence errors. All control GFP vectors and transient assay protocols with leaf protoplasts were described previously (Chang et al., 1999; Beardslee et al., 2002). For transient assays in intact Arabidopsis leaf tissues, rosette leaves of 10-d-old greenhouse-grown seedlings (Columbia ecotype) were excised and arranged abaxial side up in a filled circle pattern on hormone-free Murashige and Skoog medium (Murashige and Skoog, 1962). Plates were bombarded with 0.75-μm tungsten particles coated with 2 μg of vector DNA per shot. Bombarded leaf samples were incubated at 23°C in the dark for 24 h. Bombarded leaf pieces were mounted in water on glass slides for microscopy. Detection of GFP fluorescence and chlorophyll autofluorescence by laser scanning confocal microscopy was performed as described by Beardslee et al. (2002).

ACKNOWLEDGMENTS

We thank Dr. Joe Zhou and Christian Elowski (University of Nebraska Center for Biotechnology Microscopy Core Facility) for assistance in confocal microscopy analysis, and Dr. Tom Clemente, Dr. Julie Stone, and Dr. Sally Mackenzie (all at University of Nebraska, Lincoln) for helpful discussion.

Received November 15, 2002; returned for revision January 13, 2003; accepted February 23, 2003.

LITERATURE CITED

Allison LA (2000) The role of sigma factors in plastid transcription. Biochimie **82:** 537–548

Allison LA, Maliga P (1996) Deletion of *rpoB* reveals a second distinct transcription system in plastids of higher plants. EMBO J **15:** 2802–2809

Beardslee T, Chowdhury SR, Chang CC, Jaiswal P, Lerbs-Mache S, Stern D, Allison LA (2002) A nucleus-encoded maize protein with sigma factor activity accumulates in both chloroplasts and mitochondria. Plant J **31:** 199–209

Burgess RR, Anthony L (2001) How sigma docks to RNA polymerase and what sigma does. Curr Opin Microbiol **4:** 126–131

Campbell EA, Muzzin O, Chlenov M, Sun JL, Olson CA, Weinman O, Trester-Zedlitz ML, Darst SA (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. Mol Cell **9:** 527–539

Chang CC, Sheen J, Bligny M, Niwa Y, Lerbs-Mache S, Stern DB (1999) Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. Plant Cell **11:** 911–926

Chiu WL, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J (1996) Engineered GFP as a vital reporter in plants. Curr Biol **6:** 325–330

Despres B, Delseny M, Devic M (2001) Partial complementation of embryo defective mutations: a general strategy to elucidate gene function. Plant J **27:** 149–159

Fujiwara M, Nagashima A, Kanamaru K, Tanaka K, Takahashi H (2000) Three new nuclear genes, *sigD*, *sigE* and *sigF*, encoding putative plastid RNA polymerase sigma factors in *Arabidopsis thaliana*. FEBS Lett **481:** 47–52

Gross CA, Chan C, Dombroski A, Gruber T, Sharp M, Tupy J, Young B (1998) The functional and regulatory roles of sigma factors in transcription. Cold Spring Harbor Symp Quant Biol **63:** 141–155

Hajdukiewicz PTJ, Allison LA, Maliga P (1997) Transcription by two distinct RNA polymerases is a general regulatory mechanism of plastid gene expression in higher plants. EMBO J **16:** 4041–4048

Hakimi M-A, Privat I, Valay J-G, Lerbs-Mache S (2000) Evolutionary conservation of C-terminal domains of primary sigma70-type transcription factors between plants and bacteria. J Biol Chem **275:** 9215–9221

Helmann JD, Chamberlin MJ (1988) Structure and function of bacterial sigma factors. Annu Rev Biochem **57:** 839–872

Hess WR, Börner T (1999) Organellar RNA polymerases of higher plants. Int Rev Cytol **190:** 1–59

Hu J, Aguirre M, Peto C, Alonso J, Ecker J, Chory J (2002) A role for peroxisomes in photomorphogenesis and development of Arabidopsis. Science **297:** 405–409

Isono K, Shimuzu M, Yoshimoto K, Niwa Y, Satoh K, Yokota A, Kobayashi H (1997) Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma^{70} factors of bacterial RNA polymerases in *Arabidopsis thaliana*. Proc Natl Acad Sci USA **94:** 14948–14953

Kanamaru K, Fujiwara M, Seki M, Katagiri T, Nakamura M, Mochizuki N, Nagatani A, Shinozaki K, Tanaka K, Takahashi H (1999) Plastidic RNA polymerase sigma factors in Arabidopsis. Plant Cell Physiol **40:** 832–842

Kanamaru K, Nagashima A, Fujiwara M, Shimada H, Shirano Y, Nakabayashi K, Shibata D, Tanaka K, Takahashi H (2001) An Arabidopsis sigma factor (SIG2)-dependent expression of plastid-encoded tRNAs in chloroplasts. Plant Cell Physiol **42:** 1034–1043

Kestermann M, Neukirchen S, Kloppstech K, Link G (1998) Sequence and expression characteristics of a nuclear-encoded chloroplast sigma factor from mustard (*Sinapis alba*). Nucleic Acids Res **26:** 2747–2753

Krysan PJ, Young JK, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. Plant Cell **11:** 2283–2290

Lahiri SD, Allison LA (2000) Complimentary expression of two plastid-localized sigma-like factors in maize. Plant Physiol **123:** 883–894

Lahiri SD, Yao J, McCumbers C, Allison LA (1999) Tissue-specific and light-dependent expression within a family of nuclear-encoded sigma-like factors from *Zea mays*. Mol Cell Biol Res Commun **1:** 14–20

Liere K, Maliga P (2001) Plastid RNA polymerases in higher plants. *In* E-M Aro, B Andersson, eds, Regulation of Photosynthesis. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 29–29

McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA et al. (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. Genetics **159:** 1751–1763

Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science **296:** 1285–1290

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant **15:** 473–479

Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics **14:** 897–911

Satoh J, Baba K, Nakahira Y, Shiina T, Toyoshima Y (1999) Developmental stage-specific multi-subunit RNA polymerases (PEP) in wheat. Plant J **18:** 407–415

Serino G, Maliga P (1998) RNA polymerase subunits encoded by the plastid rpo genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol **117:** 1165–1170

Shirano Y, Shimada H, Kanamaru K, Fujiwar M, Tanaka K, Takahashi H, Unno K, Sato S, Tabata S, Hayashi H et al. (2000) Chloroplast development in *Arabidopsis thaliana* requires the nuclear-encoded transcription factor Sigma B. FEBS Lett **485:** 178–182

Skinner DJ, Baker SC, Meister RJ, Broadhvest J, Schneitz K, Gasser CS (2001) The Arabidopsis *HUELLENLOS* gene, which is essential for normal ovule development, encodes a mitochondrial ribosomal protein. Plant Cell **13:** 2719–2730

Small I, Wintz H, Akashi K, Mireau H (1998) Two birds with one stone: genes that encode products targeted to two or more compartments. Plant Mol Biol **38:** 265–277

Springer PS, Holding DR, Groover A, Yordan C, Martienssen RA (2000) The essential Mcm7 protein PROLIFERA is localized to the nucleus of dividing cells during the G1 phase and is required maternally for early Arabidopsis development. Development **127:** 1815–1822

Springer PS, McCombie WR, Sundaresan V, Martienssen RA (1995) Gene trap tagging of PROLIFERA, an essential MCM2-3-5-like gene in Arabidopsis. Science **268:** 877–880

Thum KE, Kim M, Christopher DA, Mullet JE (2001) Cryptochrome 1, cryptochrome 2, and phytochrome a co-activate the chloroplast psbD blue light-responsive promoter. Plant Cell **13:** 2747–2760

Tsugeki R, Kochieva EZ, Fedoroff NV (1996) A transposon insertion in the Arabidopsis *SSR16* gene causes an embryo-defective lethal mutation. Plant J **10:** 479–489

Tsunoyama Y, Morikawa K, Shiina T, Toyoshima Y (2002) Blue light specific and differential expression of a plastid sigma factor, Sig5 in Arabidopsis thaliana. FEBS Lett **516:** 225–228