



Title	The regulatory function of the upstream sequence of the beta-conglycinin alpha subunit gene in seed-specific transcription is associated with the presence of the RY sequence
Author(s)	Yoshino, Michiko; Nagamatsu, Atsushi; Tsutsumi, Ken-ichi; Kanazawa, Akira
Citation	Genes & Genetic Systems, 81(2), 135-141 https://doi.org/10.1266/ggs.81.135
Issue Date	2006-04
Doc URL	http://hdl.handle.net/2115/14641
Type	article (author version)
File Information	GGs2006-81-2.pdf



[Instructions for use](#)

To *Genes & Genetic Systems* (short communication)

The regulatory function of the upstream sequence of the β -conglycinin α subunit gene in seed-specific transcription is associated with the presence of the RY sequence

Michiko Yoshino¹, Atsushi Nagamatsu¹, Ken-ichi Tsutsumi² and Akira Kanazawa^{*,1}

¹*Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*

²*Cryobiosystem Research Center, Faculty of Agriculture, Iwate University, Ueda 020-8550, Japan*

*Corresponding author. E-mail: kanazawa@res.agr.hokudai.ac.jp

Note: Nucleotide sequence data reported is available in the DDBJ/EMBL/GenBank databases under the accession number AB237643.

Running head: Promoter of the β -conglycinin α subunit gene

Key words: β -conglycinin, promoter, RY sequence, seed storage protein, soybean

Correspondence to: Akira Kanazawa

Graduate School of Agriculture, Hokkaido University

Kita-ku, Kita 9, Nishi 9, Sapporo 060-8589, Japan

Fax: +81-11-706-4933 / Phone: +81-11-706-4933

e-mail: kanazawa@res.agr.hokudai.ac.jp

β -conglycinin, a major component of seed-storage proteins in soybean, comprises three subunits: α , α' , and β . Expression of these genes is spatially regulated in a stringent manner and occurs during seed development. To understand the mechanisms that control expression of the α subunit gene, we analyzed the nucleotide sequence of the 2.9-kb region upstream of the gene. The upstream sequence up to -1357 or a series of its 5'-deleted derivatives was fused to the β -glucuronidase (GUS) gene. These reporter gene constructs were introduced into *Arabidopsis thaliana* plants via *Agrobacterium*-mediated gene transfer. Prominent GUS activity was detected in developing seeds of the T3 generation when 245 bp or longer sequences of the upstream region were fused to the GUS gene. We found a clear association of decreased GUS activity with a stepwise deletion of a region containing the RY sequence from the original construct. These results are consistent with the notion that multiple sequence elements including the RY sequences are involved in the seed-specific transcriptional activation of the β -conglycinin α subunit gene in soybean.

β -conglycinin, a major component of seed-storage proteins in soybean (*Glycine max*), is a trimeric protein composed of various combinations of three subunits: α , α' , and β . These proteins occupy up to 30% of the total seed proteins in soybean (Thanh and Shibasaki, 1976, 1978; Higgins, 1984). The genes for these proteins appear to be transcribed in a coordinated but not identical manner during seed development. Regulatory elements involved in the transcription of the α' and β subunit genes have been studied extensively by means of a reporter gene assay in transgenic plants and a binding assay using nuclear extracts (Chen et al., 1986, 1988, 1989; Allen et al., 1989; Lessard et al., 1991; Chamberland et al., 1992; Fujiwara and Beachy, 1994). However, despite its importance in terms of nutritional value and allergenic activity in humans, little is known about the regulatory elements involved in the transcriptional control of the α subunit gene, mainly because the genomic DNA sequence of the gene has not been known.

Based on hybridization analyses, Harada et al. (1989) identified a family of 15 distinct β -conglycinin genes, designated CG-1 to CG-15, several of which are tandemly arrayed in limited chromosomal regions. Based on the sizes of mRNA detected by Northern blot analyses, CG-2 and CG-3 in the chromosomal “region A” have been considered to be the α subunit or α' subunit genes (Harada et al., 1989). By sequencing genomic DNA clones, we demonstrated that CG-3 actually encodes the α subunit protein (Yoshino et al., 2001). We also found that CG-2 encodes a protein that is very similar to the α subunit by analyzing mutant lines of soybean that exhibit quantitative variation of the subunit protein (Yoshino et al., 2002).

Here, we present the first report on the upstream sequence of the α subunit gene. We describe possible regulatory elements for transcription of the gene based on the results of reporter assays in transgenic *Arabidopsis thaliana* plants.

The materials and methods used in this study are as follows. For nucleotide sequencing, the 7.6 kb *Eco*RI fragment of the genomic DNA clone carrying the α subunit gene (see Yoshino et al., 2001) was subcloned into a plasmid vector pBluescript SK+

(Stratagene), and sequential deletions were introduced using exonuclease III from either end of the fragment. A series of deletion clones were used for sequence analysis. Nucleotide sequencing was carried out on both strands of DNA using ALF express DNA sequencer (Pharmacia Biotech) and Thermo Sequenase Fluorescent Labelled Primer Cycle Sequencing kit (Amersham). Possible regulatory DNA elements present in the 5' upstream sequence of the gene were identified using Plant Cis-Acting Regulatory DNA Elements (PLACE) database (<http://www.dna.affrc.go.jp/PLACE/>) (Higo et al., 1999). The upstream DNA sequences of the α subunit gene from positions -1357, -867, -545, -402, -245, -161, and -73 to +27 (positions are numbered relative to the transcription start site) were amplified by PCR using plasmid DNA containing the gene (Yoshino et al., 2001) as a template. Primers for PCR are as follows:

5'-AAGCTTTGCTTGGATTTGGACCAGAC-3' (-1357F),
5'-AAGCTTCGAACTACGAGTTATGAAGTG-3' (-867F),
5'-AAGCTTGTACTCACCAAGGTGCAATC-3' (-545F),
5'-AAGCTTGTCTCTTGGATCATGCATGC-3' (-402F),
5'-AAGCTTATGCCATGCACATCAACACG-3' (-245F),
5'-AAGCTTACTGCCTATGCGACTCTAAC-3' (-161F),
5'-AAGCTTCCATGCATGCAAGTTAACAAG-3' (-73F), and
5'-TCTAGATAGGATATTGAACTAGTTCTCG-3' (+27R). The first six nucleotides of the forward primers (termed 'F') provide a *Hind*III site and those of the reverse primer (termed 'R') provide an *Xba*I site, which were used in subsequent plasmid construction. The PCR cycling conditions were: 94°C for 30 sec, 54°C for 30 sec, and 72°C for 1 min. This cycle was repeated 29 times, and the reaction mixture was then further incubated at 72°C for 10 min. After cloning the PCR products into the pGEM-T Easy vector (Clontech), the *Hind*III-*Xba*I fragment of the plasmid containing the upstream sequence was force-cloned across the *Hind*III and *Xba*I sites located in the upstream of the GUS gene on the pBI101 plasmid vector (Clontech). An *Agrobacterium tumefaciens* strain

GV3101 harboring the plasmid DNA that contains the α subunit gene promoter-GUS gene fusions was used for transformation of *A. thaliana* plants (ecotype Columbia) by a vacuum infiltration procedure (Bechtold et al., 1993). *A. thaliana* transformants (T1 plants) were selected according to resistance to kanamycin (50 mg/l), and the T2 seeds were collected. The T2 seeds were sown on soil and plants were grown under 16-hr light and 8-hr dark conditions at 24°C. GUS activity was measured essentially as described by Jefferson et al. (1987). Siliques or leaves of *A. thaliana* plants were ground with mortar and pestle in the presence of extraction buffer (Jefferson et al., 1987). After centrifuging microtubes that contain the suspensions at 10,000 $\times g$ for 5 min at 4°C, the supernatant was subjected to GUS enzyme assay using 4-methylumbelliferone (4-MU) as a substrate. The level of 4-MU was measured 0 min, 30 min, and 60 min after commencing the reaction using a plate reader (ARVO MX-1, Amersham). The protein concentration was quantified using Bradford Reagent and BSA as standard (Bradford, 1976). The activity was calculated as pmol 4-MU $\text{min}^{-1} \text{mg}^{-1}$ of protein. Histochemical staining of GUS activity was carried out as described by Kosugi et al. (1990). Surface-sterilized T2 seeds of *A. thaliana* were sown on half a concentration of standard MS medium (Murashige and Skoog, 1962) that contained 30 g/l sucrose. The pH of the medium was adjusted to 5.8 before autoclaving. Media were solidified with 0.8% (w/v) agar. Plants were grown on this medium for 10 days. Embryos were dissected from T3 seeds 10 days after flowering. The young plants or embryos were stained in GUS staining solution at 37°C for overnight. After staining, the plants or embryos were immersed in 95% ethanol to remove chlorophyll.

We previously reported the genomic DNA sequence of the α subunit gene including a portion of the sequence corresponding to the proximal promoter of the gene (Yoshino et al., 2001). A primer extension analysis revealed a major transcription start site located in the 56-bp upstream of the ATG codon (Yoshino et al., 2001). In the present study, we analyzed regions farther upstream of the gene by sequencing the genomic clone

that contains the α subunit gene (Fig. 1). A computer-based search of the database revealed that this region contains multiple elements, including those presumed to be involved in transcriptional control of genes during seed development. These include 2 SEF1-binding sites [ATATTTA(T/A)(A/T)], 8 SEF4-binding sites [(A/G)TTTTT(A/G)], 2 core sequences [AACCCA] of the SEF3-binding site [AACCCA---AACCCA] (see Allen et al., 1989), 6 RY sequences [CATGCA, CATGCA(C/T), or CATGCATG] (Dickinson et al., 1988), 11 E-box [CANNTG] (Kawagoe and Murai, 1992), 1 G-box [CACGTG] (Kawagoe et al., 1994), and 1 Dc3 promoter-binding factor (DPBF) binding site [ACACNNG] (Kim et al., 1997). The SEF4-binding sites were located more than 1.2 kb upstream from the transcription start site. The E-box sequences were mostly detected in the region between -0.4 kb and -1.9 kb. The RY sequences were found within several hundred bp upstream from the transcription start site but was not present farther upstream.

To examine seed-specific promoter activity, a DNA fragment that covered a region -1357 to +27 was inserted upstream of the GUS reporter gene on the plasmid pBI101 vector. The DNA construct was introduced into *A. thaliana* plants by an *Agrobacterium*-mediated gene transfer. In plants of the T2 generation, GUS activity was detected in siliques that contained T3 seeds at increasing levels depending on the stage of development, while no GUS activity was detected in leaves (Fig. 2A). Strong GUS expression was detected in seeds within siliques by histochemical staining (Fig. 2B). In seeds, GUS activity was detected throughout the embryo (Fig. 2C), but not in the remaining portions after the embryo was removed (Fig. 2D). No GUS activity was detected in vegetative tissues in young plants except for cotyledons in which GUS protein that was produced during seed development can remain (Fig. 2E). These results indicate that spatial regulation of the transcription of the α subunit gene is maintained in transgenic *Arabidopsis* plants and that transgenic *Arabidopsis* plants are useful for analyzing the upstream regulatory elements of the gene as previously shown for the

promoters of seed-storage protein genes, such as the β -conglycinin α' subunit and β subunit genes of soybean (Hirai et al., 1994; Naito et al., 1994) and the β -phaseolin gene of bean (*Phaseolus vulgaris*; Chandrasekharan et al., 2003).

To identify regions that are necessary for the transcriptional control of the gene, the upstream sequence, with a series of deletions from the 5' end, was fused to the GUS gene, and the resultant reporter constructs were introduced into *A. thaliana* plants. GUS activity in siliques was measured when the siliques reached a fully mature stage. For each promoter-reporter construct, 20 siliques containing T3 seeds were harvested from six individual T2 plants of a transformed line and used for quantification of GUS activity. This was repeated for five independently transformed lines for each construct.

Substantial GUS activity was detected when sequences covering up to -245 (construct -245) or farther upstream were located upstream of the GUS gene, whereas no major GUS activity was observed when constructs -161 or -73 were used for the analysis (Fig. 3). The GUS activity basically decreased with a decrease in the length of the upstream sequence, except that construct -545 conferred lower activity than construct -402. These results indicate that sequences that activate GUS expression in seeds are located at positions between -1357 and -545, -402 and -245, and -245 and -161. This also suggests that a negative regulatory element is located at positions between -545 and -402. No major GUS activity was detected in leaf tissues for all reporter constructs (data not shown).

Early studies on the α' and β subunit genes of β -conglycinin suggested an involvement of protein factors, designated SEF3 and/or SEF4, and their binding sites located within several hundred bp upstream of these genes in the control of seed-specific gene expression (Allen et al., 1989; Lessard et al., 1991). Later studies, however, did not support this hypothesis, because mutations in the SEF3- and SEF4-binding sites affecting the binding of these protein factors had little effect on the activity of the α' subunit promoter in transgenic tobacco plants (Fujiwara and Beachy, 1994). This was also the

case for the SEF1-binding site, the deletion of which did not cause a great decrease in transcriptional activity of the α' subunit gene promoter (Chen et al., 1986). In the α subunit gene, SEF3- and SEF4-binding sites were not present within 1.2 kb upstream of the transcription start sites, while the reporter constructs actually conferred seed-specific transcription without carrying SEF3- and SEF4-binding sites. The SEF4-binding sites were mostly detected within 1.9 to 2.6 kb upstream of the transcription start site in the α subunit gene, which is farther upstream than the locations in the α' and β subunit genes. Considering that α' and α subunit genes are controlled in a similar manner during seed development, lack of SEF3- and SEF4-binding sites in the proximal region of the α subunit gene promoter is consistent with the postulate that these sequences are not directly involved in seed-specific transcriptional control.

A DNA element that may be involved in seed-specific transcriptional activation of the α subunit gene is the RY sequence. RY sequences with the consensus sequence CATGCA(C/T) are widely distributed in seed-specific gene promoters (Dickinson et al., 1988). The RY sequence has also been defined as the legumin box [CATGCATG] (Bäumlein et al., 1992) or a 6-bp sequence [CATGCA] (Ezcurra et al., 1999). Base substitutions of the last two nucleotides of the CATGCATG maintain interaction with protein factors at a lower efficiency (Reid et al., 2000). Protein factors such as ABI3 and FUS3 are known to bind and control seed-specific gene expression (Reidt et al., 2000). Six RY sequences (including CATGCA located at -218 and -68) were detected upstream of the α subunit gene in this study. In the reporter assay, a stepwise decrease in GUS activity was detected with a deletion in the 5' upstream sequences that contained the RY sequence. The regions -867 to -545, -402 to -245, and -245 to -161 each contained a single RY sequence, and deletions of these regions from the 5' upstream resulted in a clear decrease in GUS activity. There was no RY element between -545 and -402, which appeared to confer negative regulatory activity. Although additional experiments that include site-directed mutagenesis on these RY sequences are necessary to demonstrate

their roles in transcriptional activation, our results indicate an association of the presence of RY sequences with transcriptional activation in the α subunit gene promoter.

Results of previous studies have suggested that while the RY sequences are important for seed-specific expression, other sequences also play critical roles in the regulation (see Fujiwara and Beachy, 1994; Sakata et al., 1997). DNase I-footprinting analysis using soybean nuclear extract revealed DNA-protein interactions on the box I sequence as well as on the RY sequence (our unpublished data). The box I sequence is highly conserved in the promoters of the α subunit gene, the α' subunit gene, and the β -phaseolin gene (Yoshino et al., 2001). Although no prominent GUS activity was detected when the -161 to $+27$ sequence was fused to the GUS gene, it is possible that the box I sequence is involved in transcriptional control in seeds by combinatorially interacting with other DNA elements located farther upstream. It should also be noted that the box I sequence is very similar to the “vicilin box” sequence (Higgins et al., 1988; Yoshino et al., 2001) and that the protein factors ROM1 and ROM2 bind the core sequence [GCCACCTCA] of the vicilin box and repress transcription (Chern et al., 1996a, b). Similarly, the box II–IV sequences that are conserved upstream of the above three seed-protein genes (see Yoshino et al., 2001) may also have functions in transcriptional activation. The GCCGCGT sequence in box III is almost identical to the complementary sequence of the ABA response elements [ABREs; ACGTG(G/T)C] (Hattori et al., 2002) with a one-base mismatch. This sequence may function in the α subunit gene promoter, and the deletion of this sequence may explain the abrupt decrease in the induction of GUS activity between construct -402 and construct -245 .

DNA-protein interactions may be more complex *in vivo* than those expected from *in vitro* assays. Indeed, *in vivo* footprinting analysis of the β -phaseolin promoter revealed that over 20 *cis*-elements within the proximal 295 bp of the promoter are protected by binding of nuclear factors in seed tissues in tobacco (Li and Hall, 1999), which represents a summation of module-specific regulation of the gene

(Chandrasekharan et al., 2003). In this study, we presented the upstream DNA sequence of the α subunit gene, which is the only promoter in the β -conglycinin gene family that has not been previously reported. Our ongoing analyses using *A. thaliana* may reveal the exact role of the RY sequences and possibly of other elements in the transcriptional control of the α subunit gene.

Acknowledgements

We thank Yasuaki Kagaya, Satoshi Naito, Ikuo Nakamura, Yoshiya Shimamoto, Jun Abe, Keisuke Kitamura for their helpful suggestions. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and a grant from Fuji Foundation for Protein Research.

REFERENCES

- Allen, R. D., Bernier, F., Lessard, P. A. and Beachy, R. N. (1989) Nuclear factors interact with a soybean β -conglycinin enhancer. *Plant Cell* **1**, 623-631.
- Bäumlein, H., Nagy, I., Villarroel, R., Inzé, D. and Wobus, U. (1992) *Cis*-analysis of a seed protein gene promoter: the conservative RY repeat CATGCATG within the legumin box is essential for tissue-specific expression of a legumin gene. *Plant J.* **2**, 233-239.
- Bechtold, N., Ellis, J. and Pelletier, G. (1993) *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Acad. Sci. Paris, Life Sciences* **316**, 1194-1199.
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
- Chamberland, S., Daigle, N. and Bernier, F. (1992) The legumin boxes and the 3' part of a soybean β -conglycinin promoter are involved in seed gene expression in transgenic

- tobacco plants. *Plant Mol. Biol.* **19**, 937-949.
- Chandrasekharan, M. B., Bishop, K. J. and Hall, T. C. (2003) Module-specific regulation of the β -phaseolin promoter during embryogenesis. *Plant J.* **33**, 853-866.
- Chen, Z.-L., Schuler, M. A. and Beachy, R. N. (1986) Functional analysis of regulatory elements in a plant embryo specific gene. *Proc. Natl. Acad. Sci. USA* **83**, 8560-8564.
- Chen, Z.-L., Pan, N.-S. and Beachy, R. N. (1988) A DNA sequence element that confers seed-specific enhancement to a constitutive promoter. *EMBO J.* **7**, 297-302.
- Chen, Z.-L., Naito, S., Nakamura, I. and Beachy, R. N. (1989) Regulated expression of genes encoding soybean β -conglycinins in transgenic plants. *Dev. Genet.* **10**, 112-122.
- Chern, M.-S., Eiben, H. G. and Bustos, M. M. (1996a) The developmentally regulated bZIP factor ROM1 modulates transcription from lectin and storage protein genes in bean embryos. *Plant J.* **10**, 135-148.
- Chern, M.-S., Bobb, A. J. and Bustos, M. M. (1996b) The regulator of *MAT2* (ROM2) protein binds to early maturation promoters and represses PvALF-activated transcription. *Plant Cell* **8**, 305-321.
- Dickinson, C. D., Evans, R. P. and Nielsen, N. C. (1988) *RY* repeats are conserved in the 5'-flanking regions of legume seed-protein genes. *Nucleic Acids Res.* **16**, 371.
- Ezcurra, I., Ellerström, M., Wycliffe, P., Stålberg, K. and Rask, L. (1999) Interaction between composite elements in the *napA* promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. *Plant Mol. Biol.* **40**, 699-709.
- Fujiwara, T. and Beachy, R. N. (1994) Tissue-specific and temporal regulation of a β -conglycinin gene: roles of the RY repeat and other *cis*-acting elements. *Plant Mol. Biol.* **24**, 261-272.
- Harada, J. J., Barker, S. J. and Goldberg, R. B. (1989) Soybean β -conglycinin genes are clustered in several DNA regions and are regulated by transcriptional and posttranscriptional processes. *Plant Cell* **1**, 415-425.

- Hattori, T., Totsuka, M., Hobo, T., Kagaya, Y. and Yamamoto-Toyoda, A. (2002) Experimentally determined sequence requirement of ACGT-containing abscisic acid response element. *Plant Cell Physiol.* **43**, 136-140.
- Higgins, T. J. V. (1984) Synthesis and regulation of major proteins in seeds. *Ann. Rev. Plant Physiol.* **35**, 191-221.
- Higgins, T. J. V., Newbigin, E. J., Spencer, D., Llewellyn, D. J. and Craig, S. (1988) The sequence of a pea vicilin gene and its expression in transgenic tobacco plants. *Plant Mol. Biol.* **11**, 683-695.
- Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. (1999) Plant *cis*-acting regulatory DNA elements (PLACE) database:1999. *Nucleic Acids Res.* **27**, 297-300.
- Hirai, M.Y., Fujiwara, T., Goto, K., Komeda, Y., Chino, M. and Naito, S. (1994) Differential regulation of soybean seed storage protein gene promoter-GUS fusions by exogenously applied methionine in transgenic *Arabidopsis thaliana*. *Plant Cell Physiol.* **35**, 927-934.
- Jefferson, R. A., Kavanagh, T. A. and Bevan, M. W. (1987) GUS fusions: β -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901-3907.
- Kawagoe, Y. and Murai, N. (1992) Four distinct nuclear proteins recognize *in vitro* the proximal promoter of the bean seed storage protein β -phaseolin gene conferring spatial and temporal control. *Plant J.* **2**, 927-936.
- Kawagoe, Y., Campbell, B. R. and Murai, N. (1994) Synergism between CACGTG (G-box) and CACCTG *cis*-elements is required for activation of the bean seed storage protein β -phaseolin gene. *Plant J.* **5**, 885-890.
- Kim, S. Y., Chung, H.-J. and Thomas, T. L. (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. *Plant J.* **11**, 1237-1251.

- Kosugi, S., Ohashi, Y., Nakajima, K. and Arai, Y. (1990) An improved assay for β -glucuronidase in transformed cells: methanol almost completely suppresses a putative endogenous β -glucuronidase activity. *Plant Sci.* **70**, 133-140.
- Lee, W., Mitchell, P. and Tjian, R. (1987) Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* **49**, 741-752.
- Lessard, P. A., Allen, R. D., Bernier, F., Crispino, J. D., Fujiwara, T. and Beachy, R. N. (1991) Multiple nuclear factors interact with upstream sequence of differentially regulated β -conglycinin genes. *Plant Mol. Biol.* **16**, 397-413.
- Li, G. and Hall, T. C. (1999) Footprinting *in vivo* reveals changing profiles of multiple factor interactions with the β -phaseolin promoter during embryogenesis. *Plant J.* **18**, 633-641.
- Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol. Plant.* **15**, 473-497.
- Naito, S., Inaba-Higano, K., Kumagai, T., Kanno, T., Nambara, E., Fujiwara, T., Chino, M. and Komeda, Y. (1994) Maternal effects of *mtol* mutation, that causes overaccumulation of soluble methionine, on the expression of a soybean β -conglycinin gene promoter-GUS fusion in transgenic *Arabidopsis thaliana*. *Plant Cell Physiol.* **35**, 1057-1063.
- Reidt, W., Wohlfarth, T., Ellerström, M., Czihal, A., Tewes, A., Ezcurra, I., Rask, L. and Bäumllein, H. (2000) Gene regulation during late embryogenesis: the RY motif of maturation-specific gene promoters is a direct target of the FUS3 gene product. *Plant J.* **21**, 401-408.
- Sakata, Y., Chiba, Y., Fukushima, H., Matsubara, N., Habu, Y., Naito, S. and Ohno, T. (1997) The RY sequence is necessary but not sufficient for the transcription activation of a winged bean chymotrypsin inhibitor gene in developing seeds. *Plant Mol. Biol.* **34**, 191-197.
- Thanh, V. H. and Shibasaki, K. (1976) Heterogeneity of β -conglycinin from soybean

- seeds. *Biochim. Biophys. Acta.* **439**, 326-338.
- Thanh, V. H. and Shibasaki, K. (1978) Major proteins of soybean seeds: subunit structure of β -conglycinin. *J. Agric. Food Chem.* **26**, 692-695.
- Yoshino, M., Kanazawa, A., Tsutsumi, K., Nakamura, I. and Shimamoto, Y. (2001) Structure and characterization of the gene encoding α subunit of soybean β -conglycinin. *Genes Genet. Syst.* **76**, 99-105.
- Yoshino, M., Kanazawa, A., Tsutsumi, K., Nakamura, I., Takahashi, K. and Shimamoto, Y. (2002) Structural variation around the gene encoding the α subunit of soybean β -conglycinin and correlation with the expression of the α subunit. *Breeding Sci.* **52**, 285-292.

Legends of figures

Fig. 1 Nucleotide sequence of the upstream region of the α subunit gene. The putative TATA box and regulatory elements for seed-specific transcriptional control are indicated below the sequence. The major transcription start site is indicated by an arrow. Nucleotide positions are numbered relative to the major transcription start site (Yoshino et al., 2001). The ATG codon is indicated by an open square.

Fig. 2 GUS expression in transgenic *A. thaliana* plants that contained a transgene comprising the upstream sequence (-1357 to +27) of the α subunit gene and the GUS gene. (A) GUS activities at different stages of seed development and in leaves in transgenic *A. thaliana* line 1357-6. Average activity with the standard error obtained from six individual plants is shown. The divisions of the scale indicate 1 mm. (B–E) Histochemical staining of GUS activity: (B) seeds within a silique; (C) an embryo dissected from a developing seed 10 days after flowering; (D) the remaining portions of the seed after the embryo was removed; (E) the whole plant grown for 10 days. Scale bars in (C) and (D) denote 100 μ m.

Fig. 3 Diagram of the upstream regions of the α subunit gene that were fused to the GUS gene and the GUS activities of respective reporter constructs in transgenic *A. thaliana* plants. Positions of nucleotides are numbered relative to the transcription start site. Locations of the RY sequences and the TATA box sequence are indicated by vertical lines. Average activity with the standard error for each gene construct obtained from each six individual plants of five independently transformed lines of *A. thaliana* is shown.

Figure 1

TAGATTTTACACAAAACTCAACTTGATATTTACTCGCTTTTGTTTGTGTATAAATGTTTGTCTACTGAACATACATTTTCTGAGTGAA -2774

ACAAGTYCCCTGTGGATACAATACTCGGTTCTNACCCTTTTATATTACTTGCACGACTCGGTGCCACTTGCCAAATAACATCGCAGAAAG -2684

ACCACGCAGGCACACCCACGATGGTGGAAATGGTTGTGGTAGATTGGTGGTGGTGGAGGAGACATCATTGATAATGTGGTGGTGGCGACAG -2594

TGGTGGGGGCAATGGTAGTAGTATGGTGTAGTGGTTAGAATGTGTTTTTAAAACTAACCCAAAATTCACAAAACCTTAGGTACTGTTTGG -2504

GGTGTTTTTTAATTTTGAAGTTTAAAAATTTAAAAAACAACAGGTTTTAGTTATAGTTTTTAAAAAATTTAAACTATTTTTACAA -2414

TATGATTAACCTTCAAATAAACTCATTTTAAAAAATTCATATCATATTAAGATTAATTTAAAAGTGTTTTAGTGTTGGTATGGCAACGATG -2324

GTGATAATGGTAATGTCAGTAGTAGCAGTGATGATAGAAATATCAATGGTGGTAACGATAGTGGCATTGGAGGTGACATTGGTGGCGGTG -2234

GTGATGAAGTCGATTGCGATAGTGGTGTGGTCAACAACCAATTATGGCCGTGACAATAATAGTGTGGTAGTGACAATGATGACAATGAT -2144

GATGGTGTGACAATGATGGAGATAGTAATGATAGTGATAACAGAAAATGTCAAATATGAGATATGTGTGTTTTTAATCTAAAAATCAA -2054

TTCAAAAAATCTTTTCTCTTGATTTTTAAATAACTGTA AAAAATTTTTAAAAATCAATTCACCCATTTTTATCAAATACCTTTTATT -1964

TTAAAAATTATTTTAAATCAAAAAATTA AAAAATTTATGATCACTCCACCCTAACTTTAGATTTTCTCTCTTTTGACTCATGTCTTCTTA -1874

GGTCTTATCAACCAATTGAAATAGTCTTTTGAATTTAAGTTTCCCTTTCTTTTCAAATCTTCCCTTGATCTTCATCGCTTAATTAAG -1784

AAAGTAATTATTTTTTAGAGCTATTATCAAAATTATAAAAAAACATTTGTAATTAATNACTTAGTTAGTTAGTTACAAACCATGCT -1694

TATATAAAATGCATTAATGATGAGATAAAAAAAATGACGTGAACTCAATTGATGTCATTCAGCAGCTGTCTTTTCAATTTTTTTGG -1604

CCACATATTATTGGGTTCTGTGACCTTTCAAATGACTGCTATTACCTCTGACCTTCTATTACATCTTGACCATCACTAGGCATTT -1514

AAAAGTATTAGTCATAGTCACATATTACTACAAGCGAGATTGATCTCTCTAATCTAATGGGTGGGAAAACACTTATAATATATGATTCA -1424

AGAAAAGAAAGTAAATAAAACAATTTTATTATATAAAGACTATTAGGATAAAAAAAACCTTAAAGTGCTTGGATTGGACCAGACTTGA -1334

ATTTTAATTTAATGATATTATAATATGTGAATATTTTTTGAGCAATTGTAAATTTTCAAGATAAAAAAATAATGTAATTAATTAATGTAAT -1244

AACTATATCGTATACTTAATTAATTATTAAATGTGACAAAAAAGATATACATCAAACTTAATGTTTCATGACTTTTTTTTTAATGTGTG -1154

TCCTAAATAGAAATTA AAAAATTAATAATATATCCAAATGAAAAAAACATTTAATACGTATTATTTAAGAAATAACAATATATTTATA -1064

TTTTAATATGATTTCACATGTAATTTAAAAACAAAAACAAATTTCTCTTTTATTGATTAATTAATAAATTTTATAACTACATTTATT -974

TTCTATTATTATCAATTTTCTCTGTTTTTTTATTGGCATATATACCTAGACAAGTCAAAAAATGACTATTCTTTAATAATCAATCATT -884

ATTCTTACATATTGTTTGAACTACGAGTTATGAAGTGTCAATTGACCTTAGTGTTTTGATAGGCCTCCATTGCCGCTCATTAATTAAT -794

TTGATAACAGCCGTACCGATCAATTAATGCTTCTTCCATCGTAATTATATGCATGTCGGTTCTTTTAATCTTGGTACTCTCGAATGC -704

CACCACAACACTGACTAGTCTCTGGATCATGAGAAAAAGCCAAAGAACA AAAAAGACAACATAAAGATATCCTTTGCCAAAAAATGTC -614

TAAGTTCATAAAATACAAACAAAAACGCAATCACACACAGTGGACCCAAAAGCCATGCACAACAACACGTAATCACCAGGTGCAATCGT -524

GCTGCCAAAAACATTCAACCACTCAATCCATGATGAGCCACACATTTGTGTTTGTAAACCAATCTCAAACGCGGTGTTCTCTTTGGA -434

AAGCAACCATATCAGCATATCACACTATCTAGTCTCTGGATCATGATGCGCAACCAAAAGACAACACATAAAGTATCCTTTCGAAAGC -344

AATGTCCAAGTCCATCAAATAAAATGAGACAAAATGCAACCTCACCCACTTCACTATCCATGGCTGATCAAGATCGCCGCTCATGT -254

AGGTCTAAATGCCATGCACATCAACACGTAATCAACATGCAGCCAAATGCTCACCATCGCTCAACACATTTCTTGTTAATTTCTAAGT -164

ACACTGCCTATGCGACTCTAACTCGATCACAACCATCTTCCGTACATCAATTTTGTTCCAATTCAACACCCGTCAAACTTGCAATGCCACC -74

CCATGCAAGTAAACAAGAGCTATATCTCTTCTATGACTATAAAATACCCGCAATCTCGGTCCAGGTTTTTCATCATCGAGAAGTAGTT +16

CAATATCCTAGTATACCTTAATAAATAATTTAAGATACTATG

Labels: SEF4, SEF3 core, SEF1, E-box, G-box, RY, Box I, Box II, Box III, Box IV, DPBF, TATA box, -1357, -867, -402, -245, -161, -73, +1

Figure 2

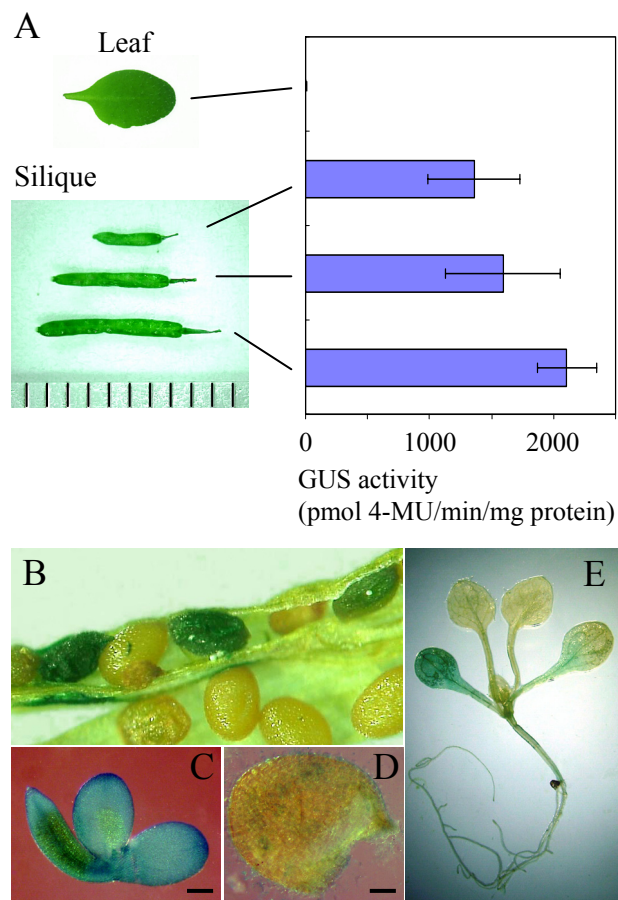


Figure 3

