Japanese Society of Breeding

Breeding Science 51 : 185-190 (2001)

Genetic Characterization of a Novel Soybean Kunitz Trypsin Inhibitor

Ke-Jing Wang[1], Yoshihito Takahata*, Kikukatsu Ito, Yun-Peng Zhao, Ken-ichi Tsutsumi and Norihiko Kaizuma

Faculty of Agriculture, Iwate University, 3-18-8 Ueda, Morioka, Iwate 020-8550, Japan

[1] *Present address*: *Institute of Crop Germplasm Resources, Chinese Academy of Agricultural Science*, Beijing 100081, China

Genetic and nucleotide sequence studies were performed on a new variant of soybean Kunitz trypsin inhibitor (SKTI) detected in wild soybean (*Glycine soja*) and showing a slightly slower electrophoretic mobility than the *Tia* type. The segregation analysis of SKTI bands in F_2 seeds from crosses of the new variant type with *Tia* or *Tib* type showed that the variant type is controlled by an allele codominant to *Tia* and *Tib* at an SKTI locus. Nucleotide sequence analysis showed that this variant has 217 amino acids composed of 181 amino acid residues of mature SKTI and extra 25 and 11 amino acids at N- and C-terminal regions, respectively. This sequence was identical to that of *Tia* (= *KTi3*), except that a G→A transitional mutation occurred at position 500 of *Tia*, which results in the translational change from Arg to Lys. The result of isoelectric forcusing-PAGE coincided with this change. In addition, three nucleotides GCT were inserted at the N-terminus, which leads to an Ala addition in the precursor of this SKTI protein. From these results, we propose the genetic symbol *Tie* for the new variant of SKTI.

Key Words: ***Glycine soja*, *G. max*, soybean Kunitz trypsin inhibitor, multiple allele, gene sequence.**

Introduction

Proteinase inhibitors have been identified from many plants, animals and microorganisms (Ryan 1981). Though their physiological functions in plants are not clear, they are considered to be regulatory agents in controlling endogenous proteinases, storage proteins and protective agents against insect and microbial proteinases (Ryan 1973). Proteinase inhibitors account for about 6% of the total protein of the seeds of soybean (*Glycine max*). These inhibitors are roughly composed of Kunitz trypsin inhibitor (SKTI) (Kunitz 1945) and Bowman-Birk trypsin inhibitor (Bowman 1946, Birk 1961). Since Singh *et al.* (1969) found electrophoretic polymorphism of the SKTI such as *Tia* and *Tib*, additional variants have been reported, namely, *Tic* (Hymowitz 1973), the slowest mobility form of *Tid* (Zhao and Wang 1992, Hua *et al.* 1999) and null type of *ti* (Orf and Hymowitz 1979). Of these, *Tia*, *Tib* and *Tic* are codominant multiple alleles at a single SKTI locus (Singh *et al.* 1969, Hymowitz and Hadley 1972, Orf and Hymowitz 1977). The polymorphism of SKTI has often been used as an index for analyzing the botanical origin, geographical diversification and phylogenetic relationship of soybean (Hymowitz and Kaizuma 1979, Kaizuma *et al.* 1980, Nakamura *et al.* 1984, Kiang *et al.* 1992, Wang *et al.* 1998a).

There is a large sequence difference in eight amino acid residues between *Tia* and *Tib* proteins (Kim *et al.* 1985). The substitutions are considered to have resulted from the accumulation of several mutations. However, no intermediate forms between them have been found in either cultivated or wild (*G. soja*) soybean. Previously, we reported two new variants of SKTI detected by electrophoretic analysis of 173 cultivars of soybean and 890 lines of wild soybean (Wang *et al.* 1996). One is a variant showing a slightly slower electrophoretic mobility than the *Tia* type (tentatively designated as *Tia-s* in Wang *et al.* (1996)), and another is one with a slightly faster mobility than the *Tib* (tentatively designated as *Tib-f*). Though we confirmed the transmissibility of these variants by the parent-offspring test and trypsin-inhibiting activity of *Tia-s* (Wang *et al.* 1996, 1998b), their allelism to other SKTI types and the gene and amino acid structure remained to be determined.

The objective of this study was to determine the mode of inheritance and the nucleotide and deduced amino acid sequence of the *Tia-s* variant of SKTI. In addition, the electrophoretic characters of this variant were examined. As described in 'Discussion', we propose the genetic symbol of *Tie* for the new variant. Here after we use *Tie* instead of *Tia-s*.

Materials and Methods

Plant materials

A new variant line for SKTI, wild soybean (*G. soja*) line 1125 (*Tie* type) and 3 soybean (*G. max*) cultivars, 'Rikuu No. 27' (*Tia*), 'Odate No. 1' (*Tib*), and 'Tachisuzunari' (*Tib*) were used in this study. *G. soja* 1125 line was hybridized with 'Rikuu No. 27', 'Odate No. 1', and 'Tachisuzunari' cultivars, respectively. Their F_1 plants were grown to obtain F_2 seeds.

Communicated by K. Kitamura
Received October 23, 2000. Accepted March 16, 2001.
*Corresponding author (e-mail: ytakahata@iwate-u.ac.jp)

NII-Electronic Library Service

Japanese Society of Breeding

Electrophoresis

Extraction of SKTI proteins from seeds was carried out according to the procedure of Hymowitz and Hadley (1972). The SKTI proteins were analyzed by Davis system of PAGE and isoelectric focusing (IEF)-PAGE as described by Wang *et al.* (1996). A urea-PAGE system (10% polyacrylamide and 2.7% Bis) containing 5 M urea was applied for analyzing the SKTI profile in immature (0.031 gfw/seed) and nearly mature (0.096 gfw/seed) seeds.

Nucleotide sequence analysis

Total DNA was extracted from 100 mg young leaf tissue of *G. soja* 1125 by the cetyltrimethylammonium bromide (CTAB) method (Rogers and Bendich 1985).

To amplify the *Tie* gene of 1125 line by PCR, a set of two 20 bp primers (5′-TAGTCCCGATTCTCCCAACA-3′, 3′-CTGTGTTCACACTCTCATGA-5′) was synthesized based on a DNA sequence of *KTi3* (= *Tia*) gene (Jofuku *et al.* 1989). The 20 μl PCR reaction mixture contained 100 ng template DNA, 1 U *Taq* DNA polymerase (TaKaRa), 2.5 mM of each dNTP, 10 pmol of each primers and 2 μl of 10X Buffer (TaKaRa). The reaction mixture was subjected to 40 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 60 s using a Thermal Cycler 480 (Perkin-Elmer Cetus). The PCR product was electrophoresed on a 1% agarose gel and stained with 1 μg/ml ethidium bromide.

The amplified DNA was purified and was then ligated into a T-vector plasmid (pBluescript (+)). The recombinant DNA (T-vector) was transferred into *E. coli* HB101 cells and the transformed *E. coli* was cultured on solid LB media with ampicilline (50 mg/ml), IPTG (30 ml in 9 cm Petri dish) and X-gal (0.6 mg/Petri dish) to select recombinant clones. The cloned DNA was isolated and sequenced by the dideoxynucleotide chain termination method (Sanger *et al.* 1977).

Results

Inheritance of the new variant of SKTI

As shown in Fig. 1, the newly detected SKTI variant (*G. soja* 1125 line, *Tie*) carries a slow mobility in comparison with the standard *Tia* type (lanes 1 and 3: *Tia*, lane 2: *Tie*). We already observed that the SKTI bands of the F_1 seeds from crosses between *Tia* and *Tie* or between *Tib* and *Tie* were composed of the parents' bands (Wang *et al.* 1996).

F_2 seeds obtained from a F_1 plant between Rikuu No. 27 (*Tia*) and *G. soja* 1125 were analyzed individually by PAGE for the segregation ratio of the new SKTI type. The genetic segregation of the F_2 seeds showed an acceptable fit to the ratio of 1 *Tia* band : 2 both *Tia* and *Tie* bands : 1 *Tie* band (Table 1). In addition, segregation of F_2 seeds from two crosses between *Tib* and *Tie* also gave a good fit to the expected ratio of 1 *Tib* band : 2 both *Tib* and *Tie* bands : 1 *Tie* band (Table 2). These results demonstrated that the *Tie* type is a newly additional codominant allele at the SKTI locus.

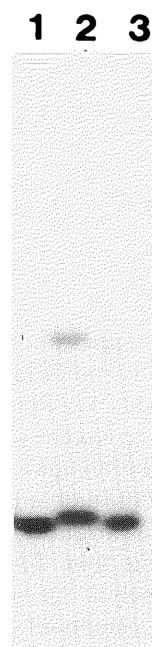

Fig. 1. A new variant of SKTI in Davis system of PAGE. Lanes 1 and 3: *Tia* type. Lane 2: A new variant of SKTI of *G. soja* 1125 line (*Tie*) showing a slightly slower electrophoretic mobility.

Electrophoretic characteristics

IEF of SKTI proteins revealed that the isoelectric point of the *Tie* band was different from that of *Tia* (Fig. 2, lanes 1 and 3: *Tie*, lane 2: *Tia*). The isoelectric point of *Tie* protein was slightly lower (acidic) than that of *Tia*.

Urea-PAGE of the *Tie* protein at two different seed-developmental stages revealed that the immature seeds had a slightly slower mobility band of SKTI protein than seeds at the nearly mature stage (Fig. 3, lane 1: immature seed, lane 2: nearly mature seed).

Table 1. Observed and expected segregation of F_2 seeds from the cross between a new variant type (*Tie*) and *Tia* type for SKTI bands

Cross		No. of F_2 seeds examined	F_2 SKTI bands			χ^2 (1:2:1)	Probability
			Tia	Both	*Tie*		
G. soja 1125 (*Tie*)	Obs.	100	25	47	28	0.54	0.7 < P < 0.8
× *G. max* cv. Rikuu No. 2 (*Tia*)	Exp.		25	50	25		

NII-Electronic Library Service

Japanese Society of Breeding

Table 2. Observed and expected segregation of F_2 seeds from the cross between a new variant type (*Tie*) and *Tib* type for SKTI bands

Cross		No. of F_2 seeds examined	F_2 SKTI bands			χ^2 (1:2:1)	Probability
			Tib	Both	*Tie*		
G. max cv. Odate No. 1 (*Tib*)	Obs.	100	29	47	24	0.86	$0.5 < P < 0.7$
× *G. soja* 1125 (*Tie*)	Exp.		25	50	25		
G. max cv. Tachisuzunari (*Tib*)	Obs.	100	22	53	25	0.70	$0.7 < P < 0.8$
× *G. soja* 1125 (*Tie*)	Exp.		25	50	25		

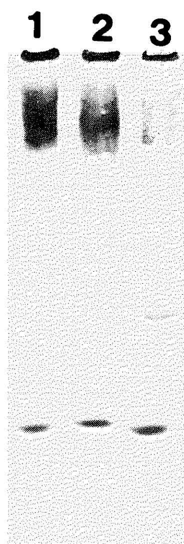

Fig. 2. IEF-PAGE banding pattern of the SKTI. Lanes 1 and 3: *Tie* of SKTI of *G. soja* 1125 line. Lane 2: *Tia* type.

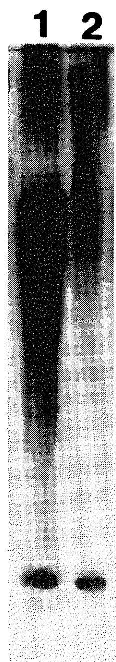

Fig. 3. Urea-PAGE banding pattern of the SKTI in the different seed developmental stages of *G. soja* 1125. Lane 1: Immature seed (0.031 gfw/seed). Lane 2: Nearly mature seed (0.096 gfw/seed).

Nucleotide sequence analysis

The genomic DNA of *G. soja* 1125 was extracted and the SKTI gene was amplified using PCR. Electrophoresis of the PCR-product showed that a fragment with the expected length was amplified (data not shown). Two PCR-products obtained through two separate PCR operations were ligated to T-vector and transferred into *E. coli* HB101 cells. Two colonies randomly selected from two groups of transformed colonies were used to analyze the gene structure.

Sequence analyses indicated that the two PCR-products showed the same sequence composed of 743 bp containing an open reading frame of 651 bp encoding 217 amino acids (Fig. 4). The nucleotide sequence of *Tie* was compared with two sequences reported for the *Tia* gene: *KTi3* (Jofuku *et al.* 1989) and *Tia* (Song *et al.* 1993). Sequences of *KTi3* and *Tia* are identical, except that three nucleotides GCT were inserted at positions 91-93 of *Tia*, which leads to an additional amino acid of Ala. *Tie* gene had the insertion of three nucleotides GCT at positions of 97-99 instead of deletions of three nucleotides GCT at positions 91-93 of *Tia*. In addition, one G→A transition occurred at position 500 of *Tie*. This mutation results in a translational change from Arg to Lys.

Discussion

So far, four types of SKTI proteins have been reported. Of these, *Tia*, *Tib* and *Tic* have been frequently investigated from the points of view of genetics, geographic distribution and chemical structure. They are codominant multiple alleles at the SKTI locus (Orf and Hymowitz 1977). The null allele *ti* was reported to be inherited as an allele recessive to the others (Orf and Hymowitz 1979). Zhao and Wang (1992) reported a fourth SKTI, *Tid* which showed slower mobility than the *Tib* type, but its inheritance mode was not clear.

The present study revealed that a new variant [tentatively named *Tia-s* by Wang *et al.* (1996)] of SKTI is controlled by a codominant allele to *Tia* and *Tib* at a SKTI locus. Although we did not investigate its genetic relationship with *Tic* and *ti*, our results indicate that the new SKTI type is one of multiple alleles at a SKTI locus. We propose the genetic symbol *Tie* for this electrophoretic form of SKTI.

Mature SKTI protein purified from seeds consists of 181 amino acid residues (Kim *et al.* 1985). When the amino

NII-Electronic Library Service

Japanese Society of Breeding

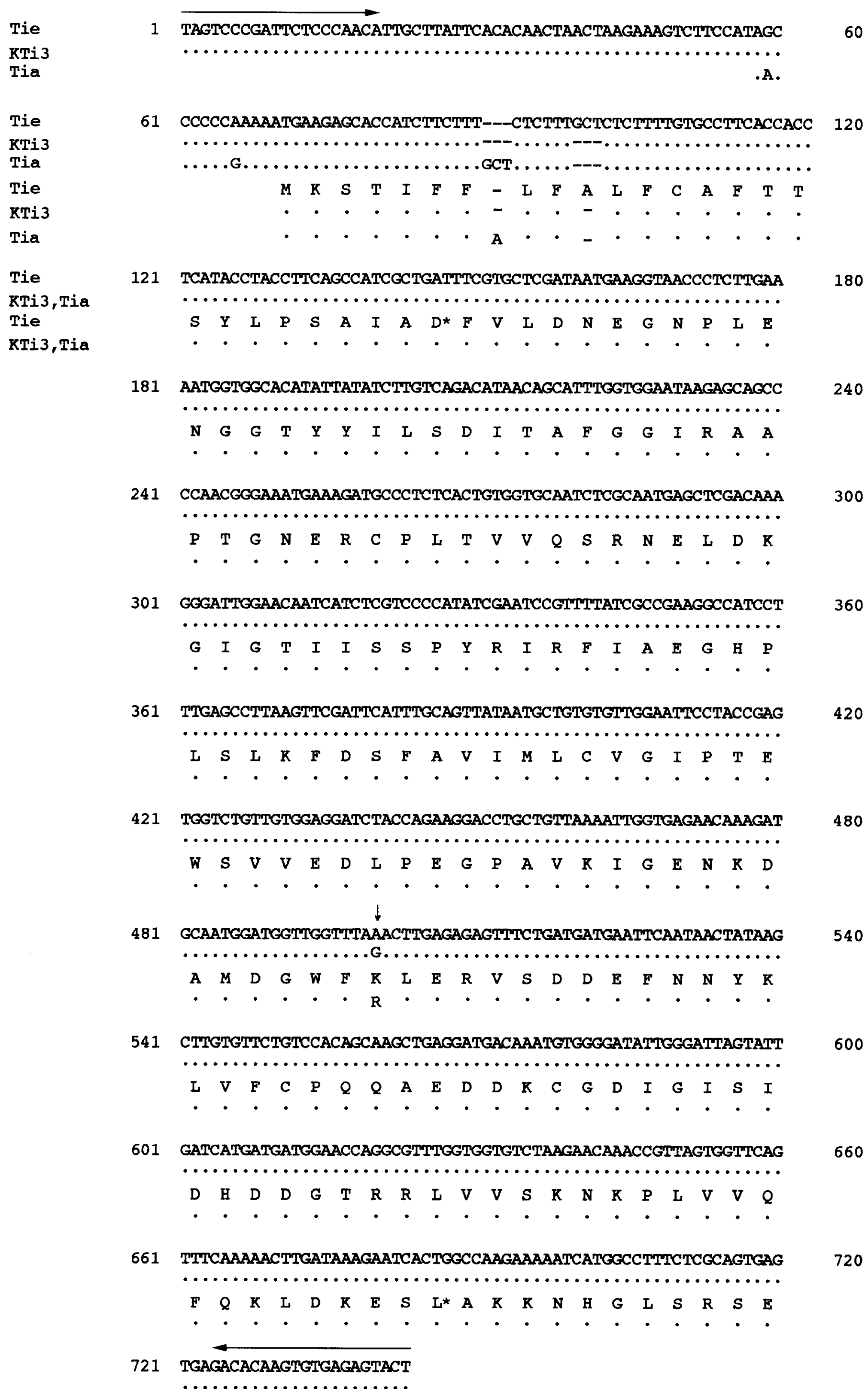

Fig. 4. The nucleotide and deduced amino acid sequence of *Tie* of SKTI in *G. soja* 1125 line. The sequence is compared with those of *KTi3* (= *Tia*) (Jofuku *et al.* 1989) and *Tia* (Song *et al.* 1993). Dots in sequence indicate identical nucleotides and amino acids. Dashes indicate gaps. The amino acid sequence of the mature SKTI proteins ranges from D* to L*. The vertical arrow shows a transitional change in *Tie*. Horizontal arrows show the positions of primers used.

NII-Electronic Library Service

Japanese Society of Breeding

acid sequence deduced from cDNA sequence of *Tia* gene was compared with that of mature SKTI from seed, they were identical (Jofuku *et al.* 1989, Song *et al.* 1993). However, the results of cDNA sequences revealed that there were an extra 24 (Jofuku *et al.* 1989) or 25 (Song *et al.* 1993) amino acids at the N terminus and 11 amino acids at the C terminus. Our results of the *Tie* gene support that the SKTI gene has some extra amino acids at the N and C terminus. The 25 amino acids at the N terminus of the *Tie* gene was identical with the *KTi3* (= *Tia*) of Jofuku *et al.* (1989) and the *Tia* gene of Song *et al.* (1993), except that Ala added at amino acid position 10 of *Tie* and Ala deleted at position 8 of *Tia* reported by Song *et al.* (1993). It is presumed that the extra amino acids at the N terminus are a signal peptide which is usually observed in many storage proteins and that those at the C terminus are cleaved during seed maturation (Jofuku *et al.* 1989, Song *et al.* 1993). Our results showed that the electrophoretic mobility of immature seed SKTI was slower than that of mature seed. This supports the idea that SKTI of immature seed may possess the extra amino acids.

Kim *et al.* (1985) reported that there is a sequence difference in eight amino acids between *Tia* and *Tib* SKTI proteins and a smaller difference in one amino acid between *Tia* and *Tic*. The amino acid sequence deduced from nucleotide sequence of the *Tie* gene revealed that the *Tie* protein of SKTI is identical to *Tia*, except that amino acid 119 of *Tie* is Lys rather than Arg of *Tia*. This change is derived from a G$\rightarrow$A transitional mutation. Since the pI and pKR of Lys (9.74 and 10.53, respectively) are lower than those of Arg (10.76 and 12.48), *Tie* protein is expected to be more acidic than *Tia*. This is in agreement with the IEF results.

Kaizuma *et al.* (1980) considered that the *Tia* type is the prototype from which the other *Tib* and *Tic* types were derived, and that differentiation of *Tib* from *Tia* occurred much earlier than that of *Tic* from *Tia*. The larger difference in the amino acid sequence between *Tia* and *Tib* than that between *Tia* and *Tic* supported this hypothesis (Kim *et al.* 1985). The SKTI protein of *Tie* is considered to be differentiated from *Tia* by a one-point mutation. This indicated that the diversification of *Tie* would have arisen relatively recently. On the other hand, it is difficult to conceive that the substitutions involved in eight amino acid residues between *Tia* and *Tib* occurred by a single mutation. Therefore, some intermediate or transitional forms are expected to be found in cultivated and/or wild soybeans by a modified method of electrophoresis such as that used by Wang *et al.* (1996), which detects any small differences in the mobility of SKTI variants.

Acknowledgment

This work was supported by a Grant-in Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan.

Literature Cited

Birk, Y. (1961) Purification and some properties of a highly active inhibitor of trypsin and α-chymotrypsin from soybean. Biochim. Biophys. Acta 54: 378-381.

Bowman, D.E. (1946) Differentiation of soy bean antitryptic factors. Proc. Soc. Exptl. Biol. Med. 63: 547-550.

Hua, X., X. Kefang, D. Aiwu, Y. Qingyan and G. Qimin (1999) The amino acid sequence determination of a new variant of Kunitz soybean trypsin inhibitor (SBTi-A2). Soybean Genetics Newsletter 26 [Online journal]: URL *http://www.soygenetics.org/articles/sgn1999-004.html*

Hymowitz, T. (1973) Electrophoretic analysis of SBTI-A2 in the USDA soybean germplasm collection. Crop Sci. 13: 420-421.

Hymowitz, T. and H.H. Hadley (1972) Inheritance of a trypsin inhibitor variant in seed protein of soybean. Crop Sci. 12: 197-198.

Hymowitz, T. and N. Kaizuma (1979) Dissemination of soybean (*Glycine max*): seed protein electrophoresis profiles among Japanese cultivars. Econ. Bot. 33: 311-319.

Jofuku, K.D., R.D. Schipper and R.B. Goldberg (1989) A frameshift mutation prevents Kunitz trypsin inhibitor mRNA accumulation in soybean embryos. Plant Cell 1: 427-435.

Kaizuma, N., K. Oikawa and M. Miura (1980) Consideration on the cause of the differential *Ti* alleles frequency distributions found among some regional populations of soybean (*Glycine max* (L.) Merrill) land varieties. J. Fac. Agr. Iwate Univ. 15: 81-96.

Kiang, Y.T., Y.C. Change and N. Kaizuma (1992) Genetic diversity in natural population of wild soybean in Iwate prefecture, Japan. J. Hered. 83: 325-329.

Kim, S.H., S. Hara, S. Hase, T. Ikenaka, H. Toda, K. Kitamura and N. Kaizuma (1985) Comparative study on amino acid sequence of Kunitz-type soybean trypsin inhibitors, *Tia*, *Tib*, and *Tic*. J. Biochem. 19: 435-448.

Kunitz, M. (1945) Crystallization of trypsin inhibitor from soybean. Science 101: 668-669.

Nakamura, I., K. Kitamura, N. Kaizuma and Y. Futsuhara (1984) Speciation of the genus *Glycine*. I. Electrophoretic properties of Kunitz trypsin inhibitor like proteins. Jpn. J. Breed. 34: 468-477.

Orf, J.H. and T. Hymowitz (1977) Inheritance of a second trypsin inhibitor variant in seed protein of soybeans. Crop Sci. 17: 811-813.

Orf, J.H. and T. Hymowitz (1979) Inheritance of the absence of the Kunitz trypsin inhibitor in seed protein of soybeans. Crop Sci. 19: 107-109.

Rogers, S.O. and A.J. Bendich (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol. Biol. 5: 69-76.

Ryan, C.A. (1973) Proteolytic enzymes and their inhibitors in plants. Annu. Rev. Plant Physiol. 24: 173-196.

Ryan, C.A. (1981) Proteinase inhibitor. *In* "Biochemisty of Plants, Vol. 6" Marcus, A. (ed.), Academic Press, New York. p. 351-370.

Sanger, F., S. Nicklen and A.R. Coulson (1977) DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Singh, L., C.M. Wilson and H.H. Hadley (1969) Genetic differences in soybean trypsin inhibitor separated by disc electrophoresis. Crop Sci. 9: 489-491.

Song, S.I., C.H. Kim, S.J. Back and Y.D. Choi (1993) Nucleotide sequences of cDNAs encoding the precursors for soybean (*Glycine max*) trypsin inhibitors (Kunitz type). Plant Physiol. 101: 1401-1402.

NII-Electronic Library Service

Japanese Society of Breeding

Wang, K.-J., N. Kaizuma, Y. Takahata and S. Hatakeyama (1996) Detection of two new variants of soybean Kunitz trypsin inhibitor through electrophoresis. Breed. Sci. 46: 39-44.

Wang, K.-J., N. Kaizuma and Y. Takahata (1998a) The polymorphisms and geographic distributions of four seed proteins and two traits in Japanese natural populations of wild soybean (*Glycine soja*). Acta Agro. Sini. 24: 147-154.

Wang, K.-J., N. Kaizuma, Y. Takahata, T. Yamashita and H. Taira (1998b) Purification and activity characterization of a Kunitz trypsin inhibitor variant in wild soybean (*Glycine soja*) in Japan. Acta Agro. Sini. 24: 665-669.

Zhao, S.W. and H. Wang (1992) A new electrophoretic variant of SBTi-A_2 in soybean seed protein. Soybean Genetics Newsletter 19: 22-24.

NII-Electronic Library Service