Theor Appl Genet (2008) 117:565–573
DOI 10.1007/s00122-008-0800-9

ORIGINAL PAPER

Allelic differentiation of Kunitz trypsin inhibitor in wild soybean (*Glycine soja*)

K. J. Wang · Y. Takahata · Y. Kono · N. Kaizuma

Received: 18 January 2008 / Accepted: 12 May 2008 / Published online: 30 May 2008

Abstract Soybean Kunitz trypsin inhibitor (SKTI) has several polymorphic types, which are controlled by co-dominant multiple alleles at a single locus. Of these types, *Tia* and *Tib* are predominant types, and there are nine differences in amino acids between *Tia* and *Tib*. Recently, an intermediate transitional type (Tib^{i5}) between them was detected. However, other transitional types have not been detected despite surveys of many cultivated and wild soybeans. One of the reasons why other transitional variants have not been found is inferred to be due to the difficulty of the detection of SKTI protein variants by polyacrylamide gel electrophoresis (PAGE). To detect novel variants of SKTI, nucleotide sequence analysis in addition to PAGE was carried out. Four new variants were found from many Japanese wild soybeans. Of these variants, three (designated as Tia^{a1}, Tia^{a2}, Tia^{b1}) were detected through gene sequence analysis on wild soybeans having the same electrophoretic mobility as *Tia*, and one (*Tig*) was detected through PAGE. The *Tig* variant showed a slightly lower electrophoretic mobility than *Tic*. The nucleotide sequences of *Tig* were identical to those of *Tib* except for one T → C transitional mutation at position +340. The sequences of Tia^{a1} and Tia^{a2} genes were identical to those of *Tia* with the exception of a G → A mutation at position +376 and a T → C mutation at +404, respectively. The sequence of Tia^{b1} differed from *Tia* by three nucleotides: C → A at position +331, T → C at +459 and A → G at +484. Of the three nucleotide changes, two were common to Tia^{b1}, Tib^{i5} and *Tib*, suggesting that Tia^{b1} is an intermediate transitional type between *Tia* and *Tib*. Our results suggest that *Tib* type has been differentiated through a series of mutations from *Tia* before the domestication of cultivated soybean.

Communicated by A. Bervillé.

K. J. Wang · Y. Takahata (✉) · N. Kaizuma
Faculty of Agriculture, Iwate University,
Morioka 020-8550, Japan
e-mail: ytakahata@iwate-u.ac.jp

K. J. Wang
The National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI), Institute of Crop Science,
Chinese Academy of Agricultural Science,
Beijing 100081, People's Republic of China

Y. Kono
Soybean Breeding, Research Subteam (Tohoku Region),
National Agricultural Research Center for Tohoku Region,
Kariwano 019-2112, Akita, Japan

Introduction

The subgenus *Soja* is an important subgenus in *Glycine*. It includes two species, an economically valuable cultivated soybean (*Glycine max* (L.) Merr.) and a wild soybean (*Glycine soja* Sieb. and Zucc.) that is a progenitor of cultivated soybean. Since the two species have the same genome (GG, $2n = 40$), the wild soybean has been accepted as an important genetic resource for improving soybean.

In genetic variation or diversity research of the subgenus *Soja*, many proteins and DNA markers have been used (Yu and Kiang 1993; Tozuka et al. 1998; Tian et al. 2000). Of these markers, a proteinase inhibitor, soybean Kunitz trypsin inhibitor (SKTI) protein has been utilized in assessing geographical variation and studying evolutionary process because of its polymorphism (Hu and Wang 1985; Fujita et al. 1997; Hymowitz and Kaizuma 1979; Kaizuma et al. 1980; Kiang et al. 1992; Wang et al. 1998). Christeller (2005) reported that proteinase inhibitors are one of the most actively evolving proteins and are a potential model

system to study evolutionary phenonomena. In addtion, the genes of proteinase inhibitors have also been paid attention to in terms of insect-resistant plant development (Lee et al. 1999; Marchetti et al. 2000).

SKTI has been found to have eight electrophoretic distinguishable forms: *Tia* and *Tib* (Singh et al. 1969), *Tic* (Hymowitz 1973), *Tid* (Zhao and Wang 1992), *Tie* (Wang et al. 1996, 2001), *ti*-null type (Orf and Hymowitz 1979), *Tif* (Wang et al. 2004) and Tib^{i5} (Wang et al. 2005). These types are controlled by co-dominant multiple alleles at a single locus. Studies of amino acid and nucleotide sequences of polymorphic variants of SKTI have uncovered that a large sequence difference in nine amino acid residues is present between *Tia* and *Tib* (Song et al. 1993; Wang et al. 2004), *Tic*, *Tid* and *Tie* differ from *Tia* by only one amino acid (Kim et al. 1985; Xin et al. 1999; Wang et al. 2001), and *Tif* differs from *Tib* by one amino acid (Wang et al. 2004). Kaizuma et al. (1980) suggested that the differentiation of *Tia* and *Tib* was very ancient and probably had already been completed before the domestication of cultivated soybean from wild soybean. They considered that the *Tia* type is hypothetically the prototype from which *Tib* and *Tic* derived, because of an absolute ascendancy of *Tia* in wild soybean. Although there are no other evidences to enforce this hypothesis, predominant frequency of *Tia* in wild soybean of China and other Asian countries, reported by Hymowitz and Kaizuma (1981) and Li (1993), supports this idea. A large difference of nine amino acid substitutions between *Tia* and *Tib* makes it difficult to visualize that *Tib* type had been differentiated from *Tia* through a single mutational event. Recently, we discovered a transitional intermediate type Tib^{i5} between *Tia* and *Tib* from a survey of a large number of wild soybeans (Wang et al. 2005). This intermediate type has five amino acid substitutions from *Tia* type. However, other transitional types have not been detected.

Many researchers have been investigating the polymorphism of SKTI proteins by polyacrylamide gel electrophoresis (PAGE), both by using the Davis system and the isoelectrofocusing (IEF) system. Though three variants, *Tia*, *Tib* and *Tic* were easily discriminated by PAGE from each other, detection of other variants was difficult because their electrophoretic mobility was very similar to that of the three variants. One of the reasons other transitional variants between *Tia* and *Tib* have not been found is due to the difficulty of detection by PAGE. We considered the possibility that there are different gene sequences in spite of the similar electrophoretic mobility of SKTI proteins. To demonstrate the possibility, we attempted the gene sequence analysis of SKTI in the same materials as a previous study (Wang et al. 2005). In this paper, we report the detection of some new variants through gene sequence analysis on many wild soybean lines having identical mobility of PAGE on SKTI protein analysis, using PAGE with a large gel. In these new variants, there is a transitional type between *Tia* and *Tib*. The evolutionary relationship among SKTI variants and the structural features of the alleles in the subgenus *Soja* are also discussed.

Materials and methods

Plant materials

A total of 720 wild soybean lines, which were the same materials used in a previous study (Wang et al. 2005), were used. The materials were collected from 77 natural habitats in 35 prefectures across four geographic regions in Japan (Fig. 1, Table 1) and were preserved at the Soybean Breeding Lab, the National Agricultural Research Center for Tohoku Region (NARCT) and the Plant Breeding Lab, Iwate University in Japan. Six soybean varieties, Rikuu No. 27 (*Tia*), Norin No. 2 (*Tib*), Odate No. 1 (*Tib*), Tachisuzunari (*Tib*), L118 (*Tic*) and Raiden (*Tic*), were used as standard types of SKTI.

Electrophoreses

Extraction of SKTI protein from seeds was carried out according to the Hymowitz and Hadley's (1972) procedure.

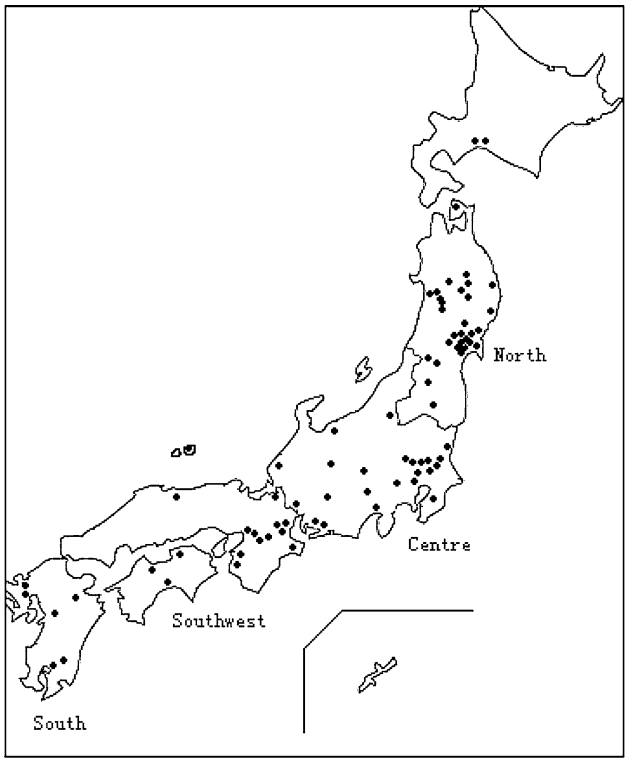

Fig. 1 Collection sites of the wild soybean lines used in this study. Distribution area of the lines covered four geographic regions of south, southwest, centre and north Japan

Table 1 Electrophoretic types and frequency distribution of the SKTI in the Japanese wild soybean lines

Source	No. of lines	Types of Davis system PAGE				IEF-PAGE
		Tia (%)[a]	*Tib* (%)[b]	*Tig* (%)	*Tia/Tib* (%)	Tib^{i5} (%)
South						
Saga	4	4				
Kagoshima	9	6	3			1
Kumamoto	2	2				
Miyazaki	5	5				
Oita	19	18	1			
Subtotal	39	35 (89.74)	4 (10.26)			1 (2.56)
Middle south						
Shimane	2	2				
Ehime	2	2				
Kochi	1	1				
Kagawa	2	2				
Tottori	1	1				
Osaka	1	1				
Nara	5	5				
Shiga	1			1		
Mie	8	5	3			1
Wakayama	6	6				
Subtotal	29	25 (86.21)	3 (10.34)	1 (3.44)		1 (3.44)
Middle						
Fukui	1		1			
Aichi	3	3				
Gifu	7	5	2			
Toyama	1	1				1
Shizuoka	1		1			
Nagano	2	2				
Yamanashi	1	1				
Saitama	38	34	4			1
Chiba	6	6				
Nigata	3	3[a]				
Ibaraki	8	4	4			2
Tokyo	6	6				
Subtotal	77	65 (84.42)	12 (15.58)			4 (5.19)
North						
Fukushima	8	8				
Miyagi	21	20	1			
Yamagata	98	89	7		2	3
Iwate	70	64	6			
Akita	328	258[a]	65	1	4	1
Aomori	13	13				
Northern Hokkaido	3	3				
Subtotal	541	455 (84.10)	79 (14.60)	1 (0.2)	6 (1.11)	4 (0.74)
Unknown	34	31	3			1
Total	720	611 (84.86)	101 (14.03)	2 (0.28)	6 (0.83)	11 (1.53)

[a] *Tia* type includes Tia^{b1}, Tia^{a1}, Tia^{a2}

[b] *Tib* type includes Tib^{i5}

The SKTI proteins were analysed by a Davis system polyacrylamide gel electrophoresis (PAGE) using 20 cm × 20 cm gel and isoelectrofocusing (IEF)-PAGE using 9 cm × 9 cm gel (Wang et al. 1996) with minor modification. IEF-PAGE was performed using a gel containing 1.2% pharmalyte (pH 4.2–4.9) and 0.4% pharmalyte

(pH 3–10). Western blot analysis of the SKTI was carried out according to the method of Wang et al. (2005).

Nucleotide sequence analysis

Gene sequence analysis was carried out according to the method of Wang et al. (2005). Genomic DNA was extracted from a seed, which is different from that extracted SKTI protein, using the DNeasy Plant Mini Kit (QIAGEN). The SKTI gene was amplified by PCR using a set of two primers (forward: 5′-TAGTCCCGATTCTCCCAACA-3′, reverse: 5′-AGTACTCTCACACTTGTGTC-3′) designed on the basis of DNA sequences of *Kti3* (= *Tia*) (Jofuku et al. 1989). After the amplified DNA was cloned using a TA Cloning Kit (Invitrogen), it was sequenced with an ABI Prism 310 Genetic Analyzer with a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). For the correctness of the sequences, DNAs isolated from four independent colonies were analysed for each plant line.

Phylogenetic tree

Nucleotide sequences were submitted to phylogenetic analysis using MEGA version 2.1 (Kumar et al. 2001), applying the neighbour-joining (NJ) method. The outgroup sequences from *G. tomentella* (DDBJ accession no. AB435659) was determined by the same method as described above. Bootstrapping with 1,000 replications for the NJ analysis was carried out.

Results

Electrophoretic polymorphism of SKTI in the Japanese wild soybean

Almost all SKTI electrophoretic mobility types of the 720 wild soybean analysed by Davis system PAGE were *Tia* (84.9%) and *Tib* (14.0%) types, except for eight lines (Table 1). The *Tib* frequency varied geographically from 10.3 to 15.6%, and the southern areas showed lower frequencies than the northern ones. Such *Tib* frequency was similar to the 10% reported in South Korea (Yu and Kiang 1993) and to the 17.6% reported in China (Wang and Li 2005). The highest frequency (19.8%) of *Tib* appeared in the Akita Prefecture. This result is consistent with that of Fujita et al. (1997), which reported a 37% frequency of *Tib* type in wild soybeans collected from Akita. The reason for the high frequency of *Tib* type in the Akita Prefecture is unknown.

Of the eight lines except for *Tia* and *Tib*, six lines showed natural-crossing type (*Tia*/*Tib*; 0.8%) and two showed a novel electrophoretic mobility (designated as *Tig*; 0.3%) (Table 1). The new variant *Tig*, which was detected in No. 210 and No. 332 lines, had a slightly slower electrophoretic mobility than *Tic* type (Fig. 2a), but its isoelectric point was more alkaline than that of *Tia*, *Tib* and *Tic* types (Fig. 2b). The new SKTI variant was also confirmed by Western blot analysis (data not shown).

Sequence polymorphism of the SKTI gene

A total of 83 randomly selected wild soybean lines (48 lines showing electrophoretic mobility of *Tia* type and 35 lines showing electrophoretic mobility of *Tib*), one novel *Tig* variant and six standard soybean cultivars were used to determine the nucleotide sequences of the SKTI gene. The analysis indicated that all 90 lines had an identical length of 743 bp, which contained an open reading frame of 651 bp encoding 217 amino acids. Six different types of allelic sequences (*Tia*, Tia^{a1}, Tia^{a2}, Tia^{b1}, *Tib* and *Tig*) were identified in the wild soybean used (Table 2, Fig. 3). Of the 48 wild soybean lines classified into *Tia* electrophoretically, 43 showed the same sequences as that of the standard *Tia* variety “Rikuu No. 27”. Sequences of the remaining five lines were different from that of *Tia*, and three variant

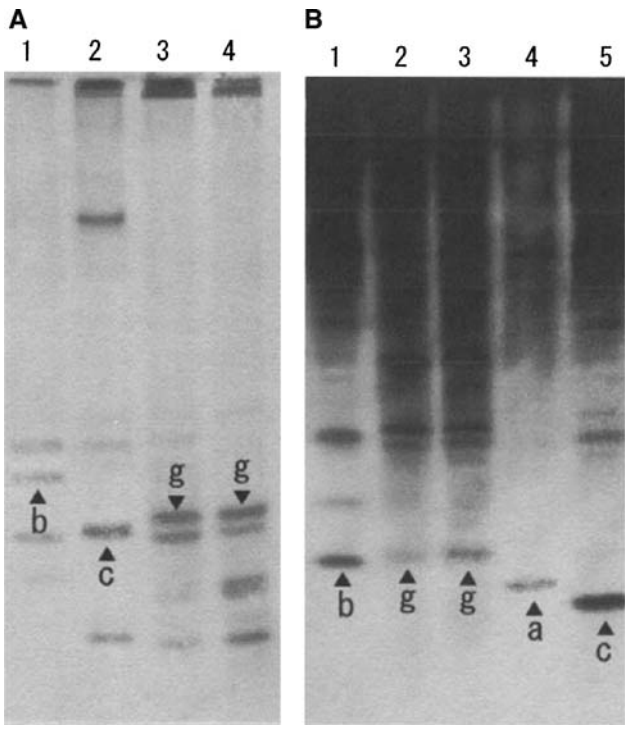

Fig. 2 Electrophoretic profiles of a novel variant (*Tig*) of SKTI proteins by a Davis system of PAGE (**a**) and a IEF-PAGE (**b**). **a** *Lane 1 G. max* “Tachisuzunari” (*Tib*), *lane 2 G. max* “Raiden” (*Tic*), *lanes 3 and 4 G. soja* “line 210” (*Tig*) and “line 332” (*Tig*), respectively. **b** *Lane 1*, “Tachisuzunari” (*Tib*), *lanes 2 and 3* “line 210” (*Tig*) and “line 332” (*Tig*), respectively, *lane 4 G. max* “Rikuu No. 27” (*Tia*), *lane 5* “Raiden” (*Tic*). A *letter* in each *lane* shows each SKTI band (*a* Tia, *b* Tib, *c* Tic, *g* Tig)

Table 2 Nucleotide and amino acid differences among nine different SKTI alleles detected in Japanese wild soybean

Alleles	Lines	Electrophoretic type	Amino acid position and genetic codon
			12 13 55 62 63 66 71 74 78 87 105 114 119 120 137 176
*Tie**	*G. soja* 1125	*Tie*	GAA-AAT-GGA-TAT-CGA-TTT-CAT-AGC-GAT-GTT-GTT-ATG-AAA CTT CCA-CTT Glu-Asn-Gly-Tyr-Arg-Phe-His-Ser-Asp-Val-Val-Met-Lys-Leu-Pro-Leu
Tia^{a1}	76	*Tia*	GAA-AAT-GGA-TAT-CGA-TTT-CAT-AGC-AAT-GTT-GTT-ATG-AGA-CTT-CCA-CTT Glu-Asn-Gly-Tyr-Arg-Phe-His-Ser-Asn-Val-Val-Met-Arg-Leu-Pro-Leu
Tia^{a2}	204, 205	*Tia*	GAA-AAT-GGA-TAT-CGA-TTT-CAT-AGC-GAT-GCT-GTT-ATG-AGA-CTT-CCA-CTT Glu-Asn-Gly-Tyr-Arg-Phe-His-Ser-Asp-Ala-Val-Met-Arg-Leu-Pro-Leu
*Tic***	L118, Riden	*Tic*	GAA-AAT-GAA-TAT-CGA-TTT-CAT-AGC-GAT-GTT-GTT-ATG-AGA-CTT-CCA-CTT Glu-Asn-Glu-Tyr-Arg-Phe-His-Ser-Asp-Val-Val-Met-Arg-Leu-Pro-Leu
Tia	Rikuu No. 27,17,25,35,44,51,52,62,86, 88,91,104,107,117,128,133,142,145,147, 165,156,169,177,183,186,195,201,203, 206,209,213,215,227,233,248,252,253, 258,302,319,320,340,358,859	*Tia*	GAA-AAT-GGA-TAT-CGA-TTT-CAT-AGC-GAT-GTT-GTT-ATG-AGA-CTT-CCA-CTT Glu-Asn-Gly-Tyr-Arg-Phe-His-Ser-Asp-Val-Val-Met-Arg-Leu-Pro-Leu
Tia^{b1}	L38, 307	*Tia*	GAA-AAT-GGA-TAT-AGA-TTT-CAT-AGC-GAT-GTT-GTC-GTG-AGA-CTT-CCA-CTT Glu-Asn-Gly-Tyr-Arg-Phe-His-Ser-Asp-Val-Val-Val-Arg-Leu-Pro-Leu
Tib^{i5}*	L24,S194,S204,S205,161-1,234,237, 250, 249,336-1, 757	Tib^{i5}	GAA-AAT-GGA-TTT-CGA-TTT-CAT-AGA-GAT-GTT-GTC-GTG-AGA-ATT-ACA-CTT Glu-Asn-Gly-Phe-Arg-Phe-His-Arg-Asp-Val-Val-Val-Arg-Ile-Thr-Leu
Tib	Odate No.1,Tachisuzunari,,Norin No. 2, K47,4,9,15,19,20,21,42,71,72,73,74,81, 87,92,171,189,190,199,210-1,217,231,277, 278,280,281,321,322,323,324,325,328, 336,337,338	*Tib*	GAT-AGT-GGA-TTT-CGA-TTT-AAT-AGA-GAT-GTT-GTC-GTG-AGA-ATT-ACA-GTT Asp-Ser-Gly-Phe-Arg-Phe-Asn-Arg-Asp-Val-Val-Val-Arg-Ile-Thr-Val
Tig	210	*Tig*	GAT-AGT-GGA-TTT-CGA-CTT-AAT-AGA-GAT-GTT-GTC-GTG-AGA-ATT-ACA-GTT Asp-Ser-Gly-Phe-Arg-Leu-Asn-Arg-Asp-Val-Val-Val-Arg-Ile-Thr-Val

Letters highlighted indicate nucleotides and amino acids that are different from those of *Tia*

* Previously reported alleles in Japanese wild soybean (Wang et al. 2001, 2005)

** Cultivated soybeans

types, which were designated as Tia^{a1}, Tia^{a2} and Tia^{b1}, were identified. The sequence of Tia^{a1} (line 76) and Tia^{a2} (lines 204 and 205) differed from *Tia* by a G → A mutation at position +376 and a T → C mutation at +404, respectively. These mutations resulted in amino acid changes at the 78th residue in Tia^{a1} (Asp Asn) and at the 87th residue in Tia^{a2} (Val Ala; Table 2). The sequence of Tia^{b1} (lines L38 and 307) differed from *Tia* by three nucleotides; C →A at position +331, T → C at +459 and A → G at +484 (Table 2, Fig. 3). Of these three mutations, the last one resulted in substitution of an amino acid from Met to Val at the 114th residue, and the other two mutations caused synonymous codons at the residue positions 63 and 105. The three variant types showed indistinguishable electrophoretic behaviours from *Tia* type not only on Davis system PAGE, but also on IEF-PAGE (data not shown).

All 35 wild soybean lines classified into *Tib* electrophoretically showed the same sequences as that of the standard *Tib* varieties. The analysis of *Tig* gene (line No. 210) revealed that nucleotide sequences of *Tig* were identical to those of *Tib* except for one T → C transitional mutation at position +340 (Fig. 3, Table 2). This mutation caused substitution of an amino acid from Phe in *Tib* to Leu in *Tig* at the 66th residue.

Structural features of SKTI gene

A total of nine alleles of SKTI (*Tia*, *Tib*, *Tic*, Tia^{a1}, Tia^{a2}, *Tig*, Tia^{b1}, *Tie* and Tib^{i5}), which were identified in the present and previous studies, were compared with each other (Tables 2, 3). According to NJ tree based on sequence structures, these alleles were grouped into three categories, *Tia*-type (*Tia* allele and *Tia*-derived *Tic*, Tia^{a1}, Tia^{a2} and *Tie*), *Tib*-type (*Tib* allele and *Tib*-derived *Tig*) and transitional type (Tia^{b1} and Tib^{i5}; Fig. 4). Each of the four alleles classified into the *Tia* category showed a point mutation at different positions on the *Tia* gene, which caused substitution of an amino acid. Of the transitional types, Tib^{i5} has been reported to have six different nucleotides from *Tia*, and the mutated nucleotides have been reported to be identical with *Tib* (Wang et al. 2005). Tia^{b1} detected in this study had three different nucleotides from *Tia*, two of which are identical with *Tib* and Tib^{i5} (Table 2, Fig. 4).

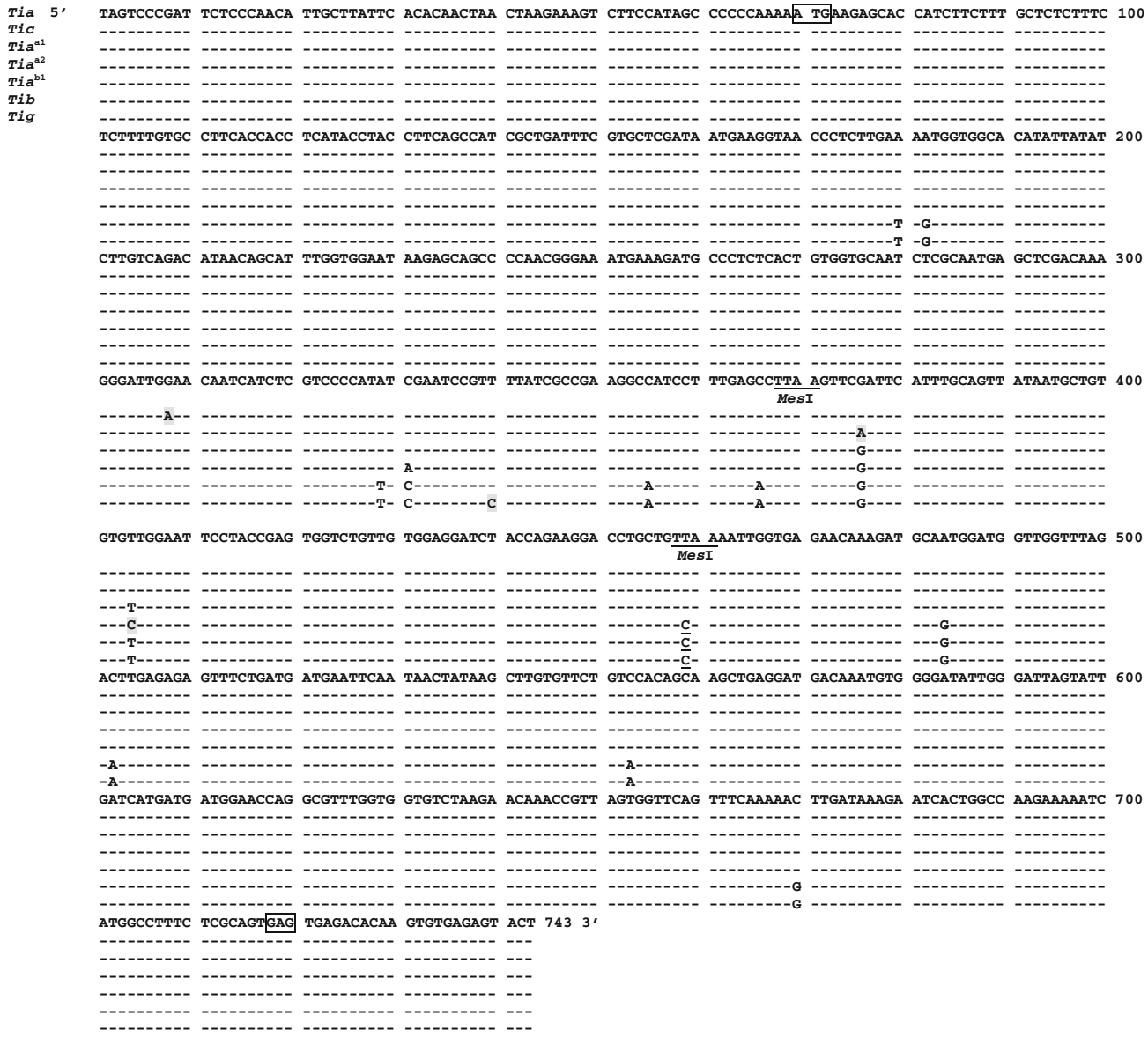

Fig. 3 Nucleotide sequences of four new SKTI genes, *Tia*^{a1} (DDBJ accession no. AB308132), *Tia*^{a2} (DDBJ accession no. AB308133), *Tia*^{b1} (DDBJ accession no. AB308134) and *Tig* (DDBJ accession no. AB308131), in Japanese wild soybean. *Highlighted Letters* indicate the nucleotides that caused amino acid changes in comparison with *Tia* or *Tib*. *Boxes* indicate the start codon and the end one of the translation. *Underlines* show *Mes* I restriction sites for PCR-RFLP discriminating *Tib* and transitional type from *Tia* type

Tib- and transitional types could be discriminated from *Tia* type by PCR-RFLP using a restriction enzyme *Mse* I (recognition site T ↓ TAA). *Tia* type has T at +459, but *Tib*- and transitional types mutated to G from T at this position. This mutation causes disappearance of the *Mse* I restriction site at this position. After PCR-RFLP with *Mse* I, *Tia*-type alleles had the two restriction sites at +377 and +457 positions (Fig. 3), resulting in three fragments of a 368, 285 and 90 bp (Fig. 5). On the other hand, *Tib*- and transitional types had no restriction site at +457, resulting in two fragments of 368 and 375 bp (Fig. 5).

Discussion

The present study revealed that there are four new polymorphic alleles (*Tia*^{a1}, *Tia*^{a2}, *Tia*^{b1} and *Tig*) in SKTI locus, and that gene sequence analysis is effective for detection of polymorphism of SKTI, though its polymorphism cannot be detected by conventional electrophoretic analysis of SKTI proteins.

Including the four new variants, a total of 11 allelic polymorphisms at the SKTI locus have been detected in the subgenus *Soja*. Of these, only *Tid* was found in a cultivated variety (Zhao and Wang 1992; Xin et al. 1999). *Tia*, *Tib*

Table 3 Features of gene sequences and nucleotide mutation positions in contrast with *Tia* type in Japanese wild soybean

Allele	Category of SKTI	Length of sequences aligned	No. of substitutive amino acids to *Tia*	Mutational position against *Tia* type
Tia	A	743	–	
Tic	A	743	1	308
Tia^{a1}	A	743	1	376
Ta^{a2}	A	743	1	404
*Tie***	A	743	1	500
Tia^{b1}	T	743	1	331*, 459*, 484
Tib^{i5}**	T	743	5	329, 366, 459*, 484, 502, 553
Tib	B	743	9	180, 182, 329,355, 366, 459*, 484, 502, 553, 670
Tig	B	743	10	180, 182, 329, 340, 355, 366, 459*, 484, 502, 553, 670

* This mutation caused a synonymous change in codon

** Quoted from previous papers (Wang et al. 2001, 2005)

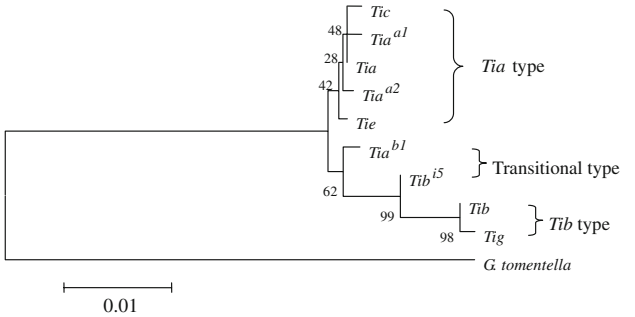

Fig. 4 Phylogenetic tree of 9 SKTI variants constructed by the neighbour-joining (NJ) method using MEGA version 2.1 program based on their nucleotide sequences. That of *G. tomentella* is used as an outgroup (DDBJ accession no. AB435659). *Numbers above branches* are bootstrap values (%)

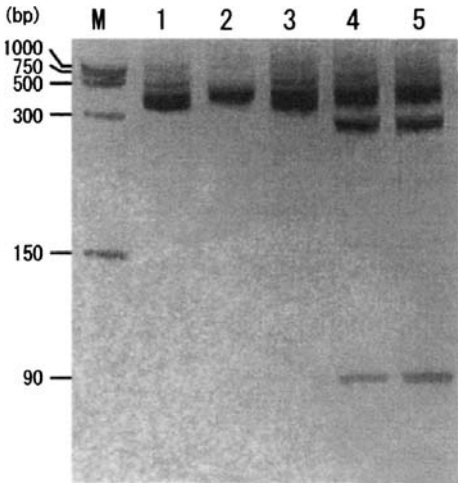

Fig. 5 PCR-RFLP profiles by *Mes* I restriction digest of the SKTI genes. *Lane 1 G. max* "Tachisuzunari" (*Tib*), *lane 2 G. soja* "line L38" (Tia^{b1}), *lane 3 G. soja* "line L24" (Tib^{i5}), *lanes 4 and 5 G. max* "Rikuu No. 27" (*Tia*). *Lanes 1–3* show two fragments of 368 and 375 bp, and *lanes 4 and 5* show three fragments of 368, 285 and 90 bp

and *Tic* are common in both wild and cultivated soybeans, and the others, *Tie* (Wang et al. 2001), *Tif* (Wang et al. 2004), Tib^{i5} (Wang et al. 2005), Tia^{a1}, Tia^{a2}, *Tig* and Tia^{b1} were detected in wild soybean. The higher level of SKTI polymorphism could be related to the function of SKTI protein, whose bio-function may be replaced by other similar functional proteins (Birk 1961, 1963; Frattali and Steiner 1968; Rachis and Anderson 1964; Yamamoto and Ikenaka 1967). Christeller (2005) explained the evolutionary mechanism occurring in such diversity of proteinase inhibitors.

The mutations or absence of SKTI in plants do not have an influence on its growth and development. Orf and Hymowitz (1979) reported that the cultivated variety lacking SKTI was able to grow and develop normally.

Table 4 Inferable differentiation route of *Tib* from prototype *Tia* of SKTI proteins

Gene	Amino acid position									
	12	13	62	71	74	105	114	120	137	176
Tia	Glu	Asn	Tyr	His	Ser	Val (GTT)	Met	Leu	Pro	Leu
						↓	↓			
Tia^{b1}	Glu	Asn	Tyr	His	Ser	Val (GTC)	Val	Leu	Pro	Leu
			↓		↓			↓	↓	
Tib^{i5}	Glu	Asn	Phe	His	Arg	Val (GTC)	Val	Ilu	Thr	Leu
	↓	↓		↓						↓
Tib	Asp	Ser	Phe	Asn	Arg	Val (GTC)	Val	Ilu	Thr	Val

Arrows indicate the substitutions of nucleotide or amino acids

Our results showed that some single nucleotide mutations caused one single amino acid substitution in *Tia* or *Tib*, which engendered rather high polymorphism of SKTI. On the other hand, there is a large difference of nine amino acid substitutions between *Tia* and *Tib* (Song et al. 1993; Wang et al. 2004). We considered that *Tib* is differentiated from *Tia* via repeated mutations, but intermediate types between them have not been found. Recently, we reported the intermediate type, Tib^{i5}, which shows a five-amino acid substitution between *Tia* and *Tib* types (Wang et al. 2005). The detection of Tia^{b1} provides an evidence to explain the mystery of how *Tia* type had changed to *Tib* type, because Tia^{b1} has an intermediate sequence trait between *Tia* and Tib^{i5}. Of the nine amino acids that differ between *Tia* and *Tib*, one amino acid at the 114th residue (Val) in Tia^{b1} was different from that of *Tia* (Met), but identical to that of *Tib* and Tib^{i5} (Table 4). In addition, in the codon of the amino acid at 105th (Val), which is synonymous between *Tia* (GTT) and *Tib* or Tib^{i5} (GTC), the codon of Tia^{b1} is the same as that of *Tib* and/or Tib^{i5}. These results suggest that *Tib* had gradually differentiated through Tia^{b1} and Tib^{i5} from *Tia* (Table 4). As mentioned above, *Tia* is considered to be hypothetically the prototype. NJ tree partly supports this hypothesis, because *Tia* type is more closely related to an outgroup.

In wild soybean, a majority of SKTI type was *Tia*, followed by *Tib* type (Hymowitz 1973; Kaizuma et al. 1980). In this study, we have observed 84.1% *Tia* type and 14.03% *Tib* type. However, *Tib* type sometimes exceeded 20%, and even 50% in different natural populations of wild soybean (Fujita et al. 1997; Xu et al. 1985). Kaizuma et al. (1980) pointed out that the genetic differentiation of *Tia* and *Tib* had occurred before the domestication of soybean from wild soybean, because high *Tib* frequency in wild soybean could not have resulted from cultivated *Tib*-type soybeans through introgressive hybridisation between the two species. However, direct and concrete proofs for this idea have not been presented. The findings of the transitional Tia^{b1} and Tib^{i5} types in wild soybean suggest that the *Tib* type of cultivated soybeans had originated directly from the *Tib* type of wild soybean along with soybean domestication. Although *Tic* has been reported in Chinese wild soybean (Xu et al. 1985), whether the rare *Tic* type of cultivated soybeans also had originated from wild soybean cannot be concluded, since we have not found any *Tic* type in more than 5,000 wild soybean accessions (Wang and Li 2005; this study). There is a possibility that *Tic* in the cultivated soybean had differentiated after soybean domestication.

Tia^{a1}, Tia^{a2}, *Tie*, *Tig*, Tia^{b1} and Tib^{i5} alleles were all identified in Japanese wild soybean. These results suggest that SKTI gene is rather highly polymorphic in the wild soybean in Japan. The finding of two transitional variants Tia^{b1} and Tib^{i5} in Japanese wild soybean suggests that certain stages of differentiation process between *Tia* and *Tib* alleles had occurred in Japan. Studies on isozyme (Hirata et al. 1999), chloroplast DNA SSRs (Xu et al. 2002) and nuclear SSRs (Abe et al. 2003) propose multiple origins of cultivated soybean in eastern Asia, especially Japan and China. Abe et al. (2003) suggest that each of the Japanese and Chinese populations form different germplasm pools of soybean. In this study, Chinese wild soybean, which is extensively distributed with a large variation and a vast genetic diversity, has not been examined. Further investigation of a large number of wild and cultivated soybean derived from eastern Asia, such as China, Japan and Korea, would provide the additional information on polymorphisms and molecular evolution of SKTI gene, including a serial of evolutionary transitional genes between *Tia* and *Tib* types, and the origin of cultivated soybean.

Acknowledgments W. K. was financially supported by a Postodoctoral Fellowship for Foreign Researchers from the Japan Society for the Promotion of Science (JSPS).

References

Abe J, Xu DH, Suzuki Y, Kanazawa A, Shimamoto Y (2003) Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor Appl Genet 106:445–453

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

Birk Y, Gertler A, Khalef S (1963) A pure trypsin inhibitor from soya beans. J Biochem 87:281–282

Christeller JT (2005) Evolutionary mechanisms acting on proteinase inhibitor variability. FEBS J 272:5710–5722

Frattali V, Steiner RF (1968) Soybean inhibitor. 1. Separation and some properties of three inhibitors from commercial crude soybean trypsin. Biochemistry 7:521–531

Fujita R, Ohara M, Okazaki K, Shimamoto Y (1997) The extent of natural cross-pollination in wild soybean. J Heredity 88:124–128

Hirata T, Abe J, Shimamoto Y (1999) Genetic structure of the Japanese soybean population. Genet Res Crop Evol 46:441–453

Hu ZA, Wang HX (1985) Genetic structure of natural population wild soybean (*Glycine soja*) on the Beijing region. Acta Bot Sin 27:599–604

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

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

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

Hymowitz T, Kaizuma N (1981) Soybean seed protein electrophoresis profiles from 15 Asian countries or regions: hypotheses on paths of dissemination of soybean from China. Econ Bot 35:10–23

Jofuku KD, Schipper RD, Goldberg RB (1989) A frameshift mutation prevents Kunitz trypsin inhibitor mRNA accumulation in soybean embryos. Plant Cell 1:427–435

Kaizuma N, Oikawa K, Miura M (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 Agric Iwate Univ 15:81–96

Kiang YT, Chang YC, Kaizuma N (1992) Genetic diversity in natural population of wild soybean in Iwate Prefecture, Japan. J Heredity 83:325–329

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

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Lee SI, Lee SH, Koo JC, Chun HJ, Lim CO, Mun JH, Song YH, Cho MJ (1999) Soybean Kunitz trypsin inhibitor (SKTI) confers resistance to the brown planthopper (*Nilaparvata lugens* Stal) in transgenic rice. Mol Breed 5:1–9

Li FS (1993) Studies on the ecological and geographical distribution of the Chinese resources of wild soybean (*G. soja*). Sci Agri Sin 26:47–55

Marchetti S, Delledonne M, Fogher C, Chiaba C, Chiesa F, Savazzini F, Giordano A (2000) Soybean Kunitz, C-II and PI-IV inhibitor genes confer different levels of insect resistance to tobacco and potato transgenic plants. Theor Appl Genet 101:519–526

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

Rachis JJ, Anderson RL (1964) Isolation of four soybean trypsin inhibitors by DEAE-cellulose chromatography. Biochem Biophys Res Commun 15:230–235

Singh LC, Wilson M, Hadley HH (1969) Genetic differences in soybean trypsin inhibitor separated by disc electrophoresis. Crop Sci 9:489–491

Song SI, Kim CH, Baek SJ, Choi YD (1993) Nucleotide sequences of cDNAs encoding the precursors for soybean (*Glycine max*) trypsin inhibitors (Kunitz type). Plant Physiol 101:1401–1402

Tian QZ, Gai JY, Yu DY, Jia JZ (2000) A study on amplified fragment length polymorphism (AFLP) in soybean. Soybean Sci 19:210–217

Tozuka A, Fukushi H, Hirata T, Ohara M, Kanazawa A, Mikami T, Abe J, Shimamoto Y (1998) Composite and clinal distribution of *Glycine soja* in Japan revealed by RFLP analysis of mitochondrial DNA. Theor Appl Genet 96:170–176

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

Wang KJ, Kaizuma N, Takahata Y (1998) The polymorphism and geographic distributions of four seed proteins and two traits in Japanese natural populations of wild soybean (*Glycine soja*). Acta Agron Sinica 24:147–154

Wang KJ, Takahata Y, Ito K, Zhao YP, Tsutsumi KI, Kaizuma N (2001) Genetic characterization of a novel soybean Kunitz trypsin inhibitor. Breed Sci 51:185–190

Wang KJ, Yamashita T, Watanabe M, Takahata Y (2004) Genetic characterization of a novel *Tib*-derived variant of soybean Kunitz trypsin inhibitor detected in wild soybean (*Glycine soja*). Genome 47:9–14

Wang KJ, Li XH (2005) *Tif* type of soybean Kunitz trypsin inhibitor exists in wild soybean of northern China. Proceedings of the 8th national soybean research conference of China, pp 167–168

Wang KJ, Takahata Y, Shimada S, Kaizuma N (2005) Entity evidence for differentiation between *Tia* and *Tib* types of soybean Kunitz trypsin inhibitor: Detection of a novel transitional variant types between *Tia* and *Tib* in wild soybean (*Glycine soja*). Theor Appl Genet 112:67–71

Xin H, Xie KF, Dong AW, Yan QY, Gu QM (1999) The amino acid sequence determination of a new variant of Kunitz soybean trypsin inhibitor (SBTi-A2). Soybean Genetics Newsletter (online journal) www.soygenetics.org/articles/sgn 1999–004.html (posted 24 March 1999)

Xu B, Zhao SW, Zou SH, Zhen HY, Hu A, Wang HX (1985) Seed protein electrophoresis profiles of wild soybean (*G. soja*) in China:The frequencies and geographical distribution of *Ti* and *Sp1* alleles and the hypothesis on the original area of soybean. Soybean Sci 4:7–13

Xu DH, Abe J, Gai Y, Shimamoto Y (2002) Diversity of chloroplast DNA SSRs in wild and cultivated soybeans:evidence for multiple origins of cultivated soybean. Thore Appl Genet 105:645–653

Yamamoto M, Ikenaka T (1967) Studies on soybean trypsin inhibitors. Purification and characterization of two soybean trypsin inhibitors. J Biochem 62:141–149

Yu H, Kiang YT (1993) Genetic variation in South Korean natural populations of wild soybean (*Glycine soja*). Euphytica 68:213–221

Zhao SW, Wang H (1992) A new electrophoretic variant of SBTi-A2 in soybean seed protein. Soyb Genet Newsl 19:22–24