Indian J. Genet., 71(4): 372-376 (2011)

Short Communication

Marker-assisted selection for development of kunitz trypsin inhibitor-free soybean varieties: I. Parental polymorphism survey using SSR markers

Vineet Kumar*, Anita Rani, Vaishali Mourya, Reena Rawal, Khushbu Verma[1], M. Shivakumar[1], S. K. Lal[1] and Akshay Talukdar[1]

Directorate of Soybean Research, Khandwa Road, Indore 452 001, Madhya Pradesh; [1]Division of Genetics, Indian Agricultural Research Institute, New Delhi 110 012

(Received: July 2011; Revised: September 2011; Accepted: October 2011)

Soybean is regarded as “miracle bean” due to its numerous uses as food, feed and health supplement. However, acceptance of soy foods has been restricted by the presence of kunitz trypsin inhibitor (KTI) in soybean seeds. Kunitz trypsin inhibitor also known as SBTI-A2 which constitutes 80% of the total trypsin inhibitor activity, has been shown to be responsible for growth inhibition, pancreatic hypertrophy and hyperplasia in experimental animals [1-3]. Though KTI is heat labile, however, heat treatment is not only cost ineffective but also results in approximately 20% decline in protein solubility [4]. Similarly, boiling of beans prior to grinding with wheat (1:9) is essential while preparing soy-supplemented *chapatti* flour. Therefore, development of KTI-free soybean varieties which are not yet available in India is expected to boost utilization of soybean in food products.

Genetic studies revealed that KTI trait in soybean is controlled by multiple alleles namely, Ti^a [5], Ti^b [5], Ti^c [6] and Ti^d [7] at a single locus. These four electrophoretic forms of soybean KTI are controlled by co-dominant multiple allelic series (Ti^a, Ti^b, Ti^c and Ti^d). A fifth form lacking kunitz trypsin inhibitor activity is controlled by a recessive allele *ti* [8]. The gene has been located on the linkage group (LG) A2 of the soybean molecular linkage map [9]. Three SSR markers Satt409, Satt228 and Satt429 have been reported to be tightly linked with *ti* locus at a distance of 4.5, 0-3.7 and 5.1 cM, respectively [10]. Soybean genotype (PI542044) which is free from KTI can effectively be utilized to transfer *ti* allele to elite soybean varieties through marker-assisted backcross selection (MABS). Hence, five elite soybean varieties *viz*., NRC7, JS97-52, MACS450, DS93-05 and DS9712 were crossed with PI542044 to introgress the *ti* allele through marker-assisted selection (MAS). This necessitates selection of the target plants with homozygous null allele of KTI (foreground selection) with maximum recovery of recurrent parent genome (background selection). Therefore, the objective of the present study was to assess the level of polymorphism at SSR loci between the donor and the five recurrent parents.

Genomic DNA was extracted from young leaves of the six genotypes following CTAB (cetyl trimethyl ammonium bromide) procedure [11]. Purified DNA was subjected to PCR amplification in 10 µl reaction mixture containing 2 µl DNA (20 ng/ µl), 1 µl PCR 10x buffer, 1.1 µl $MgCl_2$ (25 mM), 0.1 µl dNTPs (25 mM), 0.4 µl each forward and reverse SSR primers (30 ng/ µl), 0.068 µl *Taq* DNA polymerase (3U/ µl) and 4.932 µl double-distilled water. In the thermocycler (MJ Research Thermocycler model PTC100) the DNA was denatured at 94^{o}C for 2 min. followed by 30 cycles each consisting of denaturation at 94^{o}C for 1 min., primer annealing at

*Corresponding author’s e-mail: vineetksahni@yahoo.com
Published by Indian Society of Genetics & Plant Breeding, F2, First Floor, NASC Complex, PB#11312, IARI, New Delhi 110 012
Online management by indianjournals.com

50°C for 2 min. primer elongation at 72°C for 3 min. and final elongation at 72°C for 10 min. Amplified products so obtained were resolved on 3% metaphore agarose gel stained with ethidium bromide and analyzed in GeneGenius Gel Imaging System from Syngene. For polymorphism study, 484 SSR markers were picked up from the 20 linkage groups [12] at an average density of one SSR marker per 5 cM. For validation and testing utility in foreground selection, the linked markers i.e. Satt228 and Satt409 were tested for its polymorphism in the five parental combinations *i.e.*, (NRC7, JS97-12, MACS450 and DS9712) x PI542044. Each of the five parental combinations exhibited distinct polymorphism for both the linked SSR markers (Fig. 1) indicating its applicability in foreground selection including testing of hybridity of the F_1 plants as well as identification of plants with null KTI allele in different generations.

For assessing the level of polymorphism between different parental combinations as well as to select SSR markers for background selection, 484 SSR markers spanning across 20 linkage groups (LG) were used. The level of polymorphism ranged from 43.38% (NRC7 x PI542044) to 48.14% (MACS450 x PI542044) (Table 1). Thus, 210 to 233 polymorphic SSR markers (10-11 markers/chromosome) left for background selection.

Level of polymorphism in Indian soybean genotype is low primarily because, it is not native to India. Secondly, Indian soybean germplasm consists mainly of introductions from China and USA. Most serious is the fact that Indian breeders have been using only a few selected genotypes in hybridization programmes thereby narrowing down the genetic base of the cultivated germplasm. Low polymorphism has also been detected in other experiments too [13, 14]. Therefore, use of diverse parents in crossing programme is advocated to generate higher level of polymorphism.

Table 1. Parental polymorphism survey of the 5 parental combination using 484 SSR markers

Cross combination	No. of mono-morphic markers	No. of polymorphic markers*
NRC7 x PI542044	274	210 (43.38%)
JS97-52 x PI542044	266	218 (45.04%)
MACS450 x PI542044	251	233 (48.14%)
JS93-05 x PI542044	271	213 (44.00%)
DS9712 X PI542044	259	225 (46.48%)

*Values in the parenthesis indicate the percentage of the total 484 SSR markers

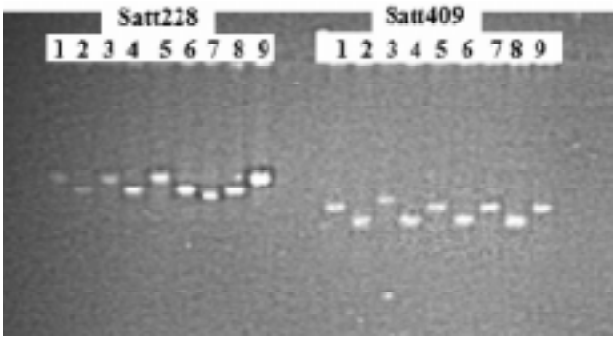

Fig. 1. Parental polymorphism for linked markers Satt228 and Satt409. Lanes (left to right) 1, 3, 5, 7 and 9 indicate PCR amplified products using genomic DNA of NRC7, JS97-52, JS93-05, MACS450 and DS9712; while lanes 2, 4, 6 and 8 indicate for PI542044 (donor of null allele of kunitz trypsin inhibitor) in case of reported linked markers Satt228 and Satt409.

While calculating the number of polymorphic markers, it was observed that 74 marker loci were polymorphic across the five cross-combinations (Table 2). However, distribution of the polymorphic loci was not uniform; it varied from one LG to the other as well as one cross-combination to the other. Highest number of polymorphic markers (11) across all the five crossing combinations was detected on linkage group F, while the least number (1) was observed on LG B1 and LG C1 (Fig. 2). Non-uniform distribution of polymorphic SSR loci has also been reported elsewhere [15]. This phenomenon reflects occurrence of conserved sequences on chromosomes across different genotypes leading to detection of reduced level of polymorphism. Therefore, a sizable number of SSR markers (at least one in every 5 cM) need to be tested for such polymorphism so as to ensure that at least one polymorphic marker is retained per 10 cM across the genome. Markers of such density can ensure efficient background screening in the segregating plants. The 74 polymorphic markers identified here has been used for background selection of the candidate plants *i.e.* plants having *titi* allele.

The number of polymorphic markers observed in each of the five cross combinations has been depicted on Table 3. Presence of polymorphic SSR markers was found to vary in different linkage groups in different crossing combinations. Linkage group C2 was detected with maximum number of polymorphic SSR markers (30) for the two parental combination *viz.*, JS97-52 x PI542044 and DS9712 x PI542044. It is thus indicated

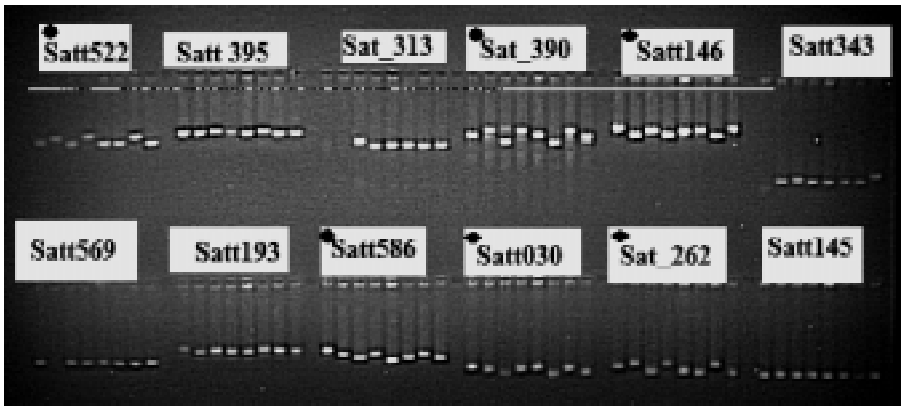

Fig. 2. Parental polymorphism using SSR markers from linkage group F. SSR markers labelled with * are polymorphic for all the 5 parental combinations. Lanes 1-8 (from left to right) for each of the SSR markers indicates amplified products of NRC7, PI542044, JS97-52, PI542044, JS93-05, PI542044, MACS450, PI542044 and DS9712, respectively

Table 2. SSR markers from different LGs revealed to be polymorphic across all the five soybean parental combinations (NRC7xPI542044; JS97-52xPI542044; MACS450xPI542044; JS93-05 xPI542044; DS9712xPI542044)

LG	Polymorphic SSR marker(s)	SSR markers
A1	2	Satt276, Satt200
A2	6	Sat_406, Sat_409, Sat_181, Satt409, Satt158, Satt228
B1	1	Satt453
B2	2	Sat_230, Sat_009
C1	1	Satt_140
C2	6	Satt227, Satt100, Sat_402, Sat_062, Satt289, Satt460
D1a	2	Sat_332, Sat_413
D1b	6	Sat_096, BE-475343, Satt005, Satt600, Satt217, Satt459
D2	2	Satt372, Satt002
E	3	Satt691, Satt720, Satt553
F	11	Satt144, Satt522, Sat_390, Satt146, Satt586, Satt030, Sat_262,Satt114, Satt657, Sat_417, Satt335
G	4	Satt612, Satt191, Satt199, Sat_290
H	2	Sat_205, Sat_216
J	2	Satt674, Sat_366
K	8	Satt178, Satt555, Sat_043, Satt552, Sat_087, Satt539, Satt375, SoyPRP1
L	3	Satt143, Sat_286, Sat_113
M	2	Sat_244, Satt175
N	5	Satt152, Satt624, Sat_084, Satt080, Satt387
O	6	Sat_318, Satt653, Satt420, Satt478, Satt477, Sat_109

that different number of SSR markers has to be used for background selection in different cross combinations; however, density of at least one SSR marker per 10 cM should be maintained for efficient background selection.

Polymorphic markers in the 11 linkage groups *viz.*, B2, D1b, D2, E, F, G, H, J, K, L and O were within the range of 5 cM approx. across all the parental combinations. However, certain regions on rest of the

Table 3. SSR markers found to be polymorphic on different linkage group with respect to each of the five parental combinations

LG	NRC7 x PI542044	JS97-52 x PI542044	MACS450 x PI542044	JS93-05 x PI542044	DS9712 x PI542044
A1	9	11	11	12	10
A2	9	13	12	10	13
B1	5	5	7	7	9
B2	10	10	7	9	9
C1	6	6	10	6	5
C2	20	30	25	17	30
D1a	7	7	13	8	4
D1b	19	18	16	16	17
D2	7	11	12	9	9
E	8	6	9	7	8
F	29	26	33	28	28
G	12	11	8	14	14
H	9	8	7	8	8
I	5	4	10	4	3
J	5	4	8	8	7
K	10	14	11	10	13
L	9	8	9	9	8
M	8	5	4	9	8
N	10	7	8	10	10
O	13	14	12	12	12

linkage groups were devoid of polymorphic markers in longer stretches (Table 4). For even representation of the genome, more SSR markers from these regions have to be analyzed. Similar is the situation in JS97-52 x PI542044 and JS93-05 x PI542044 (Table 4). However, marker distribution in MACS450 x PI542044 was nearly uniform and was without any large intervening gap between any two consecutive polymorphic SSR markers.

For incorporating new SSR markers in poorly polymorphic regions of the genome, two approaches may be followed. The first approach would be to select SSR markers from already available resources, like map published in 2008 by USDA and test for polymorphism and, the second approach would be to design new SSR marker from the clonal sequences of those regions. In this case, the physical map as well as the complete genome sequence of soybean [16] would facilitate clone identification and designing of new primers.

Detection of polymorphism often depends upon the type of motif (di, tri, tetra, etc.) in the SSRs as well as in the number of repeats of the motif (n = 5, 10, 15 etc.). In the present study, it has been observed that

Table 4. Intervening regions between two consecutive polymorphic SSR markers on different linkage groups (LG) for different parental combinations in soybean

LG	NRC7 x PI542044	JS97-52 x PI542044	JS93-05 x PI542044	DS9712 x PI542044
A1	Satt200 – Sat_267 (30.93 – 78.45 cM)		Satt385 – Sat_267 (31.07 – 78.45 cM)	Satt200 – Sat_267 (30.93 – 78.45 cM)
A2	Sat_181 – Satt329 (38.06 – 110.94 cM)		Satt089 –Satt158 (87.57 – 115.25 cM)	Sat_181 - Sat_089 (38.06– 87.57 cM)
B1	Satt197 – Satt359 (46.39 – 102.5 cM)	Satt197 – Satt453 (46.39 – 123.96 cM)	i) BE806308 – Satt509 (0.00 – 32.51 cM) ii) Sat_149 – Satt359 (54.01 – 102.56 cM)	
C1		Satt713_Satt682 (88.95 - 127.06 cM)		
C2	Sat_062 – Satt305 (30.8 - 69.67 cM)		Sat_153 – Sat_076 (61.98 - 99.18 cM)	
D1a	Sat_413- Satt603 (5.932 - 54.50 cM)		Sat_413- Satt532 (5.932 - 49.07 cM)	i) Sat_413- satt502 (5.932 - 49.84 cM) ii) Satt502- Satt129 (49.84 - 109.67 cM)
H		Vo8405- Sat_118 (7.26 - 61.86 cM)		
I	Sat_170 – Sct_189 (75.00 - 113.77 cM)		Satt354- Sct_189 (46.22- 113.77 cM)	Satt367- Sat_104 (27.98 – 65.62 cM)
M	Satt697- Satt308 (85.35 - 130.76 cM)	Satt697- Satt308 (85.35 - 130.76 cM)		

SSR motif with tri-nucleotide repeats were more polymorphic than others. Out of 74 polymorphic markers, 45 were with tri-nucleotide repeats while 29 were with di-nucleotide repeats. Among the di-nucleotide repeats, SSR motif with 13-32 repeats i.e. $(AT)_{13-32}$ appeared to be more polymorphic than others. In case of tri-nucleotide repeats, motifs with 9-23 repeats i.e. $(ATT)_{9-23}$ detected more polymorphism among the five cross combinations. Similar results in soybean have already been reported [17]. This information is useful in picking up and/or designing SSR primers in future research.

It thus could be concluded from above study that Indian soybean genotypes are less diverse and hence necessitates screening of more SSR markers for its polymorphism study. The polymorphic loci are not uniformly distributed across the genome and vary from one genotype to the other needing different sets of markers for different cross-combinations. The level of polymorphism detected by different SSR markers depends upon the motifs size and number of repeats of a motif. These principles would be helpful in precision molecular breeding in soybean.

Acknowledgements

This work is supported by a financial grant from Department of Biotechnology, Govt of India for the project “Marker Assisted selection for development of Kunitz trypsin inhibitor free soybean varieties” under the aegis of Accelerated Programme on Crop Improvement.

References

1. **Liener I. E. and Kakade M. L.** 1980. Protease inhibitors. In: Toxic constituents of Plant Foodstuffs, Liener, I.E. (ed.) 2nd edn. Academic Press Inc., New York, London, pp 7-72.
2. **Furukawa F., Toyoda K., Aba H., Hasegawa R., Sato H., Takahashi M. and Hayashi Y.** 1987. Short-term effects of feeding crude soybean trypsin inhibitor on pancreas of rat, hamster and mouse. British National Institute of Hygiene Science (Tokyo), pp. 46-60.
3. **Ge Y. C. and Morgan R. G. H.** 1993. The effect of trypsin inhibitor on the pancreas and small intestine of mice. British J. Nutr., **70**: 333-345.
4. **Anderson R. L.** 1992. Effect of steaming on soybean proteins and trypsin inhibitors. J. Am. Oil Chem. Soc., **69**: 1170-1176.
5. **Singh L. C., Wilson M. and Handley H. H.** 1969. Genetic differences in soybean trypsin inhibitor separated by disc electrophoresis. Crop Sci., **9**: 489-491.
6. **Hymowitz T.** 1973. Electrophoretic analysis of SBTI-A2 in the USDA soybean germplasm collection. Crop Sci., **13**: 420-421.
7. **Zhao S. W. and Wang H.** 1992. A new electrophoretic variant of SBTi-A2 in soybean seed protein . Soyb. Genet. Newsl., **19**: 22-24.
8. **Orf J. H. and Hymowitz T.** 1979. Inheritance of the absence of the Kunitz trypsin inhibitor in seed protein of soybeans. Crop Sci., **19**: 107-109.
9. **Cregan P. B., Jarvik T., Bush A. L., Shoemaker R. C., Lark K. G., Kahler A. L., Kaya N, VanToai T. T., Lohnes D. G., Jongil C. and Specht J. E.** 1999. An integrated genetic linkage map of the soybean genome. Crop Sci., **39**: 1464-1490.
10. **Kim M. S., Park M. J., Jeong W. H., Nam K. C. and Chung J.** 2006. SSR markers tightly linked to the *ti* locus in soybean [*Glycine max* (L.) Merr.]. Euphytica, **152:** 361-366.
11. **Doyle J. J. and Doyle J. L.** 1990. Isolation of plant DNA from the fresh tissue. Focus, **12**: 14-15.
12. **Anonymous USDA.** 2008. bldg6.arsusda.gov/cregan/soymap.htm
13. **Singh R. K., Raipuria R. K., Bhatia V. S., Rani A., Pushpendra, Husan S. M., Chauhan D., Chauhan G. S. and Mohapatra T.** 2008. SSR markers associated with seed longevity in soybean. Seed Sci. Tech., **36**(1): 162-167.
14. **Talukdar Akshay**, **Verma K., Gowda D. S. Samrat, Lal S. K., Sapra R. L., Singh R., Singh K. P. and Sinha P.** 2009. Molecular breeding of charcoal rot in soybean. I. Screening and mapping population development. Indian J. Genet., **69**(4): 367-370.
15. **Talukdar A. and Zhang G. Q.** 2007. Construction and characterization of 3-S lines, an alternative population for mapping studies in rice. Euphytica, **156**: 237-246.
16. www.sciencedaily.com/releases/2010/01100113131457.htm
17. **Kumar B., Talukdar A. and Verma K.** 2011. Diversity analysis in soybean genotypes using SSR markers. *In*: Book of Abstracts, National Seminar on “Contemporary approaches to crop improvement”, 22-23 April, 2011, UAS, Bangalore. pp126.