



Absence of the A4 peptide in the G4 glycinin subunit of soybean cultivar Enrei is caused by a point mutation in the *Gy4* gene

Kangfu Yu, Vaino Poysa, Margaret Haffner, Bailing Zhang and Lorna Woodrow

Agriculture and Agri-Food Canada, Greenhouse and Processing Crops Research Center, Harrow, Ontario, Canada.

Abstract

Functional properties of soy proteins for food are closely related to the composition of their storage protein subunits. Using base excision sequence scanning (BESS), we show that the absence of the A4 peptide in the G4 glycinin subunit of the soybean (*Glycine max* L.) cultivar Enrei was caused by the same point mutation in the *Gy4* gene as previously reported in the soybean cultivar Raiden. Although the genetic relationship between Raiden and Enrei is not known, the same point mutation in their *Gy4* genes may indicate that they probably share a related origin. The application of BESS to identify single nucleotide polymorphisms (SNPs) as co-dominant markers for marker-assisted selection (MAS) of a recessive null allele is also discussed.

Key words: base excision sequence scanning (BESS), marker-assisted selection (MAS), single nucleotide polymorphisms (SNPs), seed storage protein, soybean.

Received: September 24, 2004; Accepted: April 18, 2005.

Glycinin (11S) and β -conglycinin (7S) are the predominant seed storage proteins in soybeans (*Glycine max* L.). To date, six functional genes encoding glycinin subunits have been identified and assigned the gene symbols *Gy1* to *Gy5* and *Gy7*, the subunits encoded by these genes being denoted G1 to G5 and G7, respectively (Nielsen *et al.*, 1989; Beilinson *et al.*, 2002). Each subunit consists of an acidic and a basic peptide chain connected by a conserved asparagine-glycine (asn-gly) peptide bond (Staswick *et al.*, 1981). Unlike the other five glycinin subunits, the G4 subunit is unique in that it has two asn-gly bonds near the terminal end of the acidic chain that are cleaved during maturation. Thus cleavage of G4 by an asparaginyl endopeptidase results in two acidic peptides (A4 and A5) and a basic peptide (B₃) (Nielsen, 1995).

Functional properties, such as gelation and emulsification of soy proteins appear to have a close relationship with some of the soy storage protein subunits or peptides (Nakamura *et al.*, 1984; Kang *et al.*, 1991). Cultivars containing the A4 peptide form glycinin gels faster than cultivars without the A4 peptide but soymilk gels prepared from cultivars without the A4 peptide are harder and firmer than those prepared from cultivars with the A4 peptide (Murphy *et al.*, 1997). Selection of soybean

lines without the A4 glycinin peptide is often practiced in soybean breeding programs for the development of tofu-type soybean cultivars (Saio *et al.*, 1969). The absence of the A4 peptide has been associated with superior tofu quality, especially the ability to make premium tofu using magnesium chloride ('nigari' in Japanese) coagulant (Murasawa *et al.*, 1991).

Polyacrylamide gel electrophoresis (PAGE) analysis of total seed storage proteins is a standard method used by breeders to select for the presence of the A4 peptide (Fontes *et al.*, 1984) but although PAGE is a useful method it suffers two shortcomings in selection efficiency in that it requires seeds for protein extraction and analysis so breeders have to wait until the seeds are harvested to select for A4 nulls in the following generation. In addition, the absence of the A4 peptide in the Raiden cultivar is conditioned by a recessive null allele resulting from one base pair mutation in the *Gy4* gene (Scallon *et al.*, 1987) which means that breeders would discard 50% of the heterozygous F₂ generation if selections were based on PAGE analysis alone.

Although the Japanese Enrei soybean cultivar has an unknown pedigree we frequently use this cultivar in our soybean breeding programs because it lacks the A4 peptide and can thus be used as the donor cultivar for the absence of the A4 peptide. The genetic cause for the absence of the A4 peptide in the Enrei cultivar has never been reported.

The objectives of the study reported in the current paper were to locate the genetic mutation that causes the ab-

Send correspondence to Kangfu Yu. Agriculture and Agri-Food Canada, Greenhouse and Processing Crops Research Center 2585, County Road 20, N0R 1G0 Harrow, Ontario, Canada. E-mail: yuk@agr.gc.ca.

sence of the A4 peptide from the G4 subunit in the Enrei cultivar and to test the effectiveness of using the base excision sequence scanning (BESS) technique (Hawkins and Hoffman, 1997; 1999) as an alternative to PAGE to identify single nucleotide polymorphisms (SNPs) as co-dominant markers for marker-assisted selection (MAS) of a recessive null allele in soybean breeding programs.

Soybean line X811-1 (an A4 null breeding line selected from a cross involving Enrei as the A4 null parent) was crossed with soybean cultivar CH1078-1 (A4 normal) to generate F₅ recombinant inbred lines (RILs) for analyzing the A4 glycinin peptide and the *Gy4* genes. The Raiden and Enrei cultivars, both lacking the A4 glycinin peptide, were used as controls. All plants were grown in a greenhouse in pots containing the same soil and under the same controlled environmental conditions.

Analysis of the A4 glycinin peptide was conducted with the PhastSystem (Amersham Biosciences) using commercially available 10-15% polyacrylamide gradient gels (Pharmacia). Prior to use, the gels were soaked in gel buffer containing 112 mM Tris-Acetate and 6 M Urea at pH 6.4 for 1 h and air-dried for 15 min. Soybean seed proteins were obtained from 0.15 g of a seed powder and extracted for 5 min in 3 mL of extraction buffer (200 mM Tris-HCl pH 7.5 and 1 mM EDTA) in a boiling water bath. After centrifugation at approximately 7000 g for 5 min at room temperature, one μ L of the supernatant was used for polyacrylamide gradient gel electrophoresis running at 250 V, 10.0 mA and 3 W at 15 °C for 20 min and stained with Coomassie blue.

The PCR primers (Figure 1) were designed based on the *Gy4* gene sequence information in Raiden (Scallion *et al.*, 1987). The PCR products were cloned into a P-GEM^T vector as described by the manufacturer (Promega) and sequenced by dideoxy sequencing with T7 DNA polymerase (Pharmacia).

The BESS analysis was conducted according to the method of Yu *et al.* (2001) using BESS Mutation Detection

and Localization Kits (Epicentre Technologies) and leaf tissue genomic DNA was isolated as described by Yu *et al.* (1999).

The Enrei cultivar has been shown by electrophoretic analysis to lack the A4 glycinin subunit (Nishinari *et al.*, 1991) but no information is available regarding the molecular basis for the absence of the A4 peptide in this cultivar, the parental material of the X811-1 line. Molecular analysis has revealed that the absence of the A4 peptide in Raiden resulted from a single base mutation at the initiation codon of the *Gy4* gene (Scallion *et al.*, 1987). Based on the *Gy4* gene sequence information in Raiden we designed a pair of gene specific primers that flank the initiation codon (Figure 1). A 307 bp DNA fragment was amplified using genomic DNA templates from the Raiden (A4 null), Enrei (A4 null), X811-1 (A4 null) and CH1078-1 (A₄ normal) cultivars and the gene specific primers. Sequencing analysis of the PCR products verified that the Enrei cultivar and X811-1 line have the same single base mutation as the Raiden cultivar, this mutation involving one base pair substitution of deoxyguanosine (G) by deoxyadenine (A) at the initiation codon in the *Gy4* gene (Figure 1). About 20% of Japanese soybean varieties lack the *Gy4* glycinin subunit (Harada *et al.*, 1983). Although the parental relationship between the Raiden and Enrei cultivars is unknown the same point mutation in their *Gy4* genes indicates that they may share a related origin.

To detect the specific point mutation in the *Gy4* gene of soybean for the development of a gene specific SNP marker for MAS, the BESS T-Scan mutation detection and localization method was conducted by using the 3' primer end-labeled with γ ³³P. Figure 2 shows that one extra band occurred in the Raiden, Enrei cultivars and X811-1 line (all having the single base mutation in their *Gy4* gene) but not in the A4 normal CH1078-1 cultivar. Analysis of F₁ seeds from a cross between A4 normal and A4 null plants showed that both the A4 peptide and the extra SNP band were present. While the expression of the A4 peptide in those F₁

```

Enrei 1 5' TTCACCAACTCCTTCAAACCTTAATTATTAACTTCTCTAGTTCAATATAGGGAAGC
Ch1078-1 1 5' TTCACCAACTCCTTCAAACCTTAATTATTAACTTCTCTAGTTCAATATAGGGAAGC

58 CCTTCACTCTCTCTCTTTCTTCCCTTTGCTTGCTACTCTTGTGCGAGTGCATGCTTTGCT
58 CCTTCACTCTCTCTCTTTCTTCCCTTTGCTTGCTACTCTTGTGCGAGTGCATGCTTTGCT

118 ATTAGCTCCAGCAAGCTCAACGAGTGCCAACTCAACAACCTCAACGCGTTGGAACCC
118 ATTAGCTCCAGCAAGCTCAACGAGTGCCAACTCAACAACCTCAACGCGTTGGAACCC

175 GACCACCGCGTTGAGTCCGAAGGTGGTTTGATTCAAACATGGAACCTCTCAACACCCT
175 GACCACCGCGTTGAGTCCGAAGGTGGTTTGATTCAAACATGGAACCTCTCAACACCCT

232 GAGCTGAAATGCGCCGGTGTCACTGTTTCCAAACTCACCCCTCAACCGCAATGGCCTC
232 GAGCTGAAATGCGCCGGTGTCACTGTTTCCAAACTCACCCCTCAACCGCAATGGCCTC

288 CACTTGCCATCTTACTCACC 3'
288 CACTTGCCATCTTACTCACC 3'

```

Figure 1 - Comparison of Enrei and CH1078-1 cultivar genomic DNA sequences amplified by *Gy4* gene specific primers. The primer sequences are in bold type and the one base pair mutation at the initiation codon of the *Gy4* gene is underlined.

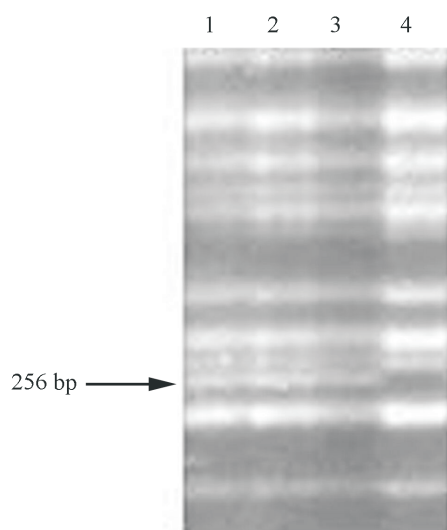


Figure 2 - The point mutation of the *Gy4* gene in the Raiden, Enrei soybean cultivars and X811-1 soybean line (lanes 1, 2 and 3) detected by BESS T-Scan compared with the CH1078-1 A4 normal cultivar (Lane 4). The DNA fragment of interest (256 bp) is indicated by an arrow.

seeds indicates that *Gy4* (A4 normal) is dominant to *gy4* (A4 null) the detection of an additional DNA band reveals that there is a single nucleotide mutation in one copy of the *Gy4* gene (Fig. 2). Figure 3 shows the co-segregation analysis of the A4 subunit with the SNP detected by BESS T-Scan in 22 of 52 F_5 RILs segregating for A4. Twenty-six of the 52 lines examined were found to have the A4

polypeptide, this number conforming to the 1:1 ratio expected in this type of segregation. Analysis of the SNPs using BESS found that 29 of them had the extra DNA band, and three also had the A4 subunit, indicating that the three lines may be heterozygous for the *Gy4* gene. Analysis of the other DNA strand by BESS G-tracker for the three lines proved that was the case.

Because PAGE analysis for the absence of the A4 subunit is targeted to the phenotype rather than the genotype, selection for A4 nulls in the F_2 generation would miss the 50% heterozygous individuals with a copy of the mutated *Gy4* gene. To avoid such loss of useful material breeders often self-pollinate the population for several generations before using PAGE to select for A4 nulls, but with the BESS T-Scan technique A4 null selection can be targeted to the mutated allele and 100% of the F_2 plants with the mutated gene can be detected at any stage in any generation and, therefore, selection can be made at the F_2 seedling stage which saves much time and space.

Co-dominant markers are more powerful than dominant markers for MAS because they can distinguish homozygous from heterozygous loci. The BESS G-Tracker can identify any point mutation involving deoxyguanosine (Hawkins and Hoffman, 1999). By end-labeling the 5' end primer, one extra band was detected in heterozygous F_1 lines that only have one copy of the normal *Gy4* gene but not in homozygous lines for the mutated gene (data not shown). In this particular case, the presence of one additional band in both the BESS T-Scan and BESS G-Tracker

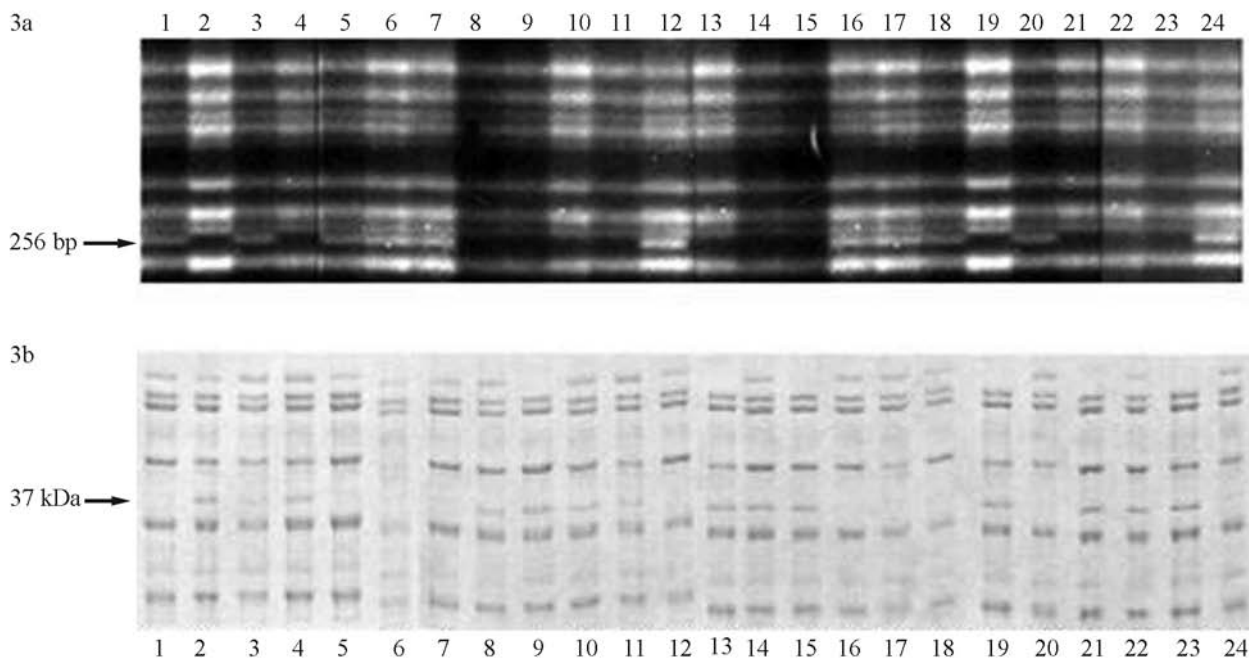


Figure 3 - Segregation analysis of the F_5 Recombinant inbred line population for the A4 protein (3b) and their corresponding single nucleotide polymorphisms (3a). Lane 1 is the A4 null parent X811-1 and Lane 2 is the A4 normal parent, CH1078-1. Lanes 3-24 are the F_5 lines derived from the parents. The DNA fragment of interest (256 bp) (3a) and the A4 subunit (37 kDa) (3b) are indicated by arrows.

analysis is an indication that the plant under investigation is heterozygous for the *Gy₄* gene. Therefore, the BESS method can be used as a co-dominant marker system for marker-assisted selection of the *Gy₄* gene and, in fact, can be used to detect any single nucleotide difference in a gene by using both strands as templates.

Acknowledgments

The authors thank the financial support of Ontario Soybean Growers, and Agriculture and Agri-Food Canada Matching Investment Initiative. The technical assistance of Dale Anderson is also acknowledged.

References

- Beilinson V, Chen Z, Shoemaker RC, Fischer RL, Goldberg RB and Nielsen NC (2002) Genomic organization of glycinin genes in soybean. *Theor Appl Genet* 104:1132-1140.
- Fontes E, Moreira M, Davies C and Nielsen NC (1984) Urea-elicited changes in relative electrophoretic mobility of certain glycinin and β -conglycinin subunits. *Plant Physiol* 76:840-842.
- Hawkins GA and Hoffman LM (1997) Base excision sequence scanning: A new method for rapid sequence scanning and mutation detection. *Nature Biotechnology* 15:803-804.
- Hawkins GA and Hoffman LM (1999) Rapid DNA mutation identification and fingerprinting using base excision sequence scanning. *Electrophoresis* 20:1171-1176.
- Kang I, Matsumura Y and Mori T (1991) Characteristics of texture and mechanical properties of heat-induced soy protein gels. *J Am Oil Chem Soc* 68:339.
- Murasawa H, Sakamoto A, Sasaki H and Harada K (1991) The effect of a glycinin subunit on tofu-making. *Proceedings of the International Conference on Soybean Processing and Utilization*, pp 53-57.
- Murphy P, Chen H, Hauck C and Wilson L (1997) Soybean protein composition and tofu quality. *Food Technology* 51:86-88.
- Nakamura T, Utsumi S, Kitamura K, Harada K and Mori T (1984) Cultivar differences in gelling characteristics of soybean glycinin. *J Agric Food Chem* 32:647.
- Nielsen N, Dickinson C, Cho T, Scallan B, Fischer R, Sims T, Drews G and Goldberg R (1989) Characterization of glycinin gene family in soybean. *The Plant Cell* 1:313-328.
- Nielsen N (1995) Soybean seed composition. In: Verma DPS and Shoemaker RC (eds) *Soybean: Genetics, Molecular Biology and Biotechnology*. CAB International, Wallingford, Oxon, pp 127-163.
- Nishinari K, Kohyama K, Zhang Y, Kitamura K, Sugimoto S, Saio K and Kawamura Y (1991) Rheological study on the effect of the A5 subunit on the gelation characteristics of soybean protein. *Agric Biol Chem* 55:351-355.
- Saio K, Kamiya M and Watanabe T (1969) Food processing characteristics of soybean 11S and 7S proteins. Part I. Effect of difference of protein components among soybean varieties on formation of tofu-gel. *Agric Biol Chem* 33:1304-1308.
- Scallan B, Dickinson C and Nielsen N (1987) Characterization of a null-allele for the *Gy₄* glycinin gene from soybean. *Mol Gen Genet* 208:107-113.
- Staswick P, Hermodson M and Nielsen N (1981) Identification of the acidic and basic subunit complexes of glycinin. *J Biol Chem* 256:8752-8755.
- Yu K, Haffner M and Poysa V (2001) Tailed Primer base Excision Sequence Scanning (TP-BESS) for detection of single nucleotide polymorphisms (SNPs). *Plant Molecular Biology Reporter* 19:49-54.
- Yu K, Park JS and Poysa V (1999) Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome* 42:27-34.

Associate Editor: Everaldo Gonçalves de Barros