

Identification of *Gy4* nulls and development of multiplex PCR-based co-dominant marker for *Gy4* and α' subunit of β -conglycinin in soybean

Mahmuda Umme Rayhan · Kyujung Van · Dong Hyun Kim · Sung Il Kim · Moon Young Kim · Yeong-Ho Lee · Suk-Ha Lee

Received: 22 November 2010 / Accepted: 05 April 2011 / Published online: 31 August 2011
© The Genetics Society of Korea and Springer 2011

Abstract

Alpha prime (α') subunit of β -conglycinin and *Gy4* subunit of glycinin are two important subunits of soybean storage protein which have negative effects on food processing, total amino acid content, and hypersensitivity reactions. It has been possible to reduce or remove some of these problems from soybean by screening or developing mutant lines. The objective of this study was to establish a simple, cheap DNA marker for *Gy4* and α' subunit for use in non-seed destructive, marker-assisted selection (MAS) that can identify these two mutants at the same time in a unique PCR reaction. To achieve this objective, we identified eight of *Gy4* mutants from diverse soybean accessions from the USDA Soybean Germplasm Collection and described a multiplex PCR based co-dominant DNA marker for *Gy4* subunit of glycinin. Then we crossed one of these *Gy4* mutants with Keburi (α' mutant) for development of double mutant variety and established a multiplex PCR based, co-dominant DNA marker for screening *Gy4* and α' mutants. Thus, using this newly developed marker to identify *Gy4* and α' mutants in breeding programs we could save our time, labor, and resources.

Keywords α' subunit; Co-dominant marker; Double mutant; *Gy4*; Multiplexing PCR; Soybean

Introduction

Soy foods have received a lot of attention as potential beneficial factors in the diet because their isoflavones are rich in calcium, iron, zinc, phosphate, magnesium, and B vitamin (Shu et al., 2001; Wu et al., 2002; Thanos et al., 2006; Wu et al., 2008; Korde et al., 2009; Lee et al., 2009). Also, soybean proteins have been reported for various health benefits, including lowering cholesterol (Fukui et al., 2002; Huang et al., 2005) and blood pressure (Lo and Li-Chan, 2005; Ringseis et al., 2005), reduction of plasma triglyceride levels (Iritani et al., 1996), appetite suppressing effect (Nishi et al., 2003), anti-obesity effect (Aoyama et al., 2000), anti-atherosclerosis effect (Huff et al., 1982), and anti-carcinogenesis effect (Tsuruki et al., 2005). However, the recognition of one of the “big 8” food allergens (Cordle, 2004) and the lower level of sulfur-containing amino acids in soybean protein (Fukushima, 1991) are setbacks of the benefits of soybean. Also, food processing like tofu gel formation face difficulties due to inappropriate functional properties of soybean seed storage proteins (Poysa and Woodrow, 2002; Salleh et al., 2004; Poysa et al., 2006).

The quantity and quality of the protein in seeds is the major biochemical component influencing soybean quality for soy food production (Poysa and Woodrow, 2002). The predominant soybean seed storage proteins are glycinin (11s) and β -conglycinin (7s) (Poysa and Woodrow, 2002; Poysa et al., 2006), accounting for 40 and 30% of the total seed proteins, respectively (Utsumi, 1992; Utsumi et al., 1997; Krishnan, 2000). β -conglycinin is a trimetric protein composed of α (~67 kD), α' (~71 kD) and β (~50 kD) subunits (Thanh and Shibasaki, 1978) and all three subunits of β -conglycinin have potentials for food allergens (Krishnan et al., 2009). Since α' subunit has a greater negative effect on tofu gel formation than whole β -conglycinin (Salleh et al., 2004), molecular markers were developed for identifying α' subunit-lacking soybean lines (Ishikawa et al., 2006; Kim et al., 2011).

M. U. Rayhan · K. Van · D. H. Kim · S. I. Kim · M. Y. Kim · Y.-H. Lee · S.-H. Lee (✉)

Department of Plant Science and Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 151-921, Korea
e-mail: sukhalee@snu.ac.kr

S.-H. Lee

Plant Genomics and Breeding institute, Seoul National University, Seoul 151-921, Korea

Six functional genes encoding glycinin subunits have been identified (*Gy1*, *Gy2*, *Gy3*, *Gy4*, *Gy5*, and *Gy7*) (Beilinson et al., 2002) and showed 45–49% of homology among the subunits (Prak et al., 2005). Each subunit consists of acidic (A) and basic (B) polypeptide components (Staswick et al., 1984; Nielsen et al., 1989). The acid and basic peptide chains are connected by a conserved Asn-Gly peptide bond (Staswick et al., 1984), except the A4 peptide in the *Gy4* subunit (Mooney and Thelen, 2004). Moreover, cleavage of *Gy4* by asparaginyl endopeptidase results in two acidic peptides (A4 and A5) and one basic peptide (B3) (Nielsen, 1995). Gels of soymilk from cultivars without the A4 peptide are harder and firmer than those prepared from cultivars with the A4 peptide (Murphy et al., 1997). Interestingly, soybean lines without the A4 glycinin peptide are often practiced in soybean breeding programs for the development of tofu (Saio et al., 1969) and hypoallergenic soybean cultivars (Rozenfeld et al., 2002).

Unwanted properties of soybean could be overcome by development of mutant lines (Samoto et al., 1997; Ogawa et al., 2000). The most popular technique to identify *Gy4*-lacking mutants is sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of total seed storage proteins (Fontes et al., 1984). Since SDS-PAGE requires seeds for protein extraction and analysis, this technique is not good for marker-assisted selection (MAS). In addition, Scallan et al., (1987) suggested that the absence of the A4 peptide is conditioned by a recessive null allele. So, 50% of the heterozygous F_2 generation could be discarded if selections were based on SDS-PAGE analysis alone. Thus, SDS-PAGE analysis is neither an accurate method nor good for MAS identifying mutants.

As an alternative to SDS-PAGE, single nucleotide polymorphism (SNP) was used for identifying A4-lacking soybean mutants (Yu et al., 2005). Recently, Ishikawa et al., (2006) described the PCR-based dominant marker for β -conglycinin α' subunit-lacking lines. However, it could not differentiate homozygotes or heterozygotes and recessive homozygotes could be detected as false negatives due to the inactive PCR process (Winter and Kahl, 1995).

In this study, several diverse *Gy4* nulls from the USDA Soybean Germplasm Collection were identified and multiplex PCR-based co-dominant marker for *Gy4* nulls was established. One of *Gy4* null genotype was crossed with α' subunit null (Keburi) for developing double mutant variety and a multiplex PCR-based co-dominant marker was established, for screening *Gy4* and α' null soybean cultivars. Here we describe a modified allele-specific PCR procedure that is sufficiently robust to easily discriminate between *Gy4* nulls and wild-type at very low cost.

Materials and Methods

Plant materials and genomic DNA extraction

Soybean lines used in this experiment are listed in Supplemen-

tary Table 1 and all lines were provided by the USDA Soybean Germplasm Collection. Ninety-five soybean accessions with high and low protein contents were included and their origins were from Japan, China, and Korea. As wild-types for the glycinin *Gy4* subunit and the β -conglycinin α' subunit, Williams 82, Daewonkong, Taekwangkong, and Danbaekkong were used. Keburi (PI 506876) genotype was used as α' mutant and PI 567395 (Kim et al., 2008) was used as a positive control of glycinin *Gy4* subunit. F_1 progenies derived from a cross between PI 238928 (*Gy4* mutant) and the Daewonkong (wild genotype) were used for examining the heterozygosity of our developed co-dominant marker for *Gy4*. Another F_1 progenies derived from a cross between PI 238928 (*Gy4* mutant) and Keburi (α' mutant) were also used for development of a multiplex PCR-based marker for *Gy4* and β -conglycinin α' subunit.

Young expanding leaves were collected in the greenhouse and ground to a fine powder in liquid nitrogen using a grinder. Genomic DNA was extracted using the procedure described by Shure et al., (1983), using urea with slight modifications. DNA was resuspended in 100 μ l of TE Buffer (Tris-EDTA buffer, pH 8.0) and incubated at 37 °C for 3 hr for RNase treatments. DNA concentration was determined using a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and DNA was diluted in TE Buffer to a concentration of 50 ng/ μ l and stored at 20 °C until use.

Screening of soybean *Gy4* mutants

A segment of the *Gy4* gene was amplified from genomic DNA of all 95 soybean accessions as a template. Forward and reverse primers were designed to amplify 307 bp of the N-terminal region of the *Gy4* gene, including 47 bp of untranslated 5' upstream sequences of the gene (Kim et al., 2008) (Table 1). PCR reactions were carried out in a final volume of 20 μ l consisted of 50 ng of genomic DNA, 10 $\times$ PCR buffer (75 mM Tris-HCl (pH 8.5), 20 mM (NH₄)₂SO₄, 2.5 mM MgCl₂), 0.2 nM dNTPs, 1.0 μ M each of the forward and reverse primers and 1.6 U *Taq* DNA polymerase (Vivagen, Sungnam, Korea). The PCR reaction was conducted in a PTC-225 Peltier Thermal Cycler (MJ Research Inc., Watertown, MA, USA) with an initial denaturation step at 94 °C for 5 min, followed by 30 cycles at 94 °C for 30 sec, 60 °C for 30 sec, and 72 °C for 30 sec, and the final extension step at 72 °C for 10 min. PCR products were resolved by electrophoresis on 1% ethidium bromide-stained gel. Highly qualified DNA sequences generated by ABI3730 sequencer (Applied Biosystems, Forest City, CA, USA) were compared with sequences of Williams 82 and analyzed using Seqscape (Applied Biosystems, Forest City, CA, USA).

Extraction of seed protein and SDS-PAGE

Soybean seeds of PI 238928 were ground to a fine powder using a mortar and pestle. A 10 mg aliquot was extracted with 1 ml of extraction solution containing 125 mM Tris-HCl (pH 6.8), 4% sodium dodecyl sulfate (w/v), 20% glycerol (v/v),

50 μl of 2-mercaptoethanol, and 0.03 mM bromophenol blue. After the sludge was heated in an incubator at 100 °C for 3 min and centrifuged at 12,000 rpm for 20 min, the supernatant was transferred into a new tube. Aliquots (10 μl) of the supernatant were utilized for 15% SDS-polyacrylamide gel electrophoresis (PAGE) analysis (Laemmli, 1970).

Uniplex and multiplex PCR for *Gy4* specific allele

The SNAPER program (<http://ausubellab.mgh.harvard.edu/>) was used for designing G allele-specific primers. The primers specific for A allele was also designed and three nucleotides from the 3' end were added (Table 1). Two different PCR reactions were performed for 'G' and 'A' specific alleles under different reaction conditions. PCR was performed in a 20 μl reaction mixture containing 10 $\times$ PCR buffer (the same as described above), 0.5, 1, and 2 mM dNTPs, 0.1, 0.3, and 1.0 μM each of the forward and reverse primers, 1.6 U *Taq* DNA polymerase and 10, 30, and 50 ng of genomic DNA of PI 238928 and Williams 82. The PCR reaction started with an initial denaturation step at 94 °C for 5 min, followed by 28 or 38 cycles using the following conditions, denaturation at 94 °C for 30 sec, and annealing and extension at 60, 62, 64 °C for 1 min, and then final extension at 72 °C for 20, 30, and 40 min. For multiplex PCR, reactions were performed according to optimal reaction mixture determined in uniplex assays and four primers of both allele were used. The products amplified by PCR were analyzed by 1.5% agarose gel electrophoresis.

Multiplex PCR for *Gy4* and α' subunit

For multiplexing PCR for *Gy4* and α' subunit of soybean storage protein, we used seven primers. For *Gy4*, both 'G' and 'A' allele-specific primer sets were used. Also, three primers designed based on the genomic sequences of Williams 82 and Keburi were used for β -conglycinin α' subunit, corresponding to each allele of the *Cgy-1* locus (Kim et al., 2011). To optimize multiplex PCR conditions, various ratios of the primer concentrations (data not shown) and genomic DNA concentrations (30, 50, and 70 ng) were tested. Performance of PCR was the same as uniplex assays except dNTPs 1.5 mM and 1.7 U of *Taq* DNA polymerase.

SSR marker analysis

SSR markers, Satt144, Satt146, Satt206, Satt510, Satt114, Satt160, Satt252, Satt269, Satt374, and Satt425, were used to confirm F_1 progeny generated from two populations. For SSR marker analysis, the PCR mixture contained 10 ng of template DNA, 0.5 μM of each forward and reverse primer, 0.25 nM dNTPs, 1 $\times$ reaction buffer (75 mM Tris-HCl (pH 8.5), 20 mM $(\text{NH}_4)_2\text{SO}_4$, 2.5 mM MgCl_2), and 0.8 U *Taq* DNA polymerase (Vivagen) in 10 μl of total volume. The SSR amplification was conducted with the following cycling conditions: an initial denaturation at 94°C for 5 min followed by 30 cycles of 94°C for 25 sec, 58°C for 25 sec and 72°C for 25 sec, and final extension at 72°C for 10 min. The PCR products were resolved with 3.5% NuSieve® GTG® agarose (Lonza Group Ltd., Rockland, ME, USA) involving a 1.0% standard agarose gel stained with ethidium bromide.

Results

Screening of *Gy4* soybean mutants

We designed primers, which could amplify a 307 bp of the N-terminal region of the *Gy4* gene including 47-bp untranslated sequences of 5' upstream of *Gy4* gene because the previous *Gy4* mutants showed polymorphism in this region (Scallan et al., 1987; Kim et al., 2008). After the part of the *Gy4* gene was directly sequenced from all 95 soybean accessions, sequences were compared with the same region of Williams 82, whose sequence is publically available at Phytozome (<http://www.phytozome.net/soybean.php>). Like other previously identified *Gy4* mutants (Kim et al., 2008), the same single point mutation, which changed to 'A' from 'G' at third nucleotide of the translation initiation codon of the *Gy4* gene on chromosome (chr.) 10, was also observed among our screened soybean accessions (Fig. 1). Further confirmation for the lack of the A4 glycinin peptide in one of these newly identified *Gy4* mutants was obtained by SDS-PAGE. The A4 peptide was clearly seen in protein extracts from Williams 82 but was missing from PI 238928, one of the eight *Gy4* mutant lines identified in this study (Fig. 1B).

Table 1. List of multiplex PCR Primer sequences used in this study for *Gy4*.

Allele	Primer sequence (5' $\rightarrow$ 3')	PCR product size (bp)
The N-terminal region of <i>Gy4</i>	F: TTCACCAACTCCTTCAAACCTTAATT R: GGTGAGTAAGATGGCAAGTGGAG	307
G-specific	F: AAGAGGATAGACTTTGAGACGCGGAGAATCT R ^a : GAGAGAGTGAAGGGCTTCCGC	324
A-specific	F ^b : ACTTAATTATTAACACTTCCTTAGTTCAATACA R: TTCTGGTGAAGTGTCTGTAGCTGCTGCT	771

^aSpecific to G allele

^bSpecific to A allele

After 95 soybean accessions were screened by direct sequencing of a part of *Gy4* gene (Supplementary Table 1), eight soybean lines, five mutants from Japan and three mutants from China, were identified (Table 2). Among newly identified mutants, four mutants were in IV maturity group, and three and one mutants belonged to maturity groups V and VII, respectively. Most *Gy4* mutants had yellow colored seed coat and others showed brown or buff. Three mutants, PI 567614A, PI 567580A, and PI 594217C, were considered as high protein content soybean with average of above 51 g/kg protein, whereas five mutants had only 35 g/kg in protein. Oil content of our identified mutants ranged from 15.3 to 25.6 g/kg (Table 2).

Development of a co-dominant marker for *Gy4* mutants

Based on a SNP between wild-type (Williams 82) and *Gy4* null (PI 238928) at the initiation codon, a multiplex

PCR-based marker for MAS was developed. After several primers suggested by SNAPER program were tested (data not shown), we optimized allele specific single nucleotide amplified polymorphism (SNAP)-PCR by using several extension time, DNA concentration, PCR amplification cycle, primer concentrations, and dNTPs concentrations (see Materials and Methods). In terms of the PCR amplification cycle, we tested 28 or 38 cycles. Specific amplicons were not produced with 38 cycles, but the specific PCR bands for each allele were generated if we used 28 cycles (data not shown). One pair of primer set for each allele ('G'- and 'A'-specific) was selected for SNAP marker development (Table 1). And, specific amplicons were produced with 30 ng of genomic DNA, 0.3 μ M of primer, 1 mM of dNTPs, and 30 min as extension time. With these optimization and the allele-specific primers, four wild-type soybean and nine *Gy4* mutants including eight newly

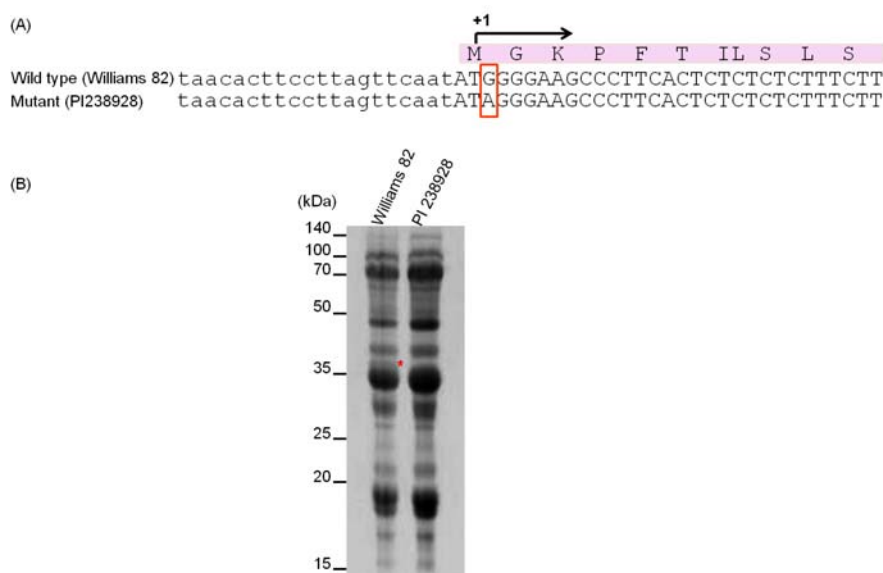


Figure 1. Comparison of *Gy4* between wild-type and mutant. (A) Nucleotide and deduced amino acid sequence comparison of a portion of the *Gy4* gene from wild-type and mutant line. The protein translation start position is marked as +1. Initiation codon “ATG” in wild-type is changed to “ATA” in mutant (indicated by red box). (B) Comparison of the storage proteins of wild-type and mutant by SDS-PAGE. Seed proteins were extracted in SDS sample buffer and resolved via 15% SDS-PAGE. * indicates the A₄ glycinin peptides and the sizes of protein markers in kilodaltons are shown at the left of the figure.

Table 2. Identified *Gy4* mutants in this study and their relevant characteristics*.

PI number	Origin	Maturity group	Seed coat color	100-seed weight (g)*	Protein	Oil (g/kg)*
PI 235347	Japan	IV	Yellow	12.4	34.7	25.6
PI 567614A	China	IV	Buff	10.8	52.2	15.3
PI 567580A	China	IV	Yellow	16.7	52.9	21.2
PI 594217C	Japan	VII	Yellow	13.8	50.5	15.8
PI 417486	Japan	V	Yellow	15.2	35.8	21.0
PI 238928	Japan	IV	Yellow	15.6	35.8	24.4
PI 181546	Japan	V	Yellow	14.3	35.9	23.4
PI 407747	China	V	Brown	23.5	36.4	21.6

*Data were compiled from USDA Germplasm Resources Information Network (<http://www.ars-grin.gov/npgs/ngo.html>).

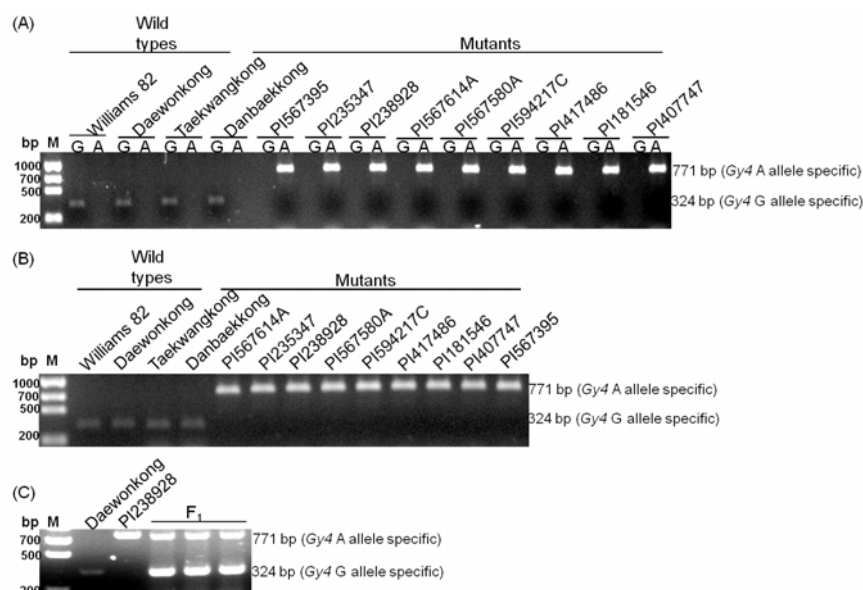


Figure 2. Optimization and evaluation of co-dominant marker for *Gy4*. (A) Optimization of specific band for specific allele of *Gy4*. G, G-specific; A, A-specific; M marker. (B) Genotyping using multiplex PCR-based co-dominant marker with other germplasm for *Gy4* gene under optimum conditions. PI 567395 was used as a positive control (Kim et al., 2008). (C) Genotyping of F₁ individuals heterozygous to *Gy4*. The F₁ individuals were derived from the cross between Daewonkong and PI238928.

identified mutants were tested (Fig. 2A). Wild-type soybean showed PCR amplicons only with ‘G’ allele specific primers and ‘A’ allele specific primers only amplified the genomic DNA from the soybean accessions, which had the mutated sequence of *Gy4*.

Since ‘G’ and ‘A’ allele specific primers amplified different lengths of DNA fragments (324 vs. 771 bp), we were able to perform multiplex PCR with a co-dominant marker for *Gy4*. We confirmed with four wild-type soybean genotypes including three cultivars and eight newly identified mutants (Fig. 2B). Three cultivars, Daewonkong, Taekwangkong, and Danbaekkong, represented the 324-bp amplicons in size the same as wild-type, Williams 82. The *Gy4* mutant, PI 567395, identified by Kim et al. (2008) produced the same size of 771 bp as shown in our eight identified *Gy4* mutants in this study. All soybean lines successfully produced the right size of PCR amplicons and the size differences were easily recognized (Fig. 2B).

The heterozygosity of co-dominant marker for *Gy4* developed in this study was also assessed using F₁ progeny from the cross of *Gy4* mutant (PI 238928) and wild-type (Daewonkong). After a total of 10 SSR markers were tested for screening polymorphisms between parents (data not shown), Satt269 exhibiting parental polymorphism was chosen for testing the heterozygosity of the F₁ plants. Using our multiplex, co-dominant marker for the *Gy4* subunit, the F₁ plants had two distinct DNA bands, indicating inheritance from each of the parental genotypes (Fig. 2C). These results revealed that the co-dominant marker for the *Gy4* gene allowed us to select heterozygous genotype due to the size difference.

Multiplex PCR-based marker for both *Gy4* and β -conglycinin α' subunit

For multiplex PCR, seven primers were used. Two pairs of primers (‘G’ and ‘A’ allele specific) were used to amplify the genomic region of *Gy4* mutant (Table 1) and three other primers were also included for amplifying the genomic region between 5’ upstream and 3’ downstream of the deletion region of *Cgy-1* in Keburi and full *Cgy-1* gene in wild-type (Kim et al., 2011).

For optimization of PCR conditions, three different conditions of annealing temperatures, primer working concentrations, and DNA amounts were tested to increase the specificity and fidelity of multiplex PCR (data not shown). These tests revealed that the annealing temperature had more influence on PCR specificity than the ratios of primer and DNA concentrations. No PCR amplicons were produced for either wild-type (Williams 82) or β -conglycinin α' subunit mutant (Keburi) at 60 °C annealing temperature. Also, β -conglycinin α' subunit mutant (Keburi) did not generate PCR products, if we used 64 °C as an annealing temperature. Therefore, 62 °C was the optimum annealing temperature and 50 ng of DNA gave the clear and specific bands with reduced concentrations of ‘G’ and ‘A’ allele specific primers and increased concentrations of *Cgy-1*_1F and *Cgy-1*_R primers. The optimum concentrations of *Cgy-1*_1F: *Cgy-1*_2F: *Cgy-1*_R: G specific F: G specific R: A specific F: A specific R were determined to be 2:1:2:0.3:0.3:2.5:2.5 respectively (Fig. 3).

Using the determined PCR conditions, four specific DNA fragments of 913, 649, 324, and 771 bp corresponding to α' wild-type, α' mutant, *Gy4* wild-type, and *Gy4* mutant, respectively, were successfully obtained. Also, four soybean

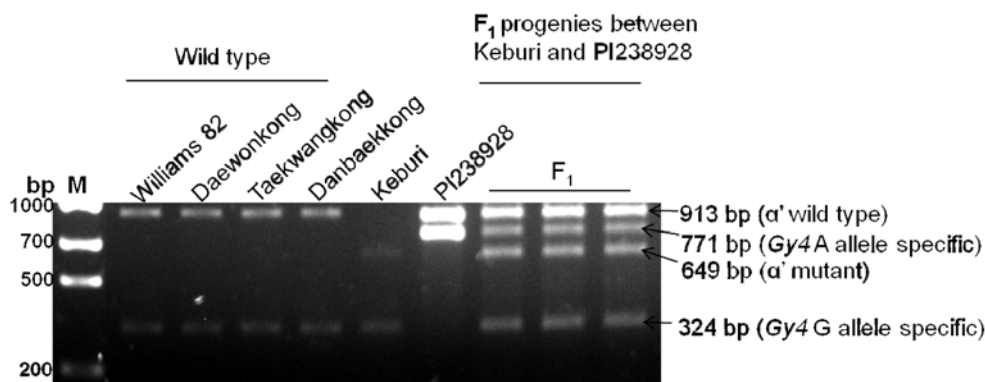


Figure 3. Genotyping using multiplex PCR-based co-dominant marker for *Gy4* and α' subunit with other germplasm and the F_1 progeny under optimum conditions. The F_1 individuals were derived from the cross between Keburi and PI 238928.

wild-type and mutants for *Gy4* and β -conglycinin α' subunits were used as controls for multiplexing two different co-dominant markers. To confirm co-dominant markers for both *Gy4* and β -conglycinin α' subunit, we generated double mutants crossed with PI 238928 (*Gy4* mutant) and Keburi (β -conglycinin α' subunit). After parental polymorphism was confirmed with the crossed samples (data not shown), all primers with the optimized PCR conditions were applied. Our established multiplex co-dominant markers for *Gy4* and β -conglycinin α' subunit enabled allelic discrimination according to the sizes of the PCR products in the same PCR reaction (Fig. 3).

Discussion

Since 7S (β -conglycinin) and 11S (glycinin) globulins are the principal storage protein in soybean and accounted for about 70% of the total seed protein (Poysa et al., 2006), screening of naturally produced mutants of β -conglycinin and *Gy4* subunit is the best way to improve soybean protein for human and animal health.

SDS-PAGE is the standard analysis for identifying *Gy4* mutants in breeding programs. Due to the need for seeds for protein extraction and the time required for the seeds to mature, this SDS-PAGE analysis is not practical. By the development of PAGE technique, many soybean glycinin subunits were reproducibly separated. But, the rapid screening method for identifying mutants is needed instead of SDS-PAGE. Since a point mutation in the initiation codon is responsible for no accumulation of *Gy4* protein (Yu et al., 2005; Kim et al., 2008), SNP analysis of *Gy4* initiation codon became the common method for identifying *Gy4* mutants.

We used 95 soybean accessions from USDA Soybean Germplasm Collection to survey *Gy4* mutants and our newly identified *Gy4* soybean lines also had the same point mutation (G/A) as the previous studies (Fig. 1A) and as a consequence, transcripts of this gene do not accumulate, A4 peptides in one

of our identified mutants PI 238928-like first observed in Enrei (Yu et al., 2005). Seed sizes of all identified mutants were the largest, which is preferable for tofu production (Liu, 1997). Up until now, few *Gy4* mutants have been utilized in breeding programs for developing cultivars suited for tofu production. The *Gy4* mutants we identified from the USDA Soybean Germplasm Collection will provide an opportunity to expand the genetic diversity of tofu as well as hypoallergenic cultivars.

To detect the specific point mutation of the *Gy4* gene in soybean, the gene-specific SNP marker for MAS was developed. After several trials of optimizations with allele specific primers suggested by SNAPER, we were able to determine four allele-specific primers (Fig. 2A). Different concentrations of dNTPs, such as 0.5, 1, and 2 mM, were tested and finally selected the lower concentration (1 mM) than usual because lower dNTPs concentration minimizes mispriming at nontarget sites (Innis et al., 1988). In addition, it is very important to use accurate extension time and concentration of genomic DNA and primer. Fig. 2B showed that a 324-bp amplicon was detected in wild-type soybean genotypes, whereas mutants produced a 771 bp amplicon. Analysis of F_1 plants from a cross between *Gy4* normal and *Gy4* null plants showed both bands specific to both alleles (Fig. 2C). It is suggest that heterozygosity of the F_1 plants to the *Gy4* gene was ascertained using the co-dominant marker.

Multiplex PCR is a variant of PCR, in which more than one locus was simultaneously amplified in the same reaction with their matching pairs of primers. This technique is conceptually simple, highly specific, sensitive, and amenable to full automation (Winn-Deen, 1996; Klapper et al., 1998). This method has been successfully applied in many areas of nucleic acid diagnostics, including gene deletion analysis (Henegariu et al., 1994), mutation and polymorphism analysis (Shuber et al., 1993; Rithidech et al., 1997), quantitative analysis (Zimmermann et al., 1996; Sherlock et al., 1998;), RNA detection (Jin et al., 1996; Zou et al., 1998), and reverse transcription PCR (Crisan, 1994). The screening of soybean lines

by MAS could be efficient using multiplex PCR. In this study, the developed co-dominant marker for β -conglycinin α' subunit (Kim et al., 2011) was used for multiplex PCR along with our primers for *Gy4* SNAP marker.

For developing multiplex PCR-based markers for both *Gy4* and β -conglycinin α' , we used F_1 progenies derived from a cross between PI 238928 (*Gy4* mutant) and Keburi (α' mutant). Comparing this double mutant with wild-type (Williams 82), seven primers were used in one PCR reaction. For optimizing the multiplex PCR conditions, we started with testing three different annealing temperatures, 60, 62, and 64 °C (data not shown). Wild-type generated false positive bands for *Gy4* and α' at 60 °C and no bands were detected for α' mutant (Keburi) in both wild and double mutants at 64 °C. But, the specific bands for each allele were amplified at 62 °C. Genomic DNA concentrations and the ratios of primer concentrations were also tested for optimizing multiplex PCR. After clear and different sizes of amplicons were generated by the proper optimization of several factors, wild-type and mutant for *Gy4* and β -conglycinin α' subunits were successfully distinguished by multiplex PCR (Fig. 3). Four PCR bands were produced in our developed double mutant; 913 bp for α' full gene, 649 bp for deleted α' gene, 324 bp for *Gy4* wild, and 771 bp for null allele of *Gy4* (Fig. 3), but wild-type showed only two bands for α' full gene and *Gy4* wild-type.

MAS is indirect selection process where a trait of interest is selected based on not only the trait itself but also a marker linked to it (Rosyara, 2006). MAS can be useful for traits that are difficult to measure, exhibit low heritability, and express late in development. Instead of difficulty in phenotyping, our developed co-dominant marker for *Gy4* and the multiplex system for *Gy4* and β -conglycinin α' subunit could easily be analyzed through simple PCR and agarose gel electrophoresis. Also, using normal PCR buffer and *Taq* DNA polymerase is one of the advantages for fast screening of MAS in soybean breeding programs. Therefore, this marker and multiplex PCR protocol are highly suitable for mutant detection and MAS for developing soybean lines and more reproducible than other genotyping methods.

Acknowledgments This research was supported by a grant from the Next-Generation BioGreen 21 Program (No. PJ0080602011), Rural Development Administration, Republic of Korea.

References

Aoyama T, Fukui K, Takamatsu K, Hashimoto Y and Yamamoto T (2000) Soy protein isolate and its hydrolysate reduce body fat of dietary obese rats and genetically obese mice (Yellow KK). *Nutrition* 16: 349–354.

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.

Cordle CT (2004) Soy protein allergy: Incidence and relative severity. *Am. Soc. Nutr. Sci.* 134: 1213S–1219S.

Crisan D (1994) Molecular diagnostic testing for determination of myeloid lineage in acute leukemias. *Ann. Clin. Lab. Sci.* 24: 355–363.

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.

Fukui K, Tachibana N, Wanezaki S, Tsuzaki S, Takamatsu K, Yamamoto T, Hashimoto Y and Shimoda T (2002) Isoflavone-free soy protein prepared by column chromatography reduces plasma cholesterol in rats. *J. Agric. Food Chem.* 50: 5717–5721.

Fukushima D (1991) Recent progress of soybean protein foods: chemistry, technology and nutrition. *Food Rev. Intl.* 7: 323–351.

Henegariu O, Hirschmann P, Kilian K, Kirsch S, Lengauer C, Maiwald R, Mielke K and Vogt P (1994) Rapid screening of the Y chromosome in idiopathic sterile men, diagnostic for deletions in AZF, a genetic Y factor expressed during spermatogenesis. *Andrologia* 26: 97–106.

Huang W, Wood C, L'Abbé MR, Gilani GS, Cockell KA and Xiao CW (2005) Soy protein isolate increases hepatic thyroid hormone receptor content and inhibits its binding to target genes in rats. *J. Nutr.* 135: 1631–1635.

Huff MW, Roberts DC and Carroll KK (1982) Long-term effects of semipurified diets containing casein or soy protein isolate on atherosclerosis and plasma lipoproteins in rabbits. *Atherosclerosis* 41: 327–336.

Innis MA, Myambo KB, Geofand DH and Brow MAD (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. *Proc. Natl. Acad. Sci. USA* 85: 9436–9440.

Iritani N, Hosomi H, Fukuda H, Tada K and Ikeda H (1996) Soybean protein suppresses hepatic lipogenic enzyme gene expression in wistar fatty rats. *J. Nutr.* 126: 380–388.

Ishikawa G, Takada Y and Nakamura T (2006) A PCR-based method to test for the presence or absence of beta-conglycinin alpha' and alpha-subunit in soybean. *Mol. Breed.* 17: 365–374.

Jin L, Richards A and Brown DWG (1996) Development of a dual target-PCR for detection and characterization of measles virus in clinical specimens. *Mol. Cell Probes* 10: 191–200.

Kim SI, Kim MY, Van K, Lee YH, Kim HS, Park BS, Seo HS and Lee SH (2011) The development of a co-dominant marker for the β -conglycinin α' subunit in soybeans. *Euphytica* 177: 355–363.

Kim WS, Ho HJ, Nelson RL and Krishnan HB (2008) Identification of several *gy4* nulls from the USDA soybean germplasm collection provides new genetic resources for the development of high-quality tofu cultivars. *J. Agric. Food Chem.* 56: 11320–11326.

Klapper PE, Jungkind DL, Fenner T, Antinozzi R, Schirm J and Blanckmeister C (1998) Multicenter international work flow study of an automated polymerase chain reaction instrument. *Clin. Chem.* 44: 1737–1739.

Korde LA, Wu AH, Fears T, Nomura AM, West DW, Kolonel LN, Pike MC, Hoover RN and Ziegler RG (2009) Childhood soy intake and breast cancer risk in Asian American women. *Cancer Epidemiol. Biomarkers Prev.* 18: 1050–1059.

Krishnan HB (2000) Biochemistry and molecular biology of soybean seed storage proteins. *J. New Seeds* 2: 1–25.

Krishnan HB, Kim WS, Jang S and Kerly MS (2009) All three subunits of soybean beta-conglycinin are potential food allergens.

- J. Agric. Food Chem. 57: 938-943.
- Laemmli JK (1970) SDS-polyacrylamide gel electrophoresis of protein. *Nature* 2: 680-685.
- Lee SA, Shu XO, Li H, Yang G, Cai H, Wen W, Ji BT, Gao J, Gao YT and Zheng W (2009) Adolescent and adult soy food intake and breast cancer risk: results from the Shanghai Women's Health Study. *Am. J. Clin. Nutr.* 89: 1920-1926.
- Liu K (1997) Soybeans: Chemistry, Technology, and Utilization, Chapman and Hall, New York.
- Lo EMY and Li-Chan EY (2005) Angiotensin I converting enzyme inhibitory peptides from in vitro pepsin-pencreatin digestion of soy protein. *J. Agric. Food Chem.* 53: 3369-3376.
- Mooney BP and Thelen JJ (2004) High-throughput peptide mass fingerprinting of soybean seed proteins: automated workflow and utility of UniGene expressed sequence tag databases for protein identification. *Phytochemistry* 65: 1733-1744.
- Murphy P, Chen H, Hauck C and Wilson L (1997) Soybean protein composition and tofu quality. *Food Technol.* 51: 86-110.
- Nielsen NC (1995) Soybean seed composition. In DPS Verma, RC Shoemaker, eds, *Soybean: Genetics, Molecular Biology and Biotechnology*, CAB International, Oxon, Wallingford, pp 127-163.
- Nielsen NC, Dickinson CD, Cho TJ, Thanh VH, Scallan BJ, Fischer RL, Sims TL, Drews GN and Goldberg RB (1989) Characterization of the glycinin gene family in soybean. *Plant Cell* 1: 313-328.
- Nishi T, Hara H, Asano K and Tomita F (2003) The soybean β -conglycinin β 51-63 fragment suppresses appetite by stimulating cholecystokinin release in rats. *J. Nutr.* 133: 2537-2542.
- Ogawa T, Samoto M and Takahashi K (2000) Soybean allergens and hypoallergenic soybean products. *J. Nutr. Sci. Vitaminol.* 46: 271-279.
- Poysa V and Woodrow L (2002) Stability of soybean seed composition and its effect on soymilk and tofu yield and quality. *Food Res. Intl.* 35: 337-345.
- Poysa V, Woodrow L and Yu K (2006) Effect of soy protein subunit composition on tofu quality. *Food Res. Intl.* 39: 309-317.
- Prak K, Nakatani K, Katsube T, Adachi M, Maruyama N and Utsumi S (2005) Structure-function relationships of soybean proglycinins at subunit Levels. *J. Agric. Food Chem.* 53: 3650-3657.
- Ringseis R, Matthes B, Lehmann V, Becker K, Schöps R, Ulbrich-Hoffmann R and Eder K (2005) Peptides and hydrolysates from casein and soy protein modulate the release of vasoactive substances from human aortic endothelial cells. *Biochem. Biophys. Acta.* 1721: 89-97.
- Rithidech K N, Dunn JJ and Gordon CR (1997) Combining multiplex and touchdown PCR to screen murine microsatellite polymorphisms. *BioTech.* 23: 36-45.
- Rosyara UR (2006) Requirement of robust molecular marker technology for plant breeding applications. *J. Plant Breed. Gr.* 1: 67-72.
- Rozenfeld P, Docena GH, Anon MC and Fossati CA (2002) Detection and identification of a soy protein component that cross reacts with caseins from cow's milk. *Clin. Exp. Immunol.* 130: 49-58.
- 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.
- Salleh MRB, Maruyama N, Takahashi K, Yagasaki K, Higasa T, Matsumura Y and Utsumi S (2004) Gelling properties of soybean beta-conglycinin having different subunit compositions. *Biosci. Biotechnol. Biochem.* 68: 1091-1096.
- Samoto M, Fukuda Y, Takahashi K, Tabuchi K, Hiemori M, Tsuji H, Ogawa T and Kawamura Y (1997) Substantially complete removal of three major allergenic soybean proteins (Gly m Bd 30 K, Gly m Bd 28K and the alpha-subunit of conglycinin) from soy protein by using a mutant soybean, Tohoku 124. *Biosci. Biotechnol. Biochem.* 61: 2148-2150.
- Scallan BJ, Dickinson CD and Nielsen NC (1987) Characterization of a null-allele for the *Gy4* glycinin gene from soybean. *Mol. Gen. Genet.* 208: 107-113.
- Sherlock J, Cirigliano V, Petrou M, Tutschek B and Adinolfi M (1998) Assessment of diagnostic quantitative fluorescent multiplex polymerase chain reaction assays performed on single cells. *Ann. Hum. Genet.* 62: 9-23.
- Shu XO, Jin F, Dai Q, Wen W, Potter JD, Kushi LH, Ruan Z, Gao YT and Zheng W (2001) Soyfood intake during adolescence and subsequent risk of breast cancer among Chinese women. *Cancer Epidemiol. Biomarkers Prev.* 10: 483-488.
- Shuber AP, Skoletsky J, Stern R and Handelin BL (1993) Efficient 12-mutation testing in the CFTR gene: a general model for complex mutation analysis. *Hum. Mol. Genet.* 2: 153-158.
- Shure M, Wessler S and Fedoroff N (1983) Molecular identification and isolation of the waxy locus in maize. *Cell* 35: 225-233.
- Staswick PE, Hermodson MA and Nielsen NC (1984) Identification of the cystines which link the acidic and basic components of glycinin subunits. *J. Biol. Chem.* 259: 13431-13435.
- Thanh VH and Shibasaki K (1978) Major proteins of soybean seeds. Subunit structure of beta-conglycinin. *J. Agric. Food Chem.* 26: 692-695.
- Thanos J, Cotterchio M, Boucher BA, Kreiger N and Thompson LU (2006) Adolescent dietary phytoestrogen intake and breast cancer risk (Canada). *Cancer Causes Control* 17: 1253-1261.
- Tsuruki T, Takahata K and Yoshikawa M (2005) Anti-alopecia mechanisms of soymetide-4, an immunostimulating peptide derived from soy β -conglycinin. *Peptides* 26: 707-711.
- Utsumi S (1992) Plant food protein engineering. In Kinsella JE(ed) *Advances in food and nutrition research*. Academic Press, San Diego, CA. pp 89-208.
- Utsumi S, Matsumura Y and Mori T (1997) Structure-function relationships of soy proteins. In S Damodaran, A Paraf, eds, *Food Proteins and Their Applications*. Marcel Dekker, New York, pp 257-291.
- Winn-Deen ES (1996) Automation of molecular genetic methods. Part 2: DNA amplification techniques. *J. Clin. Ligand Assay* 19: 21-26.
- Winter P and Kahl G (1995) Molecular marker technologies for plant improvement. *World J. Microb. Biot.* 11: 438-448.
- Wu AH, Wan P, Hankin J, Tseng CC, Yu MC and Pike MC (2002). Adolescent and adult soy intake and risk of breast cancer in Asian-Americans. *Carcinogenesis* 23: 1491-1496.
- Wu AH, Yu MC, Tseng CC and Pike MC (2008) Epidemiology of soy exposures and breast cancer risk. *Br. J. Cancer* 98: 9-14.
- Yu K, Poysa V, Haffner M, Zhang B and Woodrow L (2005) Absence of the A4 peptide in the G4 glycinin subunit of soybean cultivar Enrei is caused by a point mutation in the *Gy4* gene. *Genet. Mol. Biol.* 28: 440-443.
- Zimmermann K, Schogl D, Plaimauer B and Mannhalter JW (1996) Quantitative multiple competitive PCR of HIV-1 DNA in a single reaction tube. *Bio Tech.* 21: 480-484.
- Zou S, Stansfield C and Bridge J (1998) Identification of new influenza B virus variants by multiplex reverse transcription-PCR and the heteroduplex mobility assay. *J. Clin. Microbiol.* 36: 1544-1548.