Theor Appl Genet (2013) 126:721–731
DOI 10.1007/s00122-012-2013-5

ORIGINAL PAPER

Genetic and chemical analysis of a key biosynthetic step for soyasapogenol A, an aglycone of group A saponins that influence soymilk flavor

Yoshitake Takada · Hiroko Sasama · Takashi Sayama · Akio Kikuchi · Shin Kato · Masao Ishimoto · Chigen Tsukamoto

Received: 6 July 2012 / Accepted: 24 October 2012 / Published online: 11 December 2012

Abstract Although certain saponins in soybean seeds have been reported to have health benefits, group A acetyl saponins cause undesirable bitter and astringent tastes in soy products. Therefore, reduction or elimination of group A saponins is an important target for soybean breeders. A wide survey of cultivated and wild soybean germplasm identified a mutant line that lacked group A saponins. The absence of soyasapogenol A, a group A saponin aglycone, is controlled by a single recessive allele, *sg-5* that mapped genetically near the SSR marker, Satt117, on soybean chromosome 15 (linkage group E). The locus is epistatic to *Sg-1*, which controls the terminal sugar variation on the C-22 sugar chain of soyasapogenol A, and allelic differences at this locus lead to changes in the amount of DDMP saponins and their derivatives group B and E products. These findings provide a new insight into the biosynthetic pathway of soybean saponins, and identify a genetic approach that can be applied to improve the quality of foods produced from soybean.

Communicated by I. Rajcan.

Y. Takada
National Agricultural Research Organization (NARO) Western Region Agricultural Research Center, 1-3-1 Senyu, Zentsuji, Kagawa 765-8508, Japan

Y. Takada · A. Kikuchi · S. Kato
NARO Tohoku Region Agricultural Research Center, 297 Uenodai, Kariwano, Daisen, Akita 019-2112, Japan

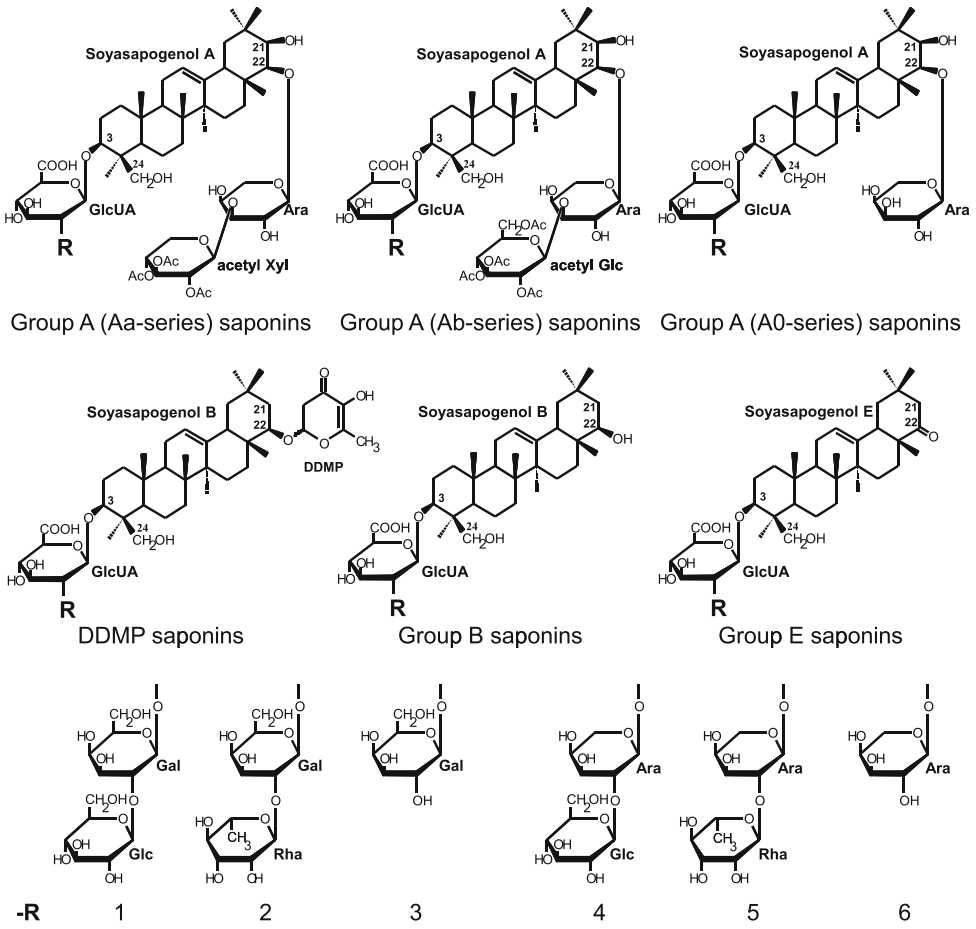
H. Sasama · C. Tsukamoto (✉)
Graduate School of Agriculture, Iwate University, 3-18-8 Ueda, Morioka, Iwate 020-8550, Japan
e-mail: chigen@iwate-u.ac.jp

H. Sasama · T. Sayama · M. Ishimoto
National Institute of Agrobiological Sciences (NIAS), 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan

Introduction

Soybean [*Glycine max* (L.) Merr.] seeds are an excellent source of nutritious proteins, oils, vitamins, and dietary fiber. They also contain other functional food components such as isoflavones, phytic acid, phytosterols, lecithins, and saponins (Sugano 2006). Among these phytochemicals, soybean saponins have become the subject of increasing research interest because of their influence on the quality of traditional foods prepared from soybean.

Soybean saponins are divided into two major groups called group A and DDMP saponins based on the chemical structures of the conjugated aglycone (Fig. 1). Group A acetyl saponins have soyasapogenol A as the aglycone and an acetylated sugar moiety attached to the C-22 hydroxyl group (Taniyama et al. 1988; Shiraiwa et al. 1991a). It is the primary cause of undesirable bitter and astringent tastes of soy foods (Kitagawa et al. 1988; Taniyama et al. 1988; Okubo et al. 1992). DDMP saponins, which contain a DDMP (2,3-dihydro-2,5-dihydroxy-6-methyl-4*H*-pyran-4-one) conjugated soyasapogenol B as the aglycone, are degraded into group B and E saponins during food processing (Kudou et al. 1992, 1993). These contain soyasapogenol B and E aglycones, respectively. Hu et al. (2004a, b) reported that ingested soybean saponins have low absorbability in the human intestine. Neither soybean saponins nor their metabolites could be detected in urine, but partially glycosylated saponin components and soyasapogenol B produced by human intestinal microorganisms in vivo were detected in feces. Consequently, DDMP saponins and their derivatives are considered to have



-R	Group A saponins			DDMP saponins	Group B saponins	Group E saponins
	Aa-series	Ab-series	A0-series			
1. β-D-glucosyl-(1→2)-*O*-β-D-galactose	Aa	Ab	A0-αg	αg	Ba	Bd
2. β-L-rhamnosyl-(1→2)-*O*-β-D-galactose	Au	Ac	A0-βg	βg	Bb	Be
3. β-D-galactose	Ae	Af	A0-γg	γg	Bb'	Be'
4. β-D-glucosyl-(1→2)-*O*-α-L-arabinose	Ax	Ad	A0-αa	αa	Bx	Bf
5. β-L-rhamnosyl-(1→2)-*O*-α-L-arabinose	Ay	Az	A0-βa	βa	Bc	Bg
6. α-L-arabinose	Ag	Ah	A0-γa	γa	Bc'	Bg'

Fig. 1 Chemical structure and nomenclature of triterpene saponins in soybean seeds. Group B and Group E saponins are derivatives of 2,3-dihydro-2,5-dihydroxy-6-metyl-*4H*-pyran-4-one (DDMP)-conjugated soyasapogenol B glycosides. The combination of saponin components in soybean seed hypocotyl is dependent on the genotypes in the *Sg-1*, *Sg-3*, *Sg-4*, and *Sg-5* loci. An arabinose residue of the second sugar moiety at the C-3 position of saponins is controlled by the *Sg-4* locus, whereas a glucose residue of the last sugar moiety at the C-3 position of saponins is controlled by the *Sg-3* locus (Tsukamoto et al. 1992, 1993; Takada et al. 2012). The structure of a sugar chain at the C-22 position of saponins is controlled by multiple alleles at the *Sg-1* locus (Shiraiwa et al. 1990; Tsukamoto et al. 1993; Kikuchi et al. 1999; Takada et al. 2010; Sayama et al. 2012). The presence of Group A saponins is controlled by the *Sg-5* locus (Tsukamoto et al. 1998; Sasama et al. 2010). The nomenclature at the bottom of Fig. 1 is related to the retention time of aglycones from HPLC columns used during this study (Aa, Ab, Ac, Ad, Ae and Af for group A saponins, and Ba, Bb, Bc, Bd and Be for group B and E saponins) (Shiraiwa et al. 1991b, c), by inserting the niche of the table with unused characters (Ax, Ay, Az, Bx, Bf, Bg and Bg′) (Fenwick et al. 1991; Tsukamoto et al. 1993) and sugar chain structures (αg, βg, γg which contain galactose and αa, βa, γa which contain arabinose at the C-3 sugar chain for DDMP and A0-line saponins) (Kudou et al. 1993; Takada et al. 2010, 2012)

beneficial effects on human digestive tract at normal dietary intake amounts (Tsukamoto and Yoshiki 2006) and have been implicated in the prevention of dietary hypercholesterolemia (Fenwick et al. 1991; Murata et al. 2005, 2006) and the suppression of colon cancer cell proliferation (Ellington et al. 2005, 2006). They are also involved in suppressing peroxidation of lipids and protecting the liver by accelerating the secretion of thyroid hormones (Ishii and

Tanizawa 2006). Genetic modification of the composition and quantity of saponins in seeds is an important target for soybean breeding in order to improve both the taste of food produced from soybean and to impart beneficial health properties.

Saponin composition is controlled genetically in soybean seeds, both spatially and temporally. A survey of extracts from hypocotyls of 800 cultivated and 329 wild soybean accessions by high performance liquid chromatography (HPLC) chromatogram resulted in the identification of eight saponin phenotypes based on the compositions of group A saponins (Shiraiwa et al. 1991a; Tsukamoto et al. 1993). The diversity of sugar moieties of soyasapogenol A glycosides is controlled by the combinations of seven alleles (*Sg-1*^{a}, *Sg-1*^{b}, *sg-1*^{0}, *Sg-3*, *sg-3*, *Sg-4*, and *sg-4*) at three independent loci (*Sg-1*, *Sg-3*, and *Sg-4*) (Shiraiwa et al. 1990; Tsukamoto et al. 1993; Kikuchi et al. 1999; Takada et al. 2010, 2012). To take advantage of this genetic diversity, a breeding program was initiated to eliminate group A acetyl saponins, and this resulted in the release of a commercial variety called 'Kinusayaka' in which group A acetyl saponins are absent. 'Kinusayaka' contains the *sg-1*^{0} allele on the *Sg-1* locus (Kato et al. 2007). While this approach successfully reduced undesirable tastes in soy-based foods, group A deacetyl saponin components still remained in the variety. In addition to the structural variations of group A saponins in soybeans described above, a wild soybean accession was identified that is unable to produce soyasapogenol A (Tsukamoto et al. 1998; Sasama et al. 2010).

Recent studies have identified glycosyltransferase and hydroxylase genes involved in saponin biosynthesis in soybean (Shibuya et al. 2006, 2010). However, most of the biosynthesis steps used to produce saponins in soybean remain uncharacterized at the molecular level. The two soybean saponin aglycones, soyasapogenols A and B, seem to be produced from β-amyrin during a series of oxidative reactions, and the aglycones are subsequently converted to the saponin glycosides by a series of glycosylations (Dixon and Sumner 2003). If soyasapogenol A were genetically removed from soybean seeds and the amount of the β-amyrin precursor remains unchanged, it might be possible to increase the amount of group B saponins, which include DDMP conjugates, in compensation for the elimination of group A saponins. In this report, genetic evidence is provided that shows the absence of soyasapogenol A is controlled by a single recessive allele that has been located on the soybean linkage map. Moreover, a quantitative analysis of saponins demonstrated that the amount of DDMP saponins is modulated in the presence of this allele, an observation that suggests this allele could permit soybean varieties to be produced that have improved flavor and enhanced health benefits.

Materials and methods

Plant and chemical materials

Cultivated soybean germplasm was obtained from the NIAS Genebank (680 entries) and NARO Tohoku Region Agricultural Research Center (155 entries). Wild soybean (*Glycine soja* Sieb. and Zucc.) germplasm was obtained from the NAIS Genebank (168 entries), as well as from the private collections of Dr. Okubo (212 entries), Tohoku University, and Dr. Shimamoto (68 entries), Hokkaido University. Soybean saponin aglycones, soyasapogenols A and B, were purchased from Koshiro Seiyaku Co., Ltd., Osaka, Japan.

Extraction of saponin components

Dry seeds of each variety or accession were separated into the hypocotyl and cotyledons with a small blade. Saponins were prepared from each hypocotyl by extracting with a tenfold volume (v/w) of aqueous 70 % ethanol that contained 0.1 % acetic acid for 24 h at 25 °C. The saponins in the extracts were analyzed by HPLC and thin layer chromatography (TLC).

Pre-screening of group A acetyl saponin deficient mutants by HPLC analyses

Extracts were analyzed by HPLC (Waters 510/484 system, Waters, Milford, MA) analysis equipped with an ODS column (YMC packed column AM303 S-5 120A, 4.6 mm ID × 250 mm, YMC, Kyoto, Japan). To detect group A acetyl saponins, an isocratic elution was performed with 34 % acetonitrile that contained 6 % 2-propanol, 0.1 % acetic acid. The solvent flow rate was 1.0 mL/min and UV absorption was monitored at 205 nm as previously described (Tsukamoto et al. 1992).

Separation, detection, and quantification of saponin components by HPLC and TLC analyses

Extracts were analyzed by HPLC (Agilent HP1100 series, Agilent Technologies, Santa Clara, CA) performed with an ODS column (TSK-GEL ODS 100 V, 2.0 mm ID × 150 mm, Tosoh, Japan). To detect group A acetyl saponins, an isocratic elution was performed by using 34 % acetonitrile that contained 6 % 2-propanol and 0.1 % formic acid. The solvent flow rate was 0.15 mL/min and UV absorption was monitored at 205 and 292 nm. Individual saponin components were detected by monitoring the ultraviolet absorbance at a wavelength of 205 nm. The quantity of group A, B, E, and DDMP saponins in each sample was determined by measuring peak areas as previously described (Takagi et al. 2011).

To detect soyasapogenols A and B, aliquots (110 μL) of the hypocotyl extract were mixed with 10 μL of 12N-HCl and hydrolyzed at 85 °C for 2 h. After cooling to room temperature, 120 μL water and 240 μL of diethylether were added, mixed well, and let stand to separate into two phases. The water (lower) layer was discarded by using a microsyringe and the remaining ether layer was evaporated to dryness. The residue was dissolved in 50 μL of ethanol. The ethanol fraction that resulted was applied directly to a pre-coated silica gel plate (Kieselgel 60F-254, Merck) with a glass capillary and developed by benzene:1,4-dioxane:acetic acid (20:5:1, v/v). Soyasapogenols were visualized by spraying with 10 % sulfuric acid and heating at 115 °C for 13 min.

Liquid chromatography–mass spectrometry (LC–MS) analysis of saponin components

To identify saponins denoted in Fig. 1, LC-photodiode array (PDA)/MS/MS on a C30 reverse phase column was employed (Develosil C30-UG-3, 2.0 mm ID × 150 mm, Nomura Chemical, Seto, Japan). Saponins were eluted from the column with a linear gradient of acetonitryl (5–60 %) that contained 0.1 % formic acid [solvent A: acetonitrile with 0.1 % (v/v) formic acid; solvent B: water with 0.1 % formic acid]. Elution was performed at a flow rate of 0.15 mL/min [solvent A was initiated at 5 % (v/v), increased to 95 % (v/v) over 90 min, and then increased to 100 % (v/v) for 5 min)]. Regeneration of the column was accomplished by elution with 5 % (v/v) solvent A for 20 min. Compounds emerging from the column were monitored with a PDA detector (Agilent 1200 LC system) and ion-trap MS (HCTultra PTM-HS, Brukar Daltonics, Billerica, MA). Source settings used for the ionization of saponins were: nebulizer gas flow, 50.00 psi; dry gas flow, 10.00 L/min; capillary temperature, 250 °C. Nitrogen (>99.99 %) and He (>99.99 %) were used as sheath and damping gas, respectively. LC–MS and MS/MS analyses were performed in the positive ion mode for electro spray ionization method using an automatic full scan mode over an *m/z* range from 50 to 2,000. The UV and MS spectra were recorded and analyzed with DataAnalysis software ver. 3.4 (Brukar Daltonics).

Soybean hybridization

The parental material used to develop segregating populations for the genetic analysis of group A saponin phenotypes are identified in Table 1. The wild accession, B01082 (Col/Hokkaido/1991/Shimamoto-9), which lacked soyasapogenol A, was crossed with 'Suzuyutaka' (Cross 1), and the breeding line 'Tohoku 152' which contained the soyasapogenol A null allele, was crossed with a normal variety, 'Shiro-sennari' (Cross 2) (Sayama et al. 2012). Both crosses were carried out and the segregating populations were grown at the NARO Tohoku Region Agricultural Research Center, Daisen, Akita, Japan.

SSR marker genotyping and mapping

Total DNA was extracted from seed flour with an Automatic DNA Isolation System according to Plant DNA Extraction Protocol version 2 (PI-50a; Kurabo, Osaka, Japan). PCR amplification and detection of SSR markers were performed as described previously (Hwang et al. 2009). Genotyping was performed with previously developed SSR markers (Cregan et al. 1999; Song et al. 2004). The SSR and sequence-tagged site (STS) markers used were developed at Chiba University (Xia et al. 2007), and EST-derived SSR primer pairs were created at the Kazusa DNA Research Institute (Hisano et al. 2007). MAPMAKER/EXP v. 3.0 software was used to analyze linkage between markers. Genetic distances (cM) were calculated with Kosambi's mapping function (Kosambi 1944). Linkage maps were drawn with MapChart software (Voorrips 2002).

Results

Detection of a mutant accession lacking group A saponins

Through wide survey of the soybean germplasm stocks, one wild accession, B01082 (Col/Hokkaido/1991/Shimamoto-9), collected in Hokkaido region in Japan, emerged as a mutant line that lacked group A saponins (Fig. 2) confirming previous results of Tsukamoto et al. (1998). Most of the cultivated and wild soybean accessions exhibited several peaks upon HPLC analysis that correspond to either Aa or Ab series saponins. Two accessions, 'A-b(F)' and JP-36121 (Col/Kumamoto/1984/Seriguchi-1), contained a different series of group A saponins that lacked the terminal acetylated sugar on the C-22 sugar chain (Kikuchi et al. 1999; Takada et al. 2010) (Figs. 1, 2). LC–MS analysis revealed that hypocotyl extracts from B01082 lacked peaks with retention times expected for group A saponins, and yielded no peaks with *m/z* 457.3 and 975.5 that are characteristic for group A saponins. TLC patterns of hydrolyzed hypocotyl extracts of B01082 also lacked a signal at Rf 0.56 due to soyasapogenol A (Fig. 3). By contrast, extracts from the normal varieties, 'Ohsuzu', 'Shiro-sennari', and another mutant, 'A-b(F)', gave two bands at 0.56 and 0.65 Rf value that correspond to

Table 1 Main component of group A saponins in soybean seeds used in this study

Variety, breeding line, and wild accession	Group A saponins	
	Main component	Predicted genotype
Shiro-sennari	Aa	*Sg-5*, *Sg-1*^{a}
Suzuyutaka, Ohsuzu	Ab	*Sg-5*, *Sg-1*^{b}
B01082 (wild accession)	Absence	*sg-5*, *Sg-1*^{a}
Tohoku 152	Absence	*sg-5*, *Sg-1*^{b}
JP-36121 (wild accession)	A0-αg	*Sg-5*, *sg-1*^{o}
A-b(F)	A0-αg	*Sg-5*, *sg-1*^{o}
Kinusayaka	A0-αg	*Sg-5*, *sg-1*^{o}

soyasapogenols A and B, which are the aglycones of group A and DDMP, respectively. Hydrolysates from B01082 did contain soyasapogenol B, but did not exhibit unique bands after TLC that were not detected in the other samples. These results establish that B01082 does not have the ability to synthesize the group A saponin aglycone, soyasapogenol A.

Genetic inheritance of the soyasapogenol A null allele, *sg-5*

The inheritance of soyasapogenol A null allele in B01082 was examined in the individual F_1 and F_2 seeds derived from a cross between 'Suzuyutaka' (Ab type) and B01082 (Table 2 Cross 1) and a cross between 'Shiro-sennari' (Aa

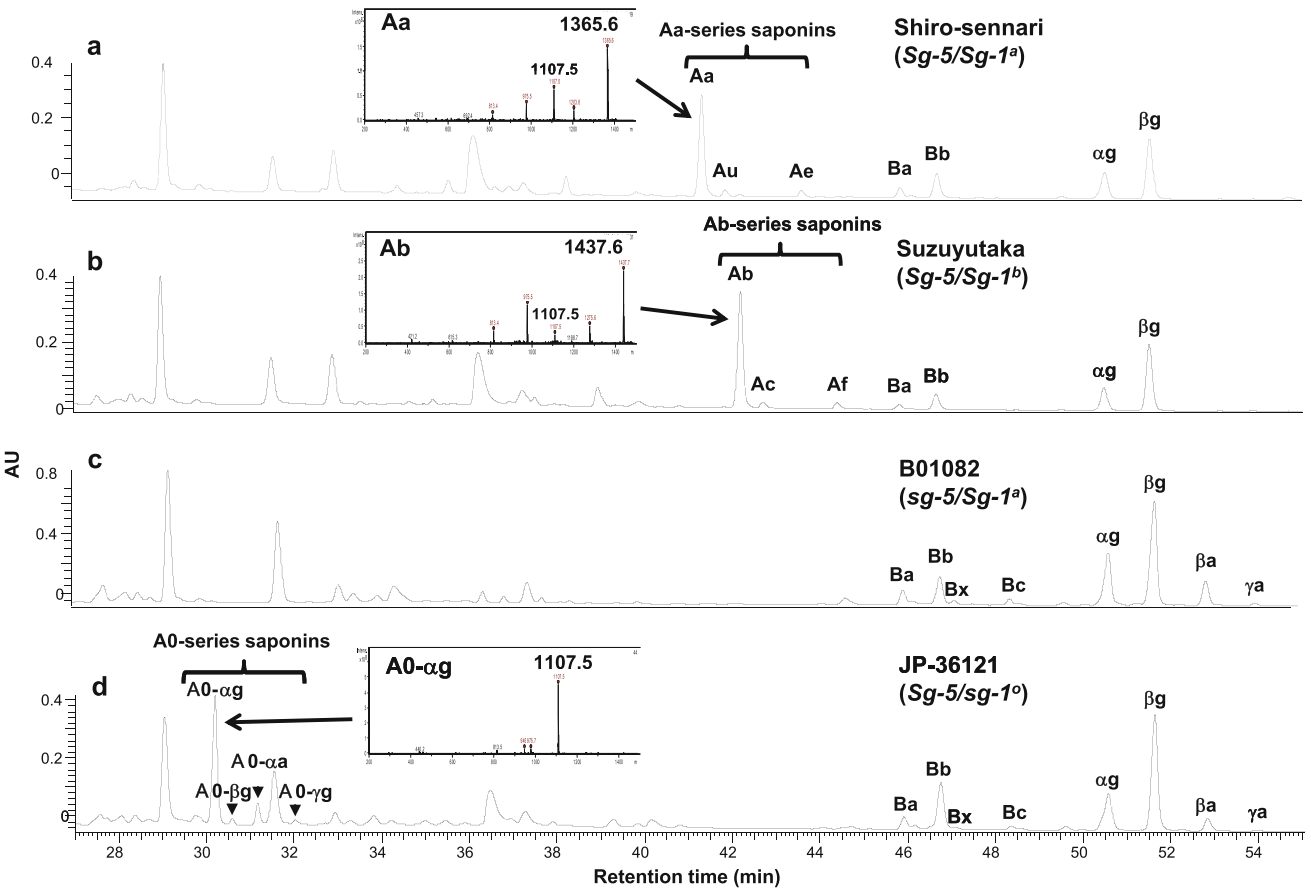


Fig. 2 Separation and detection of saponin components in hypocotyl extracts of soybean seeds used in this study. Extracts from seed hypocotyls by aqueous 70 % ethanol were separated by high performance liquid chromatography (HPLC) and monitored by the UV absorption at 205 nm. **a** 'Shiro-sennari' (*Sg-5/Sg-1*^{a} genotype for Aa series of group A saponins), **b** 'Suzuyutaka' (*Sg-5/Sg-1*^{b} genotype for Ab series of group A saponins), **c** B01082 (*sg-5* genotype for absence of group A saponin), **d** JP-36121 (*Sg-5/sg-1*^{o} genotype for A0 series of group A saponins). The chemical structures and names of saponin components were indicated in Fig. 1. As the scale of the absorbance unit (AU) of sample c (B01082) is a half of the others, the total amount of soyasapogenol B-conjugated saponins (such as saponin αg, βg, Ba and Bb) should be two times or more than the others

type) and 'Tohoku 152' (Table 2 Cross 2). 'Tohoku 152' is a breeding line that contains the *sg-5* null allele from B01082 as well as null alleles for each of the three lipoxygenase isozymes that cause grassy flavors to be produced during the preparation of soy product (Sasama et al. 2010). All F_1 seeds from the two crosses contained both saponin Aa and Ab. This result is consistent with the conclusion that the lack of soyasapogenol A is controlled by a single recessive allele and the parents each had different co-dominant alleles at the *Sg-1* locus. The segregation pattern of individuals in the two F_2 populations were consistent statistically with a 3 (Aa type):6 (AaAb type):3 (Ab type):4 (absent) ratio. These results show that *Sg-5* and *Sg-1* loci are independently inherited and the *Sg-5* locus is epistatic to the *Sg-1* locus that controls the structure of a sugar chain at C-22 position of group A saponin.

Linkage analysis of the *Sg-5* locus that control the presence of soyasapogenol A

To locate the *Sg-5* locus in the genome, a linkage analysis was performed using SSR markers (Fig. 4). As used previously, the symbol *sg-5* denoted the null allele that caused soyasapogenol A to be absent (Sasama et al. 2010). The *Sg-5* locus mapped 1.2 cM from the SSR marker Satt117 on chromosome 15 (linkage group E) of the soybean genome in the segregating population of Cross 2. By contrast, the *Sg-1* locus was shown previously to be located on chromosome 7 (linkage group M) (Takada et al. 2010; Sayama et al. 2012).

Effect of allele combinations at the *Sg-5* locus on saponin composition

Because B01082 differed substantially from cultivated soybeans in seed size, maturity and other agronomic traits, the amount of group A and DDMP saponins in seed hypocotyls from the progeny from the cross 'Shiro-sennari' and 'Tohoku 152' and the two parents were compared (Fig. 5a). The female parent, 'Shiro-sennari', contained 1.25 nmol of group A saponins and 2.2 nmol of DDMP saponins/100 g seed hypocotyl dry weight. In contrast, the male parent, 'Tohoku 152,' contained only DDMP saponins, and had a total content as high as 4.06 nmol/100 g. F_1 hybrid seeds from the cross had an intermediate amount of group A and DDMP saponins between the parents, but the total saponin content of 4.00 nmol/100 g was comparable to that in 'Tohoku 152'. These results are interpreted as showing the allele combination at the *Sg-5* locus determines the amount of group A saponins, and eventually the amount of DDMP.

To assess this phenomenon further, F_2 seeds derived from F_1 hybrid plants of a cross between 'Shiro-sennari' and 'Tohoku 152' were divided into three groups according to the allele combination at the *Sg-5* locus (Fig. 5b). The genotypes of the *Sg-5* locus were determined based on the marker genotypes used for the linkage analysis. Some F_2 seeds exhibited recombination between the two proximal co-dominant markers, Satt117 and GMES0332, and they were excluded from the analysis. Because all the seeds identified to have the *sg-5/sg-5* allele combination lacked group A saponins. We concluded that the two marker genotypes classified the allele combination at the *Sg-5* locus accurately. The F_2 seeds with heterologous alleles, *Sg-5/sg-5*, as well as the F_1 hybrid seeds, contained group A and DDMP saponins at intermediate levels between the two genotypes that contained homologous allele combinations. We conclude from these results that the reduction and elimination of group A saponins result in an increase of DDMP saponins and their derivatives group B and E products.

Discussion

Triterpenoid saponins are common secondary metabolites in soybean seeds, and these are a group of diversified chemicals that are the consequence of a variety of hydroxylations and subsequent glycosylations (Fig. 1). Soybean saponins are divided into group A and DDMP saponins based on aglycon structures. The decrease or elimination of group A saponins in soybean seeds is an important target of soybean breeding, because they cause a bitter and astringent taste in foods prepared from the seeds (Okubo et al. 1992). While there have been earlier reports that soyasapogenol A-conjugated saponins (group A saponins) were exclusively localized in hypocotyls (Taniyama et al. 1988; Shimoyamada et al. 1990), recent results have demonstrated the presence of group A saponins in cotyledons (Tsukamoto et al. 2011). As a considerable amount of partially and fully acetylated group A saponins was detected in cotyledons, genetic elimination of group A saponins from cotyledons could improve the taste of soy foods.

A wild soybean accession, B01082, was identified that lacks group A saponins. It is unable to synthesize soyasapogenol A, the discriminative aglycone of group A saponins. Two mutant lines have been already identified from cultivated and wild soybeans that are free from ordinary group A saponins such as Aa and Ab series (Tsukamoto et al. 1993; Kikuchi et al. 1999). A single recessive allele, *sg-1^o*, causes the accumulation of group A deacetyl saponins (A0 series saponins) instead of the group A acetyl saponins, and has been successfully incorporated into the breeding program to yield a new commercial variety, 'Kinusayaka'. 'Kinusayaka' does not produce a bitter or astringent taste in traditional soy-based foods (Kato et al. 2007). However, group A saponins, including group A

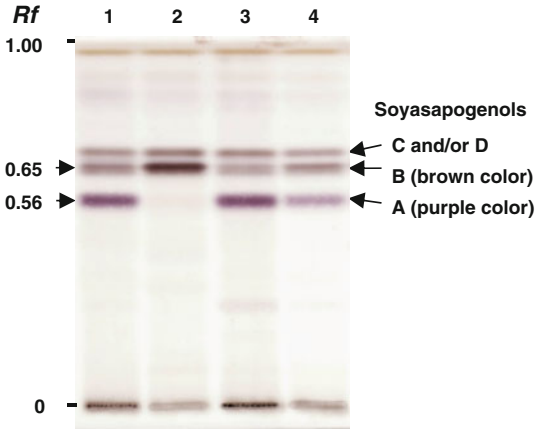


Fig. 3 Separation and detection of soyasapogenols in soybean seed hypocotyls by thin layer chromatography (TLC). Aqueous 70 % ethanol extracts were hydrolyzed and then separated by TLC. *1* 'Shiro-sennari' (*Sg-5/Sg-1*^{a} genotype for Aa series of group A saponins), *2* B01082 (*sg-5* genotype for absence of group A saponin), *3* 'Ohsuzu' (*Sg-5/Sg-1*^{b} genotype for Ab series of group A saponins), *4* A-b(F) (*Sg-5/sg-1*^{o} genotype for A0 series of group A saponins). *Arrows* indicate the positions of standard soyasapogenols A (Rf = 0.56) and B (Rf = 0.65) in this analytical condition (Shiraiwa et al. 1991b, c; Sasama et al. 2010). Rf means relative mobility to the front. Soyasapogenol A shows specific *purple* color which is different from soyasapogenols B, C and D (*brown* color) (color figure online)

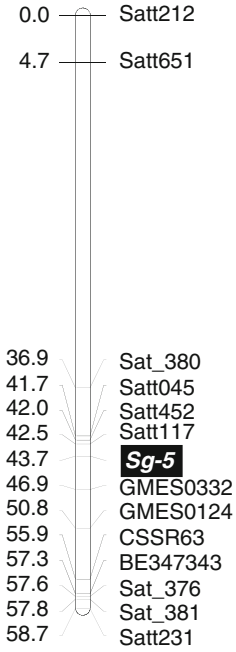


Fig. 4 A genetic linkage map of *Sg-5* locus positioned on soybean chromosome 15 (linkage group E) along with SSR markers. The markers and *Sg-5* locus were mapped in the F_2 population derived from Cross 2 (Table 2) between 'Shiro-sennari' (*Sg-5*) and 'Tohoku 152' (*sg-5*). Genetic distances of SSR markers and *Sg-5* from Satt212 are shown in centimorgans (cM)

deacetyl saponins, have almost no beneficial effects on human health (Tsukamoto and Yoshiki 2006). Thus, an increase of DDMP (group B) saponins achieved simultaneously with a decrease of group A saponins has become an objective in our soybean breeding program. To achieve this objective, the genetic trait responsible for the lack of soyasapogenol A in B01082 was successfully introduced into a new breeding line 'Tohoku 152'. 'Tohoku 152' and its progeny accumulated an increased amount of DDMP saponin in response to eliminating soyasapogenol A from the seed (Fig. 5). These breeding lines exhibited no detectable adverse agronomic traits such as poor germination, sterility and decreased seed production. They are therefore very promising germplasm for use in generating soybeans destined for the soy-based food production.

Table 2 Segregation of F_1 and F_2 individuals for group A saponin phenotypes in soybean seed hypocotyl

Parents generation	Group A saponin		Number of seeds	Expected ratio	Chi-square value	*df*	Probability
	Phenotype	Predicted genotype					
Cross 1							
P1: Suzuyutaka	Ab	*Sg-5*, *Sg-1*^{b}					
P2: B01082	Absence	*sg-5*, *Sg-1*^{a}					
F_1 individuals	AaAb	*Sg-5*, *Sg-1*^{a}/*Sg-1*^{b}	4				
F_2 individuals	Aa	*Sg-5*, *Sg-1*^{a}	21	3	3.69	3	0.29
	AaAb	*Sg-5*, *Sg-1*^{a}/*Sg-1*^{b}	43	6			
	Ab	*Sg-5*, *Sg-1*^{b}	17	3			
	Absence	*sg-5*, *Sg-1*^{a} and/or *Sg-1*^{b}	38	4			
Cross 2							
P1: Shiro-sennari	Aa	*Sg-5*, *Sg-1*^{a}					
P2: Tohoku 152	Absence	*sg-5*, *Sg-1*^{b}					
F_1 individuals	AaAb	*Sg-5*, *Sg-1*^{a}/*Sg-1*^{b}	7				
F_2 individuals	Aa	*Sg-5*, *Sg-1*^{a}	31	3	1.50	3	0.68
	AaAb	*Sg-5*, *Sg-1*^{a}/*Sg-1*^{b}	66	6			
	Ab	*Sg-5*, *Sg-1*^{b}	40	3			
	Absence	*sg-5*, *Sg-1*^{a} and/or *Sg-1*^{b}	45	4			

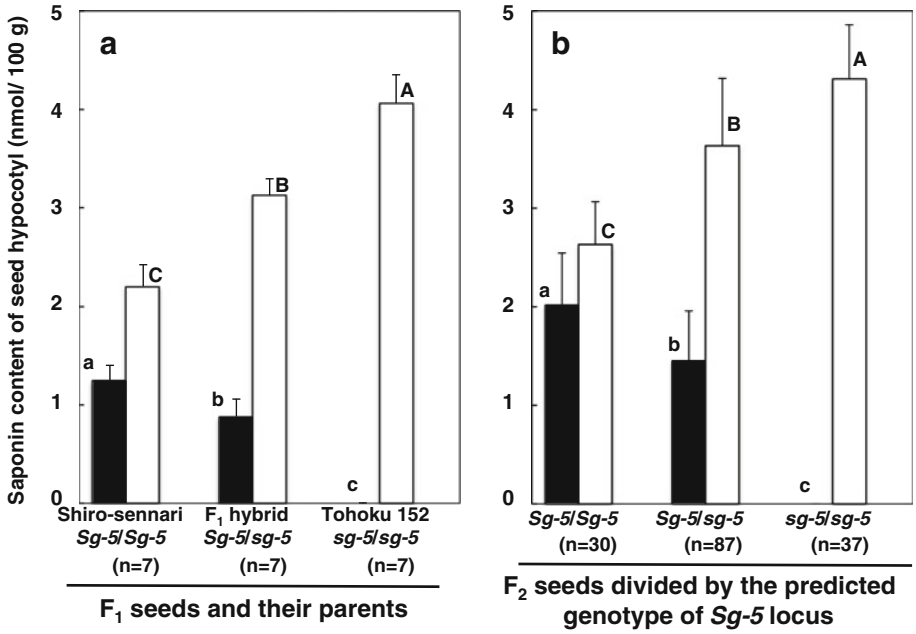


Fig. 5 Amount of group A and DDMP saponins in seed hypocotyls of the progenies from Cross 2 (Table 2) between 'Shiro-sennari' (*Sg-5*) and 'Tohoku 152' (*sg-5*). *Black bars* show a total amount of group A saponins, whereas *open bars* show a total amount of DDMP saponins and their degraded products (group B and E saponins). Data are mean ± SD of values from the described number of seed hypocotyls. *Different letters* indicate statistically significant ($P < 0.05$) differences among parental line and the F_1/F_2 seeds (Tukey–Kramer multiple comparison test). **a** 'Shiro-sennari' (*Sg-5*), 'Tohoku 152' (*sg-5*), and their F_1 hybrid seeds, **b** F_2 seeds divided by the predicted genotype of *Sg-5* locus. Numbers of the respective seeds are shown in *parentheses*

Experiments reported earlier (Kato et al. 2007) have demonstrated that modification of group A saponins in soybean seeds improves soymilk taste. In a sensory test, soymilk made from 'Kinusayaka' was assessed to have no beany flavor and a plain taste. Soymilk from this variety lacked the astringent and bitter tastes typical of products made from soybeans that contain unmodified group A saponins. By contrast, exploratory sensory evaluations of the soymilk made from 'Tohoku 152' also lacked the beany flavor, but had a rich taste flavor that was different from the ones associated with soymilk prepared from either normal soybeans or 'Kinusayaka' (unpublished data). Presently, there is no information on the causal relationship between the change in taste of soymilk prepared from either 'Tohoku 152' or 'Kinusayaka' and the modification of saponin composition. However, it may be possible to use the soybean lines that have now been generated with different saponin compositions for this purpose.

The selection of a target phenotype for saponin composition required an immense expenditure in time and effort, because seed germination is disabled when the hypocotyls are excised to evaluate saponin composition. However, as reported in this communication, a genetic linkage analysis of the *Sg-5* locus that controls the presence of soyasapogenol A was completed and soybean SSR markers on Chr. 15 (Hwang et al. 2009) that are linked to it were identified. This positional information permits identification of saponin phenotypes by marker-assisted selection as demonstrated by the selection of the F_2 seeds derived from a cross between 'Shiro-sennari' and 'Tohoku 152' (Fig. 5). In this regard, three independent loci (*Sg-1*, *Sg-3*, and *Sg-4*) that control the sugar moieties of soyasapogenol A glycosides have previously been characterized (Takada et al. 2010, 2012), and they, together with the *Sg-5* locus, can be used to manipulate saponin composition.

Figure 6 presents a putative biosynthetic pathway for soyasapogenol A and B. Of particular importance to this discussion are the hydroxyl groups attached to C-21, C-22 and C-24 of the basic β-amyrin molecule. Kitagawa et al. (1982) established that C-21 of soyasapogenol A is hydroxylated, whereas soyasapogenol B is not. More recently, Shibuya et al. (2006) reported that an expressed sequence tag of a cytochrome P450 gene (CYP93E1) in soybean had the ability to hydroxylate the C-24 position of β-amyrin and sophoradiol (22-hydroxy-4-amyrin). The same group identified a cytochrome P450 gene (CYP72A61) that has hydroxylating activity for the C-22 position of β-amyrin (Ebizuka et al. 2011). In this communication, the *Sg-5* genetic locus was characterized that modulates the presence or absence of the C-21 hydroxyl to produce soyasapogenol A and B, respectively. The LC–MS/MS analysis made it clear that no extra and unique saponins were found in the mutant seeds with the *sg-5* null allele that lack group A saponins (Sasama et al. 2010).

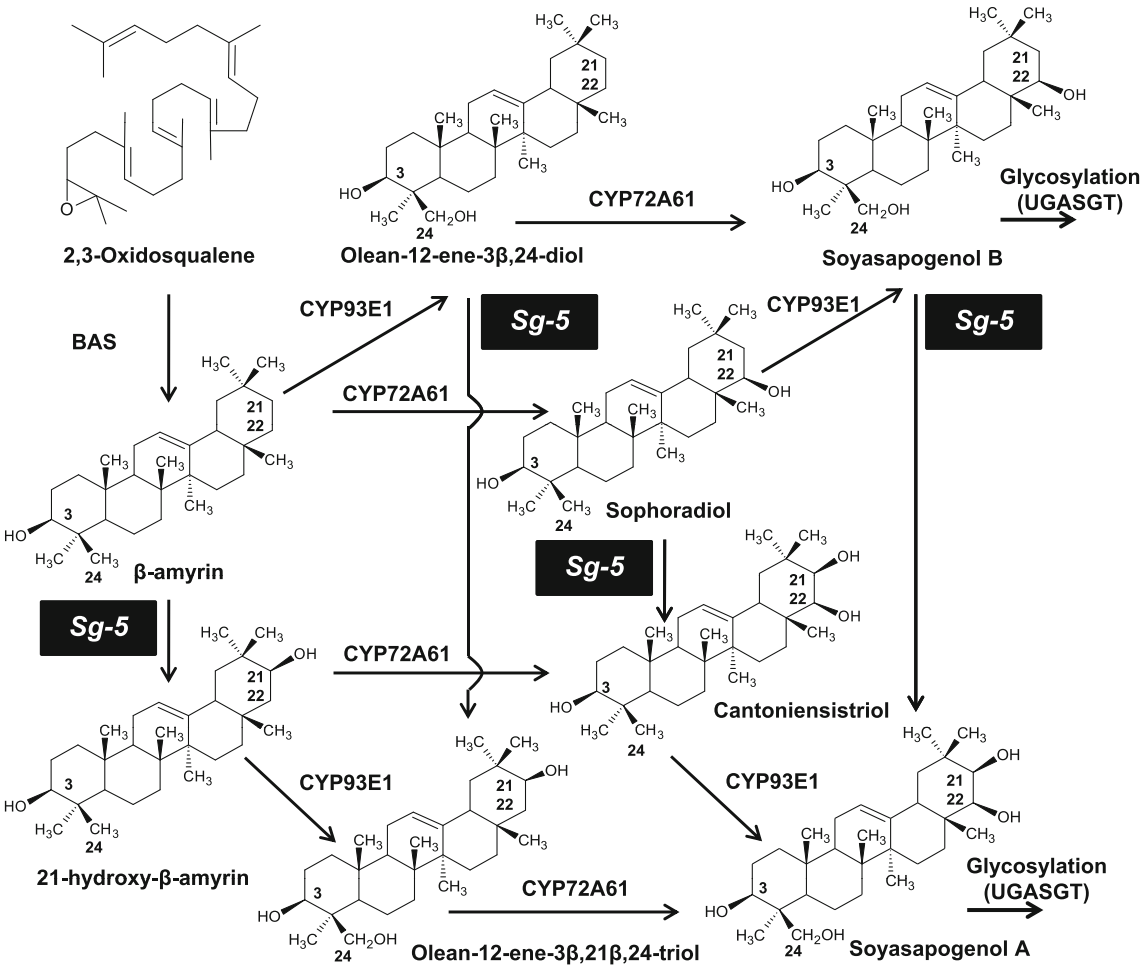


Fig. 6 Proposed biosynthetic pathway of soyasapogenols A and B in soybean. The *Sg-5* gene might encode a 21-hydroxylase gene, which hydroxylates the C-21 position of the four candidate precursors, β-amyrin, sophoradiol, olean-12-ene-3β,24-diol, and soyasapogenol B to produce soyasapogenol A

In addition, the glycosides of soyasapogenol B increased in mutant seeds. It therefore seems likely that the gene at the *Sg-5* locus encodes a 21-hydroxylase gene. As indicated in the putative biosynthetic pathway of Fig. 6, such an enzyme could potentially hydroxylate β-amyrin, sophoradiol, olean-12-ene-3β,24-diol and soyasapogenol B to produce soyasapogenol A. To establish the synthetic pathway for group A saponins, it will be necessary to isolate the gene product from *Sg-5* and characterize its specificity against these candidate precursors.

Acknowledgments We thank Dr. Okubo (Tohoku University) and Dr. Shimamoto (Hokkaido University) for providing private germplasm collections. This work was supported by a grant from the Ministry of Agriculture, Forestry, and Fisheries of Japan (Research Project for Utilizing Advanced Technologies in Agriculture, Forestry and Fisheries, 18063) as well as, in part, by the Program for Promotion of Basic and Applied Researches for Innovations in Bio-oriented Industry (BRAIN). We thank Dr. Niels C. Nielsen (Professor Emeritus, Purdue University) for critically reading this manuscript.

References

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, VanToai TT, Lohnes DG, Chung J, Specht JE (1999) An integrated genetic linkage map of the soybean genome. Crop Sci 39:1464–1490

Dixon RA, Sumner LW (2003) Legume natural products: understanding and manipulating complex pathways for human and animal health. Plant Physiol 131:878–885

Ebizuka Y, Shibuya M, Wakita E (2011) C-22 hydroxylase. Jpn Pat Appl 269:P2011–24534A

Ellington AA, Berhow MA, Singletary KW (2005) Induction of macroautophagy in human colon cancer cells by soybean B-group triterpenoid saponins. Carcinogenesis 26:159–167

Ellington AA, Berhow MA, Singletary KW (2006) Inhibition of Akt signaling and enhanced ERK1/2 activity are involved in induction of macroautophagy by triterpenoid B-group soyasaponins in colon cancer cells. Carcinogenesis 27:298–306

Fenwick GR, Price KR, Tsukamoto C, Okubo K (1991) Saponins. In: D'Mello JPF, Duffus CM, Duffus JH (eds) Toxic substances in crop plants. The Royal Society of Chemistry, Cambridge, pp 285–327

Hisano H, Sato S, Isobe S, Sasamoto S, Wada T, Matsuno A, Fujishiro T, Yamada M, Nakayama S, Nakamura Y, Watanabe S, Harada K, Tabata S (2007) Characterization of the soybean genome using EST-derived microsatellite markers. DNA Res 14:271–281

Hu J, Reddy MB, Hendrich S, Murphy PA (2004a) Soyasaponin I and sapongenol B have limited absorption by Caco-2 intestinal cells and limited bioavailability in women. J Nutr 134: 1867–1873

Hu J, Zheng YL, Hyde W, Hendrich S, Murphy PA (2004b) Human fecal metabolism of soyasaponin I. J Agric Food Chem 52:2689–2696

Hwang TY, Sayama T, Takahashi M, Takada Y, Nakamoto Y, Funatsuki H, Hisano H, Sasamoto S, Sato S, Tabata S, Kono I, Hoshi M, Hanawa M, Yano C, Xia Z, Harada K, Kitamura K, Ishimoto M (2009) High-density integrated linkage map based on SSR markers in soybean. DNA Res 16:213–225

Ishii Y, Tanizawa H (2006) Effects of soyasaponins on lipid peroxidation through the secretion of thyroid hormones. Biol Pharm Bull 29:1759–1763

Kato S, Yumoto S, Takada Y, Kono Y, Shimada S, Sakai T, Shimada H, Takahashi K, Adachi T, Tabuchi K, Kikuchi A (2007) A new soybean cultivar 'Kinusayaka' lacking three lipoxygenase isozymes and group A acetyl saponin. Bull Natl Agric Res Cent Tohoku Reg 107:29–42

Kikuchi A, Tsukamoto C, Tabuchi K, Adachi T, Okubo K (1999) Inheritance and characterization of a null allele for group A acetyl saponins found in a mutant soybean (*Glycine max* (L.) Merrill). Breed Sci 49:167–171

Kitagawa I, Yoshikawa M, Wang HK, Saito M, Tosirisuk V, Fujiwara T, Tomita K (1982) Revised structures of soyasapogenol A, soyasapogenol B, and soyasapogenol E, oleanene sapogenols from soybean, structures of soyasaponin I, soyasaponin II, and soyasaponin III. Chem Pharm Bull 30:2294–2297

Kitagawa I, Taniyama T, Nagahama Y, Okubo K, Yamauchi F, Yoshikawa M (1988) Saponin and sapogenol XLII. Structures of acetyl-soyasaponins A_1, A_2, and A_3, astringent partially acetylated bisdesmosides of soyasapogenol A, from American soybean, the seeds of *Glycine max* Merrill. Chem Pharm Bull 36:2819–2828

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Kudou S, Tonomura M, Tsukamoto C, Shimoyamada M, Uchida T, Okubo K (1992) Isolation and structural elucidation of the major genuine soybean saponin. Biosci Biotech Biochem 56:142–143

Kudou S, Tonomura M, Tsukamoto C, Uchida T, Sakabe T, Tamura N, Okubo K (1993) Isolation and structural elucidation of DDMP-conjugated soyasaponins as genuine saponins from soybean seeds. Biosci Biotech Biochem 57:546–550

Murata M, Houdai T, Morooka A, Matsumori M, Oishi T (2005) Membrane interaction of soyasaponins in association with their antioxidation effect. Soy Protein Res 8:81–85

Murata M, Houdai T, Yamamoto H, Matsumori M, Oishi T (2006) Membrane interaction of soyasaponins in association with their antioxidation effect—analysis of biomembrane interaction. Soy Protein Res 9:82–86

Okubo K, Iizima M, Kobayashi Y, Yoshikoshi M, Uchida T, Kudou S (1992) Components responsible for the undesirable taste of soybean seeds. Biosci Biotech Biochem 56:99–103

Sasama H, Takada Y, Ishimoto M, Kitamura K, Tsukamoto C (2010) Estimation of the mutation site of a soyasapogenol A-deficient soybean [*Glycine max* (L.) Merr.] by LC–MS/MS profile analysis. In: Cadwallader KR, Chang SKC (eds) Chemistry, texture, and flavor of soy, ACS symposium series 1059. Oxford University Press, Washington, pp 91–102

Sayama T, Ono E, Takagi K, Takada Y, Horikawa M, Nakamoto Y, Hirose A, Sasama H, Ohashi M, Hasegawa H, Terakawa T, Kikuchi A, Kato S, Tatsuzaki N, Tsukamoto C, Ishimono M (2012) The *Sg-1* glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean. Plant Cell 24:2123–2138

Shibuya M, Hoshino M, Katsube Y, Hayashi H, Kushiro T, Ebizuka Y (2006) Identification of β-amyrin and sophoradiol 24-hydroxylase by expressed sequence tag mining and functional expression assay. FEBS J 273:948–959

Shibuya M, Nishimura K, Yasuyama N, Ebizuka Y (2010) Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*. FEBS Lett 584:2258–2264

Shimoyamada M, Kudou S, Okubo K, Yamauchi F, Harada K (1990) Distribution of saponin constituents in some varieties of soybean plant. Agric Biol Chem 54:77–81

Shiraiwa M, Yamauchi F, Harada K, Okubo K (1990) Inheritance of "group A saponin" in soybean seed. Agric Biol Chem 54:1347–1352

Shiraiwa M, Harada K, Okubo K (1991a) Composition and content of saponins in soybean seed according to variety, cultivation year and maturity. Agric Biol Chem 55:323–331

Shiraiwa M, Harada K, Okubo K (1991b) Composition and structure of "group B saponin" in soybean seed. Agric Biol Chem 55:911–917

Shiraiwa M, Kudou S, Shimoyamada M, Harada K, Okubo K (1991c) Composition and structure of "group A saponin" in soybean seed. Agric Biol Chem 55:315–322

Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, Delannay X, Specht JE, Cregan PB (2004) A new integrated genetic linkage map of the soybean. Theor Appl Genet 109:122–128

Sugano M (2006) Soy in health and disease prevention. CRC Press, Taylor and Francis Group, New York

Takada Y, Sayama T, Kikuchi A, Kato S, Tatsuzaki N, Nakamoto Y, Suzuki A, Tsukamoto C, Ishimoto M (2010) Genetic analysis of variation in sugar chain composition at the C-22 position of group A saponin in soybean, *Glycine max* (L.) Merrill. Breed Sci 60:3–8

Takada Y, Tayama I, Sayama T, Sasama H, Saruta M, Kikuchi A, Ishimoto M, Tsukamoto C (2012) Genetic analysis of variations in the sugar chain composition at the C-3 position of soybean seed saponins. Breed Sci 61:646–652

Takagi K, Nishizawa K, Hirose A, Kita A, Ishimoto M (2011) Manipulation of saponin biosynthesis by RNA interference-mediated silencing of β-amyrin synthase gene expression in soybean. Plant Cell Rep 30:1835–1846

Taniyama T, Nagahama Y, Yoshikawa M, Kitagawa I (1988) Saponin and sapogenol XLIII. Acetyl-soyasaponins A_4, A_5, and A_6, new astringent bisdesmosides of soyasapogenol A, from Japanese soybean, the seeds of *Glycine max* Merrill. Chem Pharm Bull 36:2829–2839

Tsukamoto C, Yoshiki Y (2006) Soy saponin. In: Sugano M (ed) Soy in health and disease prevention. CRC Press, Taylor and Francis Group, New York, pp 155–172

Tsukamoto C, Kikuchi A, Kudou S, Harada K, Kitamura K, Okubo K (1992) Group A acetyl saponin-deficient mutant from the wild soybean. Phytochemistry 31:4139–4142

Tsukamoto C, Kikuchi A, Harada K, Kitamura K, Okubo K (1993) Genetic and chemical polymorphisms of saponins in soybean seed. Phytochemistry 34:1351–1356

Tsukamoto C, Kikuchi A, Harada K, Kitamura K, Iwasaki T, Okubo K (1998) Genetic improvement of saponin components in soybean seed. In: Chainuvati C, Sarobol N (eds) Soybean feeds the world. Proceedings (Suppl) of 5th World soybean research conference, Kasetsart University Press, Bangkok, p 80–85

Tsukamoto C, Takada Y, Omizu Y, Kon T, Kikuchi A, Kato S, Ishimoto M, Kitamura K (2011) Soyasapogenol A-conjugated saponin glycosides detected in the seed cotyledon of soybeans. Proceeding CD of 12th ASEAN food conference 2011, BITEC, Bangkok, p 391–394

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93:77–78

Xia Z, Tsubokura Y, Hoshi M, Hanawa M, Yano C, Okamura K, Ahmed TA, Anai T, Watanabe S, Hayashi M, Kawai T, Hossain KG, Masaki H, Asai K, Yamanaka N, Kubo N, Kadowaki K, Nagamura Y, Yano M, Sasaki T, Harada K (2007) An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using a single F_2 population. DNA Res 14:257–269