Journal of Plant Biology, March 2005, 48(1) : 106-112

Quantitative Trait Loci Associated with Oligosaccharide and Sucrose Contents in Soybean (*Glycine max* L.)

Hyeun-Kyeung Kim[1]*, Sung-Taeg Kang[2], Jun-Hyeun Cho[1], Myoung-Gun Choung[3], and Duck-Yong Suh[1]

[1]Yeongnam Agricultural Research Institute, NICS, RDA, Milyang 627-130, Korea
[2]Research Coordination Division, RDA, Suwon 441-707, Korea
[3]Department of Pharmacognosy Material Development, Samcheok National University, Samcheok 245-711, Korea

Oligosaccharides and sucrose are very important nutritional components in soybean seeds. However, little information is available about their inheritance. We used molecular markers to identify the genomic regions significantly associated with the quantitative trait locus (QTL) that controls oligosaccharide and sucrose contents in segregating $F_{2:10}$ RI lines. Two related, but independent, QTLs were identified for oligosaccharides -- near marker satt546 on linkage group (LG) D1b+W and satt278 on LG L. Four others, for sucrose content, were located at LG B1 (satt197), D1b+W (satt546), and L (satt523 and satt278). Finally, we found two common QTLs, on LG D1b+W and L, that are associated with both oligosaccharides and sucrose.

Keywords: oligosaccharides, QTL, soybean, SSR marker, sucrose

Soybean (*Glycine max*), a legume crop, is an excellent food and feed source all over the world, with a total production in 2001 of about 130 million tons. The primary constituents of its seed are protein, oil, and soluble carbohydrates. Although soybean oligosaccharides are a very important nutritional component, little is known about the inheritance of carbohydrates in this species. Greater than 90% of the sugars in ripe seeds are present as sucrose, raffinose, and stachyose (Kawamura, 1954). These three are important determinants for consumer acceptance of soybean products.

Although soybean oligosaccharides have long been considered undesirable components that cause flatulence (Omosaiye et al., 1978; Borejszo and Khan, 1992), they are also a safe, 'probiotic material' approved by the FDA for use in the USA (Hoover, 1993). Because of their economic value to the food and feed industries, attempts have been made to recover them from defatted soybean meal (Kim et al., 2003). For example, in the sweetener industry, new technologies have been applied so that high-oligosaccharide syrups can be commercially produced from the soybean (Koga et al., 1993).

Sucrose is the most common form of reduced carbon product generated from photosynthesis. Once synthesized, it is imported into the phloem and transported long distances to organs, such as developing leaves, flowers, fruits, and roots, that are unable to reduce carbon themselves (Aldape et al., 2003). Sucrose supplementation has positive effects on soybean growth but also suppresses photosynthesis (Abdin et al., 1998). Developing seeds synthesize storage compounds from imported sucrose during their maturation phase. Phloem-unloading and transport and transfer processes may play important roles during seed formation (Patrick and Offler, 1995; Weber et al., 1998).

Recent progress in the use of molecular markers has increased the efficiency of plant breeding programs (Paterson and Tanksley, 1991). Information concerning the inheritance and underlying genetic control of oligosaccharides and sucrose content in soybean will provide better understanding of the formation of those seed constituents, and may help breeders accelerate the development of specialty varieties destined for food-quality purposes. Several studies have already utilized molecular markers to identify the genomic regions particularly associated with quantitative trait locus (QTL), including the contents of sucrose (Maughan et al., 2000), seed protein and oil (Chung et al., 2003; Csanadi et al., 2001), and seed isoflavone (Meksem et al., 2001). Marker-assisted selection (MAS) is a tool that can substantially increase the efficiency of obtaining appropriate genotypes. Advances in molecular genetics have already made possible the genetic dissection and characterization of many quantitatively inherited seed-quality traits.

* Corresponding author; fax +82-55-353-3050
e-mail hkkkim@freechal.com

Maughan et al. (2000) have recently reported on sucrose content in *G. max* × *G. soja*, but the QTLs they identified could not be directly used in breeding and interpretation because of the different background crosses used. Therefore, the primary objective of our current research was to improve the breeding efficiency of oligosaccharide and sucrose contents in *G. max* by identifying the simple sequence repeat (SSR) markers associated with QTLs as well as through MAS.

MATERIALS AND METHODS

Plant Materials and Field Evaluation

Two well-characterized soybean cultivars, 'Keunolkong' and 'Iksan10', were used to study the QTLs associated with various plant components, e.g., growth habit and tendency to lodge. 'Keunolkong', a pure line derived from a locally selected Korean variety, is susceptible to early maturation, and has short stems and large seeds. In contrast, 'Iksan10' is the typical cultivar released from systematic breeding programs through the deliberate crossing of 'KW552' × 'Pangsakong'. This cultivar shows late maturity, long stems, and small seeds.

We used genetic materials from the F_2-derived recombinant inbreed line (RIL) population that were developed by single seed descent (SSD) from reciprocal crosses of 'Keunolkong' × 'Iksan10'. This population, designated K/I, consisted of 115 F_{10} lines. The F_{10} seeds of each line were planted in a randomized complete block design, with two replications, at the Yeongnam Agricultural Research Institute, NICS, RDA, Milyang, Korea, on 12 June 2001. Each entry was planted in a 1.5-m-long paired-row plot with two seeds per hill. Spacing was 60 cm between rows and 10 cm between plants.

Analysis of Free Sugars

Seeds were crushed to a meal consistency, and 1.0 g of this soybean powder was extracted with 70% EtOH (10 ml) at room temperature for 24 h. The extract was then processed through Whatman No. 2 paper and a 0.45 mm membrane filter. To eliminate any non-polar compounds, the filtered extraction was passed through a Sep-Pak C18 plus cartridge (Waters, USA), and then diluted 1:1 with distilled water. Afterward, to analyze oligosaccharide and sucrose contents, 20 μl of the diluted extract was injected into an HPLC column (300 mm × 6.5 mm, Sugar Pak 1; Waters/Hewlitt Packard 1100; Germany) that was equipped with a Shodex RI Detector. HPLC conditions included a temperature of 90°C and 0.0001 M EDTAPCaNa$_2$ used as solvent, with a flow rate of 0.5 ml min^{-1}.

SSLP Analysis

Genomic DNA was isolated from healthy leaves following the procedure described by Keim et al. (1988). DNA quality and intactness were checked by electrophoresis through a 1% agarose gel. The DNA solution was then diluted to a working concentration with TE buffer (pH 8.0) and stored at −20°C. A total of 199 soybean SSR markers (Cregan et al., 1999) were used with primer pairs to screen polymorphisms between parental genotypes. PCR was performed in a total volume of 10 μl containing 25 ng of template DNA, 0.15 μM of each primer, 200 μM of each dNTP, 2 mM $MgCl_2$, 0.1% Triton X-100, 1× reaction buffer [10 mM Tris-HCl (pH 8.5) and 100 mM KCl] and 0.5 U of *Taq* DNA polymerase (BioBasic *Taq* Polymerase, Applied BioBasic, Canada). Template DNA was initially denatured at 94°C for 2 min, followed by 40 cycles for amplification under the following conditions: denaturation at 94°C for 25 s, annealing at 47°C for 25 s, and extension at 68°C for 60 s, all on a 96-well GeneAmp PCR system 9700 (Applied Biosystems, USA). The segregation patterns for each SSR marker were determined by electrophoresis on a 4% polyacrylamide gel. Afterward, the gel was stained with a silver sequencing kit (Promega, USA) and scored for map construction. Pigmentation colors for the flower and hilum were also scored as morphological markers.

Map Construction and Statistical Analysis

Traits means, correlation, and analysis of variance were determined by SAS (Statistical Analysis Systems, USA). Narrow-sense heritability (h^2) was calculated on a per-plot basis, using an estimate of the variance component (Frey and Horner, 1957). Based on the segregation data subsets for SSRs and morphological markers, we constructed a linkage map with MapManager QT version 2.8 software (Manly and Olson, 1999). Recombination fractions were converted to map distances by applying the Haldane map function (Haldane, 1919). Where possible, linkage groups were named according to their designations from a consensus USDA map (Cregan et al., 1999).

The association between marker and QTL was

tested according to the interval mapping methods of Whittaker et al. (1996), using MapManager QT and single-factor ANOVA. For each SSR and morphological marker, the class means for growth habit and lodging were compared for significance ($P < 0.05$) using an F-test from the TypeIII mean squares, as obtained from the GLM Procedure SAS program. In addition, a two-way ANOVA was used to detect significant ($P < 0.05$) interactions (i.e., epistasis) among all pairs of significant markers.

If SF-ANOVA identified two or more linked markers associated with the oligosaccharide and sucrose contents, a multiple regression analysis was conducted by including all the significant markers on that linkage group in the model (i.e., SLG-Regr). Forward and step-wise selection procedures were applied in the regression analysis. Any significant ($P < 0.05$) markers that were retained in the SLG-Regr analysis were assumed to identify unique QTLs in that linkage group. All significant markers from the SLG-Regr analysis were then combined in a multiple-linkage group regression (MLG-Regr) at $P < 0.05$ to determine the combination of independent markers that explained the greatest amount of phenotypic variation in a given trait. This probability level was selected to enhance our ability to detect QTLs associated with growth traits and lodging. Finally, the coefficient of determination (R^2) obtained from MLG-Regr was used to provide an estimate of the percentage phenotypic variation explained by the markers.

RESULTS

Phenotypic Evaluation of Oligosaccharide and Sucrose Contents

Oligosaccharide and sucrose contents were normally distributed in the F_2-derived F_{10} RIL populations (Fig. 1), with the former ranging from 67.9 to 102.4 g kg^{-1} (mean of 83.8 g kg^{-1}) and the latter measuring between 35.2 and 69.0 g kg^{-1} (mean of 50.9 g kg^{-1}). Average values for oligosaccharides and sucrose were higher from Keunolkong (87.5 g kg^{-1} and 55.7 g kg^{-1}, respectively) than from 'Iksan10' (76.7 g kg^{-1} and 38.0 g kg^{-1}, respectively). Overall, the contents of oligosaccharide and sucrose were highly correlated (0.966**), a trend that was also reported previously (Hartwig et al., 1997; Wilcox and Shibles, 2001). Narrow-sense heritabilities of oligosaccharide and sucrose contents were 80.4% and 78.3%, respectively (Table 1). This trend has also been reported by Openshaw and Hadley (1981) and Maughan et al. (2000).

Distribution of QTLs Associated with Oligosaccharide and Sucrose Contents in Soybean Seed

The linkage map for our K/I population consisted of 20 linkage groups, which spanned 1590 cM and were defined by 99 SSRs and two morphological

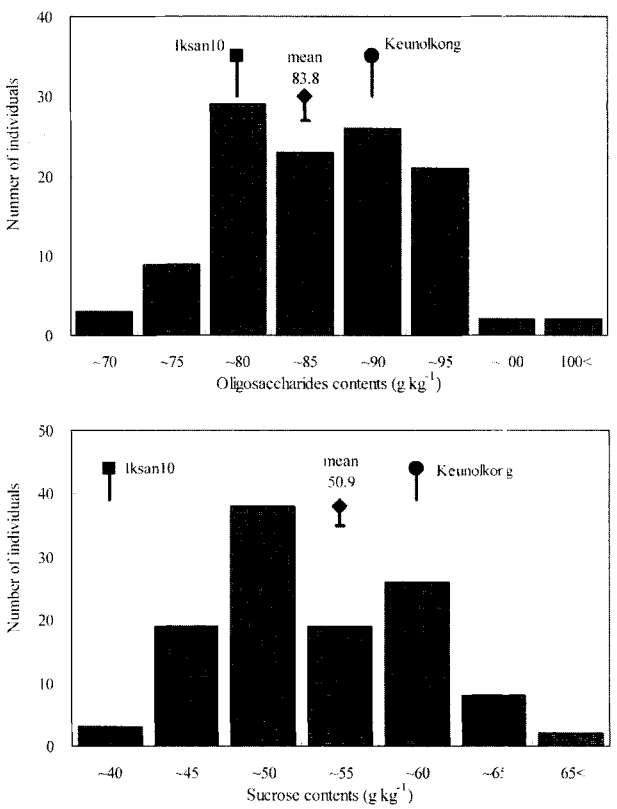

Figure 1. Frequency distribution of oligosaccharide and sucrose contents in 115 RILs of 'Keunolkong' × 'Iksan10'.

Table 1. Estimates of narrow-sense heritabilities (percentage) for oligosaccharide and sucrose contents in F_2-derived F_{10} RIL population of 'Keunolkong' × 'Iksan10'.

Trait	df	MS_1	MS_2	Pr>F	R^2	h^2
Sucrose	114	0.117	0.960	<0.0001	0.892	78.33
Oligosaccharides	114	0.102	0.939	<0.0001	0.902	80.40

$h^2 = (MS_2\text{-}MS_1) / (MS_1+(r\text{-}1)MS_2)$, MS_1 = line mean square, MS_2 = error mean square, r = replication.

markers. On average, this map revealed a marker density of 1.0 per 15.9 cM. Here, we present only those LGs in which tentative QTLs were positioned (Fig. 2, 3, and 4).

Our SF-ANOVA analysis identified eight markers as potentially associated with oligosaccharide content (Table 2). The phenotypic variation explained by individual markers ranged from 4.92% to 14.21%. Two QTLs were identified in LG D1b+W and L, based on MLG-Regr analysis, accounting for 20.45% of the total phenotypic variation (Table 2). The QTLs located at the marker intervals of satt278 and satt313 in LG L appeared to be the major controls of oligosaccharide content (Fig. 2). Another QTL was located in LG D1b+W, but its contribution to oligosaccharide content was relatively low (Fig. 3). Finally, tentative loci on the marker intervals of satt216 and satt269 in LG D1b+W could not be quantified by MLG-Regr analysis, and further study is needed to confirm whether they are real or simply caused by an environmental factor.

In determining sucrose content, we noted 11 poten-

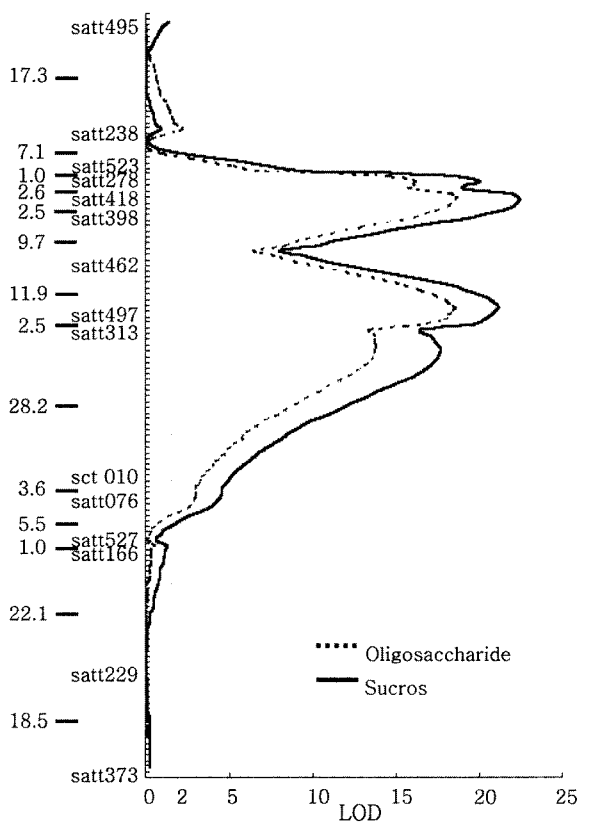

Figure 2. Distribution of QTLs associated with oligosaccharide and sucrose contents in LG L from 115 RILs of 'Keunolkong' × 'Iksan10'. Open vertical bar indicates LOD score of 2.0.

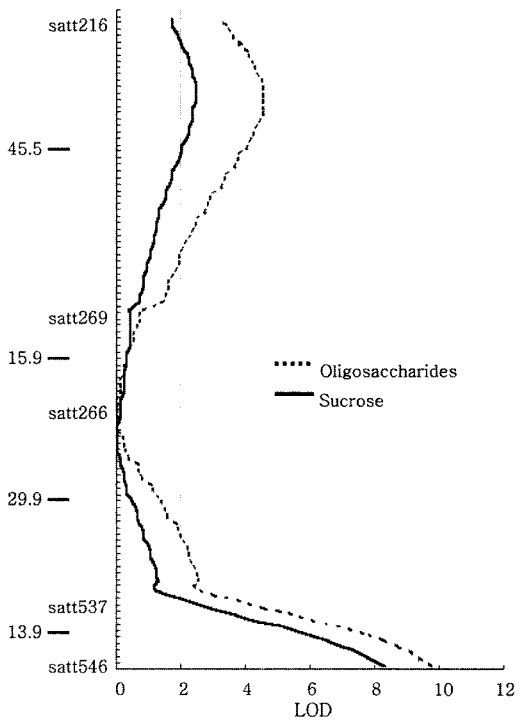

Figure 3. Distribution of QTLs associated with oligosaccharide and sucrose contents in LG D1b+W from 115 RILs of 'Keunolkong' × 'Iksan10'. Open vertical bar indicates LOD score of 2.0.

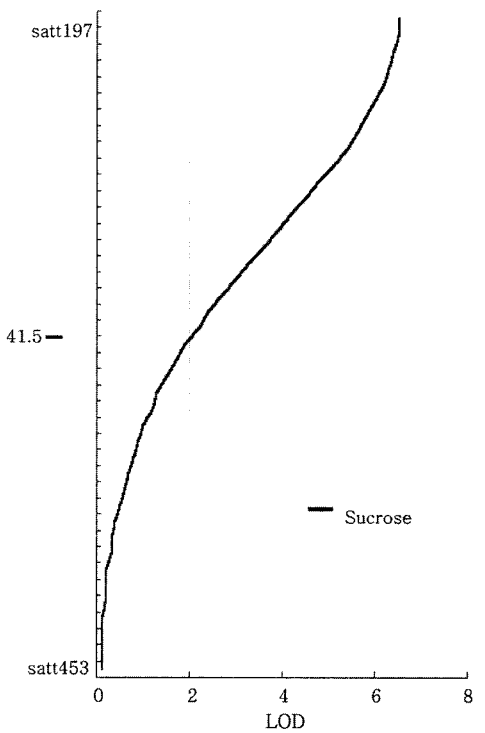

Figure 4. Distribution of QTLs associated with sucrose contents in LG B1 from 115 RILs of 'Keunolkong' × 'Iksan10'. Open vertical bar indicates LOD score of 2.0.

Table 2. Marker distributions and QTLs associated with oligosaccharide and sucrose contents in 115 RILs of 'Keunolkong'×'Iksan10'.

Trait	Marker	LG	SF-ANOVA[a]		Allelic means		MLG-Regr[b]	
			P	R^2 (%)	K/K[c]	I/I[c]	P	R^2 (%)
Oligosaccharide content	satt546	D1b+W	0.0018	8.82	81.6	85.8	0.0071	6.48
	satt523	L	0.0022	8.85	84.0	79.2	-	-
	satt278	L	<0.0001	13.78	84.7	78.2	0.0002	13.97
	satt418	L	<0.0001	14.14	85.1	78.4	-	-
	satt398	L	0.0002	12.82	85.0	78.2	-	-
	satt462	L	0.0211	4.92	84.7	81.3	-	-
	satt497	L	<0.0001	14.21	85.1	77.7	-	-
	satt313	L	0.0004	10.97	84.8	79.1	-	-
Sucrose content	satt197	B1	0.0202	4.94	49.2	52.3	0.0352	3.61
	satt546	D1b+W	0.0033	7.84	48.8	52.7	0.0074	6.43
	satt523	L	0.0006	11.04	51.3	45.8	0.0280	4.10
	satt278	L	<0.0001	17.51	52.1	44.5	<0.0001	17.29
	satt418	L	<0.0001	16.43	52.3	45.0	-	-
	satt398	L	<0.0001	15.31	52.3	44.8	-	-
	satt462	L	0.0087	6.32	52.0	48.2	-	-
	satt497	L	<0.0001	15.86	52.2	44.4	-	-
	satt313	L	<0.0001	13.42	52.1	45.7	-	-
	sct010	L	0.0235	4.62	52.4	49.4	-	-
	satt076	L	0.0493	3.70	51.8	49.2	-	-

[a]SF- ANOVA, single-factor analysis of variance; [b]MLG-Regr, multiple regression with all significant markers from the SLG-Regr model; [c]K/K, 'Keunolkong'; I/I, 'Iksan10'.

Table 3. Allele interactions associated with oligosaccharide and sucrose contents in 'Keunolkong'×'Iksan10'.

Trait	SSR locus	Allele	Allele		*P*	R^2 (%)
			Keunolkong	Iksan10		
Oligosaccharides	satt462		satt523		0.0034	16.54
		Keunolkong	83.7	87.3		
		Iksan10	85.1	77.3		
Sucrose	satt462		satt523		0.0015	20.00
		Keunolkong	51.1	54.7		
		Iksan10	52.6	43.9		

tial markers that individually accounted for 3.70% to 17.51% of the phenotypic variation. MLG-Regr analysis identified four QTLs in LG L -- D1b+W and B1 (Table 2; Fig. 2, 3, and 4). One major QTL, located on marker interval of satt278 and satt523 on LG L, had a relatively high phenotypic variation of 21.39% while the other two had only minor effects. Those sucrose-associated QTLs in LG D1b+W showed the same trends as those controlling oligosaccharide content.

For both oligosaccharides and sucrose, we observed that different allele effects depended on the chromosomal background. For example, the 'Keunolkong'-derived marker allele in LG L increased the contents of both compounds. The same effect was seen with the 'Iksan10'-derived marker alleles in LG D1B+W and B1. In contrast, the allele effects derived from 'Iksan10' in LG D1b+W and B1 were weaker that that of 'Keunolkong' for increasing either oligosaccharide or sucrose content.

To study the allele-interaction effects at a given QTL, we examined each marker set that manifested trait significance (Table 3). At satt523/satt462 combination, marker alleles that originated from 'Iksan10' were associated with the lowest amount of oligosaccharides (77.3 g kg^{-1}) and sucrose (43.9 g kg^{-1}); other allele combinations resulted in much higher contents.

DISCUSSION

Sucrose, followed by stachyose and raffinose, is the

most abundant carbohydrate in mature soybean seeds. The ability to extract it from defatted soybean meal has increased its economic value to the food and feed industries (Kim et al., 2003).

Oligosaccharide and sucrose contents in seed are, in general, negatively correlated with protein content, but positively correlated with oil content (Hymowitz et al., 1972; Wilcox and Shibles, 2001). However, unlike in other studies, our data demonstrated a negative correlation with oil content; i.e., r= -0.554** for oligosaccharides and r= -0.473** for sucrose. This may be explained in two ways. First, we used a different genetic background than what was derived from the *Glycine soja* research. Second, different characteristics were targeted in those previous breeding programs, i.e., the production of soybean paste and oil.

Our MLG-Regr analysis identified two and three linkage groups, respectively, that were related to oligosaccharide content (LG D1b+W and LG L) and sucrose content (LG B1 and D1b+W; and LG L). LG L appeared to be the major QTL control of these two compounds. However, those loci extended across 61.0 cM of the chromosome. Moreover, only the marker interval of satt278 could explain the phenotypic variations in content.

Interestingly, we identified one other QTL related to sucrose, in on LG B1 (Fig. 4). We considered this very unusual because sucrose content is known to be closely correlated with oligosaccharide content, and both those contents are controlled in the same direction under the same genes. Nevertheless, we detected no tentative QTLs in LG B1 that were related to oligosaccharides. Therefore, further study is needed to confirm whether those QTLs are real or that, instead, environmental factors affected those results. In previous research, the interspecific cross of *G. max* and *G. soja* was used to investigate the relationship among oil, protein, and sucrose. For example, Brummer et al. (1997) identified one QTL in LG I that was associated with protein and sucrose contents. Maughan et al. (2000) also detected seven linkage groups (A1, A2, E, F, I, L, and M) that were associated with sucrose. Here, however, we located the QTLs related to oligosaccharide and sucrose contents on different chromosomes, an aberration that may have arisen because *G. soja* had been used previously. A well-fitting pleiotypic theory on the occurrence of QTLs for oil, protein, and sucrose contents has been proposed by Maughan et al. (2000), and may be used to explain the key metabolic importance of seed constituents from soybean.

In this study, we confirmed the positioning of the QTLs, and achieved relatively high LOD scores for both oligosaccharide and sucrose contents. However, the phenotypic variations explained by the QTLs were fairly low, and showed narrow spectrums compared with the wide-spreading QTL peak. These results lead us to suspect that environmental factors affected the trends reported previously (Brummer et al., 1997; Wilcox and Shibles, 2001). Therefore, high quality soybean seeds could be identified through a concerted effort of conformable QTL studies that are consistent with environmental effects and genotypic backgrounds.

Received June 26, 2004; accepted January 7, 2005.

LITERATURE CITED

Abdin OA, Zhou X, Coulman BE, Cloutier D, Faris MA, Smith DL (1998) Effect of sucrose supplementation by stem injection on the development of soybean plants. J Exp Bot 49: 2013-2018

Aldape MJ, Elmer AM, Chao WS, Grimes HD (2003) Identification and characterization of a sucrose transporter isolated from the developing cotyledons of soybean. Arch Biochem Biophys 409: 243-250

Borejszo Z, Khan K (1992) Reduction of flatulence-causing sugars by high temperature extrusion of pinto bean high starch fraction. J Food Sci 57: 771-772

Brummer EC, Graef GL, Orf J, Wilcox JR, Shoemaker RC (1997) Mapping QTL for seed protein and oil content in eight soybean populations. Crop Sci 37: 370-378

Chung J, Babka HL, Graef GL, Staswick PE, Lee DJ, Cregan PB, Shoemaker RC, Specht JE (2003) The seed protein, oil and yield QTL on soybean linkage group I. Crop Sci 43: 1053-1067

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

Csanadi G, Vollmann J, Stift G, Lelley T (2001) Seed quality QTLs identified in a molecular map of early maturing soybean. Theor Appl Genet 103: 912-919

Frey KJ, Horner T (1957) Heritability in standard units. Agron J 49: 59-62

Haldane JBS (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. J Genet 8: 299-309

Hartwig EE, Kuo TM, Kenty MM (1997) Seed protein and its relationship to soluble sugars in soybean. Crop Sci 37: 770-773

Hoover DG (1993) Bifidobacteria activity and potential benefit. Food Tech 6: 120-124

Hymowitz T, Collins FI, Panczner J, Walker WM (1972) Relationship between the content of oil, protein and sugar in soybean seed. Agron J 66: 613-616

Kawamura S (1954) Studies on soybean carbohydrates IV. Determination of oligosaccharides in soybeans. Nippon Nogei Kogaka Kaishi 28: 851-852

Keim P, Olson TC, Shoemaker RC (1988) A rapid protocol for isolating soybean DNA. Soybean Genet News 15: 150-154

Kim S, Kim W, Hwang IK (2003) Optimization of the extraction and purification of oligosaccharides from defatted soybean meal. Int J Food Sci Tech 38: 337-342

Koga Y, Shibuta T, O'Brien R (1993) Soybean oligosaccharides, *In* T Nakakuki, ed, Oligosaccharides Production, Properties and Applications. Gordon and Beach Science, New York

Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QT. Mamm. Genome 10: 327-334

Maughan PJ, Saghai Maroof MA, Buss GR (2000) Identification of quantitative trait loci controlling sucrose content in soybean (*Glycine max*). Mol Breed 6: 105-111

Meksem K, Njiti VN, Banz WJ, Iqbal MJ, Kassem MM, Hyten DL, Yuang J, Winters TA, Lightfoot DA (2001) Genomic regions that underlie soybean seed isoflavone content. J Biomed Biotech 1: 38-44

Omosaiye O, Cheryan M, Metthews ME (1978) Removal of oligosaccharide from soybean water extracts by ultrafiltration. J Food Sci 43: 354-359

Openshaw SJ, Hadley HH (1981) Selection to modify sugar content of soybean seeds. Crop Sci 21: 805-808

Paterson AH, Tanksley SD (1991) DNA markers in plant improvement. Adv Agron 46: 39-89

Patrick R, Offler CE (1995) Post-sieve element transport of sucrose in developing seeds. Aust J Plant Physiol 22: 681-702

Weber H, Heim U, Golomblek S, Borisjuk L, Wobus U (1998) Assimilate uptake and the regulation of seed development. Seed Sci Res 8: 331-345

Whittaker JC, Thompson R, Visscher DM (1996) On the mapping of QTL by regression of phenotype on marker type. Heredity 77: 23-32

Wilcox JR, Shibles RM (2001) Interrelationships among seed quality attributes in soybean. Crop Sci 41: 11-14