Theor Appl Genet (1996) 93: 1011–1016

M. A. R. Mian · M. A. Bailey · J. P. Tamulonis
E. R. Shipe · T. E. Carter Jr. · W. A. Parrott
D. A. Ashley · R. S. Hussey · H. R. Boerma

Molecular markers associated with seed weight in two soybean populations

Received: 28 April 1996 / Accepted: 24 May 1996

Abstract Seed weight (SW) is a component of soybean, *Glycine max* (L.) Merr., seed yield, as well as an important trait for food-type soybeans. Two soybean populations, 120 F_4-derived lines of 'Young' × PI416937 (Pop1) and 111 F_2-derived lines of PI97100 × 'Coker 237' (Pop2), were mapped with RFLP makers to identify quantitative trait loci (QTLs) conditioning SW across environments and populations. The genetic map of Pop1 consisted of 155 loci covering 973 cM, whereas Pop2 involved 153 loci and covered 1600 cM of map distance. For Pop1, the phenotypic data were collected from Plains, GA., Windblow, N.C., and Plymouth, N.C., in 1994. For Pop2, data were collected from Athens, GA., in 1994 and 1995, and Blackville, S.C., in 1995. Based on single-factor analysis of variance (ANOVA), seven and nine independent loci were associated with SW in Pop1 and Pop2, respectively. Together the loci explained 73% of the variability in SW in Pop1 and 74% in Pop2. Transgressive segregation occurred among the progeny in both populations. The marker loci associated with SW were highly consistent across environments and years. Two QTLs on linkage group (LG) F and K were located at similar genomic regions in both populations. The high consistency of QTLs across environments indicates that effective marker-assisted selection is feasible for soybean SW.

Key words Soybean · *Glycine max* · Seed weight · RFLP · QTL · Markers

Communicated by J. Mac Key

M. A. R. Mian · M. A. Bailey[1] · J. P. Tamulonis · W. A. Parrott · D. A. Ashley · H. R. Boerma (✉)
Department of Crop and Soil Sciences, University of Georgia, Athens, GA 30602-7272, USA

E. R. Shipe
Department of Agronomy, Clemson University, Clemson, SC 29634-0359, USA

T. E. Carter Jr.
USDA-ARS, Dep. of Crop Science, North Carolina State University, Raleigh, NC 27695-7631, USA

R. S. Hussey
Department of Plant Pathology, University of Georgia, Athens, GA 30602-7274, USA

Present address:
[1] Pioneer Hi-Bred International, Inc., 7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131, USA

Introduction

Seed weight (SW) is one of the yield components of soybean and, in general, is positively correlated with seed yield (Burris et al. 1973; Smith and Camper 1975). SW is especially critical for various soybean food-types. For example, soybean seed used for sprouts should possess small SW, whereas soybean seed used for cooking with rice should have large SW. In South Korea, the average weight of a black soybean for cooking with rice is 300–350 mg seed^{-1} whereas that of a soybean for sprouts is about 120 mg seed^{-1} (Hong et al. 1994). The demand for food-type soybean is increasing in the international market. Thus, the need for development of productive soybean cultivars with either large or small seed is increasing in importance.

Soybean SW is a highly heritable trait with heritability ranging from 44 to 94% (Brim 1973). Recently, Paterson et al. (1995) detected seven, eight, and eight QTLs for the SW of *Sorghum bicolor* (L.), *Oryza sativa* (L.), and *Zea mays* (L.), respectively. These QTLs explained from 52 to 78% of the variation in SW. Christian et al. (1992) identified two QTLs for SW in *Vigna unguiculata* (L.) and four QTLs for SW in *Vigna radiata* (L.) explaining 53 and 50% of total variation in the trait, respectively. Mansur et al. (1993), however, detected only one unlinked marker for soybean SW explaining 13% of the variation for the trait.

Many earlier studies examining molecular marker associations with QTLs have used phenotypic data from a single environment. Recent studies using phenotypic data from multiple environments have shown inconsistencies among environments in the detection of QTLs (Paterson

et al. 1991; Dudley 1993; Tanksley 1993). Stuber et al. (1992), however, reported a consistency of QTLs across environments for seven quantitative traits in maize. Ajmone et al. (1995) also reported a consistency of QTLs across environments for three quantitative traits in maize. Lee et al. (1996 a), in a study of QTLs associated with soybean plant height, lodging, and maturity over four locations, found that the consistency of QTLs was trait specific.

Paterson et al. (1991) suggested that the studies conducted in a single environment are likely to underestimate the number of QTLs which can influence a trait. It is also possible to have environmentally sensitive QTLs, meaning that the expression of these QTLs will only occur under certain environments (e.g., a particular range of temperatures). In the latter case, one would only be able to identify the QTLs at a location where these environmental conditions are satisfied. Lee et al. (1996 a) suggested that the phenotypic data for a polygenic trait should be collected over a range of locations from within the base population of environments to identify putative QTLs.

Until now, very few studies have been reported on associations of molecular markers with soybean SW, and these studies did not use multiple environments or populations for phenotypic data. Even though Mansur et al. (1993) collected the phenotypic data from two environments, they reported the QTLs based on the combined data and did not discuss the consistency of QTLs across environments. Thus, the objectives of the present study were to: (1) detect QTLs conditioning SW in soybean, and (2) investigate the consistency of the QTLs across populations and environments.

Materials and methods

Two soybean populations, Young × PI416937 (Pop1) and PI97100 × Coker 237 (Pop2), were used for this study. Pop1 was derived from a cross between the high-yielding cultivar Young (Burton et al. 1987) and drought-tolerant PI416937 (Sloane et al. 1990). The 120 F_4-derived lines of this population were obtained by single-seed-descent, with each line originating from a different F_2 plant. Pop2 consisted of 111 F_2-derived lines, with each line originating from a different F_2 plant. This population was segregating for growth habit. PI97100 is indeterminate in its growth habit, whereas Coker 237 has a determinate growth habit.

The procedures for RFLP mapping of the two populations have been described previously (Lee et al. 1996 a, b; Mian et al. 1996). Linkage maps were constructed with marker data using the Kosambi map function of GMENDEL (Holloway and Knapp 1993) for Pop1 and Mapmaker (Lander et al. 1987) for Pop2. For grouping markers into linkage groups, a minimum LOD of 3.0 (rmax of 0.38 for GMENDEL) and a maximum distance of 50 cM were employed.

In 1994, the parents and the F_4-derived lines of Pop1 were grown at Plains (Univ. of Georgia Southwest Branch Exp. Stn.), GA., and Windblow (North Carolina State Univ. Sandhills Res. Stn.) and Plymouth (North Carolina State Univ. Plymouth Res. Stn.), N.C. The parents and the F_2-derived lines of Pop2 were grown at Athens (Univ. of Georgia Plant Sciences Farm), GA., in 1994; and at Athens as well as at Blackville (Clemson Univ. Edisto Res. and Educ. Ctr.), S.C. in 1995.

To reduce experimental error due to soil heterogeneity within an individual experiment, the lines in each population were divided into three groups. For Pop1, the lines in each of the groups were placed in separate tests along with three entries of Young and one entry of PI416937 as reference genotypes. Similarly, PI97100, Coker 237, and Stonewall were used as reference genotypes in each test of Pop2.

In both populations, each test was planted in an incomplete block design. Each test of Pop1 was replicated twice, except for the tests at Windblow, N.C., where three replications were used. For Pop2, two tests had four replications each, while the third test had only three replications due to insufficient seed availability. The plot size at different locations varied from 5.56 to 14.05 m^2 (Table 1). Additional details on plant culture for Pop1 and Pop2 can be found in Lee et al. (1996a, b). The SW was determined in mg $seed^{-1}$ based on a 100-seed-sample per plot.

The association between markers and QTLs was evaluated by using two different procedures. One method was interval mapping (Lander and Botstein 1989) with Mapmaker QTL software (Lincoln et al. 1992) for Pop2. The QTL analyses were done on the standardized genotypic means of SW data within each environment, as well as across environments. For standardization across the three tests, the plot values within each test were divided by the mean of the reference entries within that test. A LOD score of 2.4 was chosen as the minimum to declare the presence of a QTL in a given genomic region. The LOD score peak was used to estimate the most likely QTL position on the RFLP linkage map. The percent of variation explained by an individual QTL, and the additive (a) and dominance (d) effects, were estimated at the maximum-likelihood QTL position. The average degree of dominance for each QTL was calculated as the ratio of d/a.

The other method was single-factor ANOVA for Pop1 and 2. Each marker locus was evaluated for linkage to a QTL affecting SW based on the standardized data. Significance among RFLP genotypic class means was determined by using an F-test from the Type-III mean

Table 1 Description of plots and soil classifications at experimental sites for two soybean populations

Population	Location	Rows/ plot	Row spacing (m)	Row length (m)	Soil classification
Young × PI416937	Plains, G.A.	4	0.76	3.66	Clayey, kaolinitic, thermic. Typic Paleudult
	Plymouth, N.C.	3	0.96	4.88	Sandy, skeletal, mixed thermic. Typic Umbracudult
	Windblow, N.C.	3	0.96	3.05	Sandy, salicious, thermic. Arenic Paleudult
PI97100 × Coker 237	Athens, G.A.	2	0.76		Clayey, kaolinitic, thermic. Typic Hapludults
	Blackville S.C.	2	0.96		Loamy, siliceous, thermic. Arenic Plinthic Kandiudult

squares obtained from the GLM procedure (SAS 1988). For Pop1, only homozygous RFLP classes at each marker locus were compared for the determination of significant difference. Due to the limited number of heterozygous lines available for a given marker in F_4-derived lines (the expected number is 15=12.5% of 120 lines), the heterozygous class was excluded from the analysis. The significant markers were determined by first testing all the markers for significant association with SW at $P<0.01$. Then the relaxed probability level of $P<0.05$ was used to detect the consistency of the previously identified markers across environments. Thus, each of the putative independent marker loci presented in this paper was significant at $P<0.01$ in at least one of the environments for each population. Putative independence was defined as a significant marker which was at least 50 cM from another marker significantly associated with the trait, and which acted in an additive manner to explain variation. Two-factor analysis of variance was used to detect significant ($P<0.01$) epistatic interactions between all possible pairs of independent marker loci.

Results and discussion

Genetic map

A total of 155 RFLP markers for Pop1 and 153 markers for Pop2 was used to construct the genetic linkage maps. In Pop1, 140 markers were genetically linked resulting in 33 linkage groups with a coverage of 970 cM. The map of Pop2 had 23 linkage groups with 143 linked markers and covered 1600 cM. Detailed descriptions of the genetic maps were provided in Lee et al. (1996 a, b) and Mian et al. (1996).

The abbreviated soybean linkage groups presented in the present paper (Fig. 1) were made possible due to recent research by Randy Shoemaker's laboratory at Iowa State University (Shoemaker et al. 1997). Using the 'Joinmap software (Stam 1993), they created an integrated RFLP map from nine different soybean populations, including both Pop1 and Pop2 of this study. Most of the markers from Pop1 and Pop2 were placed on this newly integrated map.

Seed-weight QTLs

Combined across environments, Young and PI416937 did not differ in SW, while Coker 237 had a 15% larger SW than PI97100 (Table 2). The progeny of both populations displayed transgressive segregation for both larger and smaller SW than did the parents. In Pop1, in which the parents did not differ in SW, the progeny differed by as much as 93 mg seed^{-1}. In Pop2, the parents differed by 19 mg seed^{-1}, and the progeny differed by 44 mg seed^{-1} (Table 2).

In Pop1, a total of seven putatively independent marker loci was associated with SW in the combined analysis (Fig. 1), of which six were identified in all three environments (Table 3). Thus, the level of consistency of the QTLs for SW across environments was very high. Marker locus

Table 2 Mean seed weight for the parents and the extreme progeny lines combined over environments in two soybean populations

Genotype	Mg seed^{-1}	
	Young×PI416937	PI97100×Coker 237
Female parent	160	128
Male parent	174	147
High progeny	227	160
Low progeny	134	116
$LSD_{0.05}$	15	5

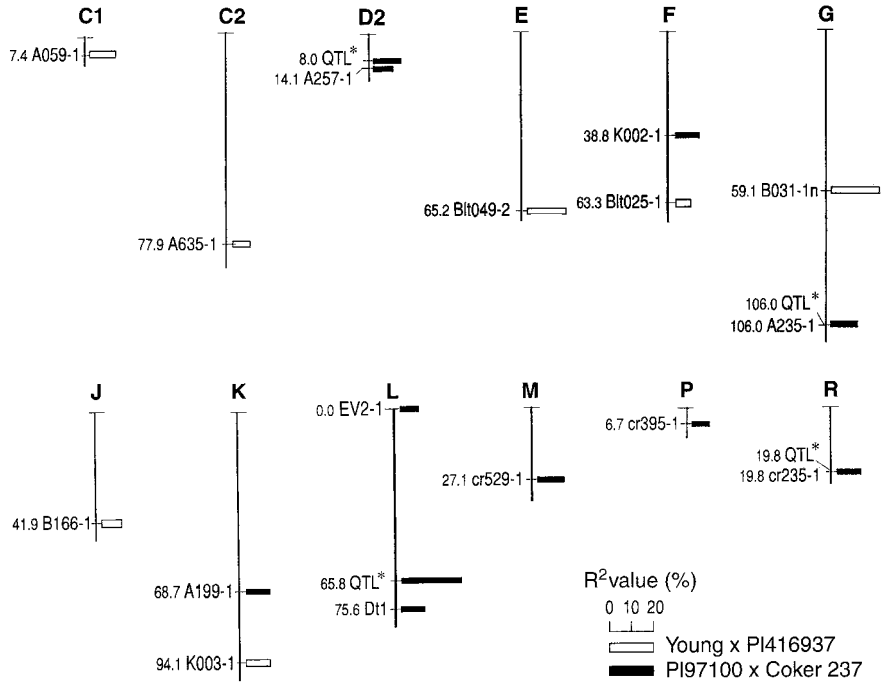

Fig. 1 Soybean linkage groups (LG) with the independent marker loci conditioning seed weight (SW) in two soybean populations, Young×PI416937 (Pop1) and PI97100×Coker 237 (Pop2). The map distance (in cM) shown on the left of each marker is according to the newly integrated USDA soybean linkage map (Shoemaker et al. 1997), and indicates the distance of a marker from the top of the LG. The length of horizontal bars indicates R^2 values for the loci associated with SW in Pop1 (*open bar*) and Pop2 (*solid bar*). QTL* shows the most likely QTL position determined by the Mapmaker/QTL analysis

Table 3 Putative independent marker loci associated with variation in seed weight of F_4-derived lines from the cross of Young × PI416937 based on single-factor analysis of variance

	Linkage group	Combined				Environment					
		P	R^2 (%)	Allelic mean (mg seed^{-1})		Plains		Plymouth		Windblow	
				Young	PI416937		R^2 (%)		R^2 (%)	*P*	R^2 (%)
A059-1	C1	0.0001	10	172	185	0.006	9	0.0001	14	0.004	10
A635-1	C2	0.0100	6	181	172	0.009	7	0.0400	4	0.009	6
Blt49-2n	E	0.0001	14	170	184	0.01	6	0.0001	16	0.0001	14
Blt025-1	F	0.0200	5	177	170	0.04	4	0.0040	8	–	–
B031-1n		0.0001	22	167	187	0.0001	18	0.0001	24	0.0001	20
B166-1	J	0.006	8	180	170	0.007	7	0.0200	6	0.009	7
K003-1	K	0.004	8	170	179	0.021	6	0.0100	6	0.002	9

Table 4
Cokker 237 based on single-factor analysis of variance

RFLP locus	Linkage group	Combined					Environment					
		P	R^2 (%)	Allelic mean (mg seed^{-1})[a]			Athens, 94		Athens, 95		Blackville, 95	
						B/B	*P*	R^2 (%)	*P*	R^2 (%)	*P*	R^2 (%)
A257-1	D2	0.0087	8	133	135	141	0.0062	9	0.0289	6	0.0164	8
K002-1	F	0.0168	7	140	134	136	0.0388	6	0.0032	10	–	–
A235-1	G	0.0040	10	132	139	136	0.0125	8	0.0277	7	0.0029	11
A199-1	K	0.0501	5	129	128	134	0.0098	8	–	–	0.0140	10
Dt1	L	0.0028	10	140	137	131	0.0143	8	0.0321	6	0.0013	12
Ev2-1	L	0.0256	7	140	133	134	–	–	–	–	0.0052	10
Cr529-1	M	0.0150	11	140	135	130	0.0329	9	0.0150	11	0.0362	9
Cr395-1	P	0.0300	6	125	129	133	0.0052	7	0.0465	4	0.0099	6
Cr235-1	R	0.0159	10	136	137	144	0.0430	8	0.0150	8	0.0470	8

[a] A/A: homozygous PI97100, A/B: heterozygous, B/B: homozygous Coker 237

Blt049-2n on LG E was mapped on LG Unk3 of the previously published map of Young×PI416937 population (Mian et al. 1996). However, the LG Unk3 of the Young×PI416937 map is now placed on LG E of the newly integrated soybean map (Shoemaker et al. 1997). No significant epistatic interactions were detected among the independent marker loci in the two-way analysis of variance.

Individually, the seven marker loci explained between 5 and 22% of the variation in SW. If combined, the seven marker loci would together explain as much as 73% of the total variation in SW. The heritability of SW in this population was 90% (based on a selection unit of three environments, and two replications/environment). Thus, most of the variation in SW could be explained by the seven marker loci. The PI416937 alleles contributed to larger SW at four of the seven independent loci, while the Young alleles did so at the remaining three loci (Table 3). Thus, both parents potentially can contribute to a larger or smaller SW of the progeny, and this provides the genetic basis for the transgressive segregation observed in this population.

In Pop2, nine independent loci were identified by single-factor ANOVA as being associated with SW in the combined analysis (Fig. 1); six of these were consistent across all three environments (Table 4). Markers K002-1 (LG F) and A199-1 (LG K) were each detected at two out of the three environments. Marker locus EV2-1 (LG L) was significant only at Blackville ($P<0.005$) and in the combined analysis ($P<0.02$). Thus, the consistency level of the QTLs across environments for Pop2 was also high. At Athens, the only location where data were collected in 2 different years, the consistency level of significant QTLs across years was very high. Of the eight marker loci detected in 1994, seven were also detected in 1995. A199-1 (LG K) was the only significant marker locus from 1994 that was not detected in 1995.

No epistatic interactions were found among the nine marker loci in the two-way analysis. Individually, these loci explained between 5 and 11% of the variation in SW. The nine loci, if combined, would explain 74% of variation in SW. The heritability of SW in this population was 91% (on the basis of a selection unit of three environments and four replications/environment). Most of the variation in SW could, therefore, be explained by these nine marker loci. At four of the nine marker loci, PI97100 alleles contributed to a larger SW, whereas the Coker 237 alleles did so at the remaining five loci (Table 4).

The interval-mapping method identified four of the nine loci detected by the single-factor ANOVA in the combined

Table 5 Genomic location, genetic effects, and percentage of variability for seed weight loci of F_2-derived soybean lines from the cross of PI97100× Coker 237 based on interval-mapping analysis

Linkage group[a]	Interval	Length (cM)	QTL position (cM)[b]	Genetic effects (mg seed^{-1})[c]			R^2 (%)	LOD[d]
				Additive (a)	Dominant (d)	d/a		
D2	A095-1–A257-1	13	6	4.8	–2.2	–0.5	12	2.4
G	A235-1–L154-2	3	0	2.4	6.4	2.7	10	2.6
L	MO109-1–Dt	60	10	–7.2	6.6	–0.9	26	3.4
R	cr235-1–cr322-1	0	0	6.3	–3.9	–0.6	10	2.6

[a] USDA linkage group (Shoemaker and Specht 1995)
[b] Most likely locus position, corresponding to LOD score peak, which represents the distance from the right marker of interval
[c] Genetic effects were estimated using Mapmaker/QTL. A negative sign indicates that the Coker 237 allele decreases the value of the trait
[d] LOD indicates how much more probable the data are to have arisen assuming the presence of a locus than assuming its absence; LOD theshhold = 2.4

analysis for Pop2 (Table 5 and Fig. 1). The likely position of the QTLs on LG D2 was 6 cM from marker A257-1, and explained 12% of the variation. The likely position of the QTL on LG G was right on the marker A235-1, and explained 10% of the variation in SW. Lander and Botstein (1989) stated that the amount of variation explained by a particular marker in a single-factor ANOVA, compared to interval mapping, is a function of the distance between the QTL and the marker locus. The results from our data support this concept. For example, the QTL identified on LG G was located at marker A235-1 and the amount of variation explained at this locus was the same for both methods of analysis. In contrast, the QTL on LG L was 10 cM away from *Dt1* and the amount of variation explained by interval-mapping analysis was much higher (R^2=26%) than that explained by *Dt1* in the single-factor ANOVA (R^2=10%) (Tables 4 and 5, and Fig. 1).

When a LOD score minimum of 2.4 was chosen, interval mapping underestimated the number of marker loci for SW compared to the number of marker loci detected by the single-factor ANOVA (Tables 4 and 5). In general, interval mapping detected the loci with larger effects and missed the loci with smaller effects, as determined by single-factor ANOVA. The loci with smaller effects could be detected by the interval-mapping method at LOD scores lower than 2.4 (data not shown). These results are in agreement with those of Lee et al. (1996 b) who used the same genetic map for detecting QTLs for plant height, lodging, and maturity. In addition to the LOD score minimum, the underestimation of the QTLs by interval mapping may be due to low saturation of the genetic map used in this study. Advantages and disadvantages of the two methods of analysis were discussed by Doerge et al. (1994).

Consistency of QTLs across populations

Linkage groups F, G, and K each had a locus for SW in both populations (Tables 3 and 4). On LG F, the two loci (Blt025-1 in Pop1 and K002-1 in Pop2) were 24 cM apart (Fig. 1) based on the concensus map of Shoemaker et al. (1997). On LG K, the two marker loci (K003-1 in Pop1 and A199-1 in Pop2) were positioned 25 cM from each other (Fig. 1). The two marker loci on LG G were 47 cM apart. Even though the QTLs on LG F and K appear to be fairly conserved across the two populations, the majority of the soybean SW QTLs are population specific.

Relationship of SW with other agronomic traits

In Pop1, SW had a positive correlation with seed yield (r=0.20*) and a negative association with maturity (r=–0.19*). There was no association of SW with plant height and lodging. Two marker loci (Blt025-1 on LG F and B166-1 on LG J) conditioned both SW and seed yield in this population. Marker locus B031-2 on LG G explained variation in both SW and maturity.

In Pop2, SW had a positive correlation with plant height (r=0.30**), lodging (r=0.32**), and seed yield (r=0.34**), but no relationship with maturity. The *Dt1* locus was associated with variation in plant height (R^2=60%), lodging (R^2=53%), seed yield (R^2=10%), and SW (R^2=10%) in Pop2. Except from *Dt1*, none of the markers conditioning SW had any significant association with plant height, lodging, or seed yield.

The apparent inconsistency between the two populations, in respect to the relationship of SW with other traits, may be due to the fact that Pop2 was segregating for growth habit while Pop1 was not. Clearly, *Dt1* (the locus for growth habit) played an important role in explaining variation in a number of agronomic traits in Pop2.

The high level of consistency of marker loci across environments (locations and years) indicates that an effective marker-assisted selection program is feasible for SW in soybean. The transgressive nature of the trait provides a real opportunity to select for progeny lines with both larger or smaller SW than either parent.

References

Ajmone MP, Monfredini G, Ludwing WF (1995) In an elite cross of maize a major quantitative trait locus controls one fourth of the genetic variation for grain yield. Theor Appl Genet 90:415–424

Brim CA (1973) Quantitative genetics and breeding. In: Caldwell BE (ed) Soybeans: improvement, production, and uses. Agronomy 16:155–186

Burris JS, Edje OT, Wahab AH (1973) Effect of seed size on seedling performance in soybeans. II. Seedling growth and photosynthesis and field performance. Crop Sci 13:207–210

Burton JW, Brim CA, Young MF (1987) Registration of Young soybean. Crop Sci 27:1093

Christian AF, Menancio-Hautea DI, Danesh D, Young ND (1992) Evidence for orthologous seed-weight genes in cowpea and mung bean based on RFLP mapping. Genetics 132:841–846

Doerge RW, Zeng Z-B, Weir BS (1994) Statistical issues in the search for genes affecting quantitative traits in populations. Proc Joint Plant Breed Symp Sen Am Soc Hort Sci and Crop Sci Soc Am, pp-15–26

Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Sci 33:660–668

Holloway JL, Knapp SJ (1993) GMendel 3.0 Users Guide. Department of Crop and Soil Sciences, Oregon State University

Hong BH, Park MW, Kim SD, Choe BH, Hong EH, Ha YW (1994) Breeding strategy for quality improvement and diversity in an upland crop. Korean J Breed 26:16–35

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP maps. Genetics 121:185–199

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ , Lincoln SE, and Newburg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lee SH, Bailey MA, Mian MAR, Carter, Jr. TE , Ashley DA, Hussey RS, Parrott WA, Boerma HR (1996 a) Molecular markers associated with soybean plant height, lodging, and maturity across locations. Crop Sci 36:728–735

Lee SH, Bailey MA, Mian MAR, Carter, Jr. TE , Ashley DA, Hussey RS, Parrott WA, Boerma HR (1996 b) Identification of quantitative trait loci for plant height, lodging, and maturity in a soybean population segregating for growth habit. Theor Appl Genet 92:516–523

Lincoln S, Daly M, Lander ES (1992) Mapping genes controlling quantitative traits with MAPMAKER/QTL. Whitehead Institute Tech Rep, 2nd edn

Mansur LM, Lark KG, Kross H, Oliveira A (1993) Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean (*Glycine max* L.). Theor Appl Genet 86:907–913

Mian MAR, Bailey MA, Ashley DA, Wells R, Carter Jr. TE , Parrott WA, Boerma HR (1996) Molecular markers associated with water-use efficiency and leaf ash in soybean. Crop Sci 36:1252–1257

Paterson AH, Damon S, Hewitt JD, Zamir D, Rabinowitch HD, Lincoln SE, Lander ES, Tanksley SD (1991) Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. Genetics 127:181–197

Paterson AH, Lin Y-R, Li Z, Schertz K, Doebley J, Pinson S, Liu S-C,Stansel J, Irvine J (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269:1714–1718

SAS (1988) SAS users guide. Statistics. SAS Institute, Cary, North Carolina

Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. Crop Sci 35:436–446

Shoemaker RC, Polzin K, Specht J, Brummer EC, Olson T, Young N, Concibido V, Wilcox J, Tamulonis JP, Kochert G, Boerma HR (1997) Genome duplication in soybean. Genetics 144(1):329–338

Sloane RJ, Patterson RP, Carter, Jr TE (1990) Field drought tolerance of a soybean plant introduction. Crop Sci 30:118–123

Smith TJ, Camper Jr HM (1975) Effect of seed size on soybean performance. Agron J 67:681–684

Stam P (1993) Construction of integrated genetic maps by means of a new computer package: Join Map. Plant J 3:739–744

Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES (1992) Identification of genetic factors contributing to heterosis in a hybrid from two elite inbred lines using molecular markers. Genetics 132:823–839

Tanksley SD (1993) Mapping polygenes. Annu Rev Genet 27: 205–233