Theor Appl Genet (1997) 95:460–467

F. J. Cerna · S. R. Cianzio · A. Rafalski
S. Tingey · D. Dyer

Relationship between seed yield heterosis and molecular marker heterozygosity in soybean

Received: 8 March 1996 / Accepted: 21 February 1997

Abstract In soybean [*Glycine max* (L.) Merr.] heterosis has been reported for seed yield. Molecular markers may be useful to select diverse parents for the expression of heterosis and yield improvement. The objective of this study was to determine if molecular markers could be used to predict yield heterosis in soybean. From each Maturity Group (MG) II and III, 21 genotypes were selected on the basis of high yield (HY), different geographic origin (GO), and isozyme loci (ISO) and for diversity in restriction fragment length polymorphisms (RFLP), and crosses were made within MGs and selection criteria groups to obtain 6 F_1 hybrids per group. The 21 parents and the 24 F_1 hybrids of each MG were evaluated for yield in replicated tests at two locations in 2 years, and midparent heterosis (MPH) and high-parent heterosis (HPH) estimates were calculated. On the basis of hybrid performance during the first year, 12 parents (3 per selection criteria group) were chosen in each MG to conduct a second RFLP analysis using 129 probes. Genetic distances (GD_M) for pairs of the 12 genotypes were calculated with this RFLP information and correlated with MPH and HPH estimates. Significant MPH averages for seed yield were observed in the combined analysis of variance in each of the four selection criteria groups of MG II, and in the HY, ISO, and GO of MG III. Significant HPH averages were observed only in the ISO and GO groups of MG II. The greatest frequency of F_1 hybrids with significant MPH was observed in the ISO and GO groups of both MGs. For HPH, the greatest frequency was observed in the ISO group of both MGs. In both MGs, the ISO group had the largest absolute MPH value; the RFLP group had generally the smallest. The observations indicated that the expression of heterosis in seed yield might be associated with diversity in the isozyme loci present in the parents. For the genotypes included in the second RFLP analysis, correlations of GD_Ms with MPH and HPH values on an entry-mean basis were low and not significant, indicating that heterosis in yield may not be associated with genetic diversity at the molecular level as determined by RFLPs. The results suggest that in soybean, parent selection on the basis of RFLPs and isozyme loci to exploit heterosis in seed yield may not be feasible. There was no association between genetic distance estimated by the RFLP analysis and seed yield heterosis, and in spite of the observed relationship between isozyme loci and heterosis for yield, the practicality of using the isozyme markers to select parents may be limited because of the reduced number of assayable isozyme loci in soybean.

Key words *Glycine max* · RFLP · Isozyme · Geographic origin · Hybrid vigor

Journal Paper No. J-16762 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project number 3107

Communicated by A. L. Kahler

F. J. Cerna (✉)
Apartado 63874, El Dorado, Panama City, Panama

S. R. Cianzio
Department of Agronomy, Iowa State University,
Ames IA 50011, USA

A. Rafalski · S. Tingey · D. Dyer
Agricultural Biotechnology, DuPont de Nemours Company,
Wilmington, DE 19898, USA

Introduction

In soybean [*Glycine max* (L.) Merr.], hybrid vigor has been observed in several traits of economic importance (Chauhan and Singh 1982; Mehta et al. 1984; Nelson and Bernard 1984; Paschal and Wilcox 1975). Burton (1987) reported that 85% of the F_1 crosses showed midparent heterosis for seed yield and 62% had high-parent heterosis. Average midparent heterosis for yield ranged from 7.9% (Nelson and Bernard 1984) to 35.5%

(Veatch 1930), and average high-parent heterosis ranged from 8% (Paschal and Wilcox 1975) to 26% (Chauhan and Singh 1982). In these studies, genotypes for hybridization were selected primarily, based on seed yield, but also on differences in morphological traits, ancestry, and geographic origin.

Molecular markers, such as restriction fragment length polymorphisms (RFLP) and isozymes, detect locus differences among genotypes and represent a powerful tool for the assessment of genetic diversity in plant species (Burr et al. 1983; Lander and Botstein 1989; Kiang and Gorman 1983; Melchinger 1993; Tanksley 1983). Because heterosis is associated with the interaction of different alleles at a locus (Castle 1946; Hull 1945; Jones 1945), it has been suggested that molecular marker diversity may be used to select parents for hybridization (Burr et al. 1983; Kiang and Gorman 1983).

Studies to determine the association between molecular markers and heterosis expression have been conducted in several crops. In maize (*Zea mays* L.), results indicated that isozyme allelic differences between lines are not predictive of hybrid performance (Lamkey et al. 1987; Price et al. 1986). In rice (*Oryza sativa* L.), Peng et al. (1988) did not find any association between the magnitude of heterosis in the F_1s and isozyme variation among parents. Deng (cited by Peng et al. 1988), however, suggested that esterase and peroxidase patterns in the parents may be of value for predicting F_1 yield heterosis. Studies conducted in maize indicate that there is no relationship between RFLP markers and heterosis expression (Dudley et al. 1991; Godshalk et al. 1990; Lee et al. 1989; Melchinger et al. 1990). A study by Smith et al. (1990), however, reported significant correlations between RFLP-based genetic distance (GD_M) and heterosis expression of yield, suggesting that measures of similarity calculated from RFLP data could allow maize breeders to predict combination of lines resulting in high-yielding single-cross hybrids.

In soybean, Kiang and Gorman (1983) observed that populations from which high-yielding lines had been identified were developed by crossing genotypes with significantly greater genetic distances than any two others drawn at random. The authors estimated genetic distances by using 12 polymorphic isozyme systems as markers on 100 soybean cultivars from the northern region of the USA. Relationships between isozyme genotypes and quantitative traits have also been determined in 2 interspecific crosses (Graef et al. 1989; Suarez et al. 1991), although no information is available for crosses with the *G. max* species. Diversity of RFLP markers was reported by Keim et al. (1989) in a survey of 58 genotypes of the wild and cultivated species using 17 RFLP markers. There is, however, no publised information in soybean on the relationship between diversity in isozymes and RFLP markers with the expression of heterosis. The objective of our study was to evaluate the association of genetic divergence estimated by RFLP analysis and isozyme loci with the expression of heterosis for seed yield in soybean. If such a relationship were to exist, diversity at the molecular level could be used as criteria in the selection of parents for hybridization.

Materials and methods

For the study, 21 soybean genotypes in each of two Maturity Groups (MG) II and III, were crossed to develop F_1 hybrids (Table 1). In each MG the genotypes were selected on the basis of high seed yield (HY), different geographic origin (GO), and for diversity in restriction fragment length polymorphisms (RFLP) markers and isozyme loci (ISO). Four groups of genotypes according to each of the four selection criterion were formed, and 6 F_1 hybrids were obtained in each group. Crosses in the GO, RFLP, and ISO groups were conducted to ensure maximal divergence between parents. The HY group simulated practical breeding situations in which populations for yield improvement are developed by crossing high-yielding cultivars. Genotypes in the HY groups were crossed in a partial diallel arrangement.

In each MG 4 high-yielding genotypes among public cultivars and experimental lines were chosen for the HY criterion (Table 1). For the GO group, 6 genotypes were selected in each MG that were introduced to the United States from different geographic areas of the world. Six genotypes in MG II and 4 in MG III were selected based on genetic distances estimated from an analysis of RFLP divergence (Keim et al. 1989). Five genotypes in MG II and 7 in MG III that differed in isozymes at six loci were selected for the ISO criterion. The isozymes were: aconitase [aconitate hydrate, enzyme commission (EC) (E.C. 4.2.1.3, *Aco3* locus), isocitrate dehydrogenase (E.C. 1.1.1.42, *Idh1* and *Idh2*, loci), acid phosphatase (E.C. 3.1.3.1, *Ap* locus), diaphorase (E.C. 1.6.4.3, *Dia1* locus), endopeptidase (E.C. 3.4.23.6, *Enp-b* locus), and superoxide dismutase (E.C. 1.15.1.1, *Sod1*).

In plantings conducted at the Iowa State University-University of Puerto Rico research site in Isabela, PR, and at the Iowa State University Agronomy and Agricultural Engineering Research Center near Ames Iowa in May 1989, 48 F_1 seeds were obtained for each cross. This seed was used to conduct replicated yield tests during 1990. The same crosses were repeated during 1990 to obtain the seed used in the yield tests in 1991. Flower, pubescence, and hilum color were used as markers to confirm F_1 hybrid plants, which were then harvested in bulk. For the combinations in which no such markers were available, the F_1 plants were harvested individually forming each a subpopulation. Subpopulations were maintained until the planting of the F_2 seed, at which time observations of segregation for plant height and maturity were used to confirm the hybrid nature of the subpopulations.

For each MG, the 21 parents and the 24 F_1 hybrids were evaluated in yield tests conducted in 2 years at each of two locations at the Iowa State University Agronomy and Agricultural Engineering Research Center (Ames and Burkey Farms). The experiments were planted on 30 May 1990 and 29 May 1991. Soil type in the two locations is a Nicollet loam (fine loamy, mixed, mesic, Aquic Hapludoll).

Plots consisted of single unbordered hills, spaced in a grid of 102×102 cm, planted with 12 seeds in a linear distance of 15 cm. Hills with less than 3 plants were considered missing plots, and seed of the cultivars 'Kenwood' (MG II) and 'Resnik' (MG III) were sown before plants in adjacent plots had one fully developed trifoliolate leaf to provide adequate competition to neighboring plots.

A randomized complete-block design with two replications was used at each location and year. Analyses of variance were conducted for individual years over locations and combined over the four

Table 1 Genotypes of Maturity Groups II and III, their origin, and hybrid combinations obtained with each selection group

Selection criterion	Genotype	Origin	Hybrid combinations
Maturity Group II			
High yield	1-Hack	USA	2×1, 3×2, 4×3
	2-Elgin 87	USA	4×2, 1×4, 3×1
	3-Conrad	USA	
	4-Kenwood	USA	
RFLP	5-Corsoy	USA	6×5, 5×10, 8×7
	6-Asgrow25AF	USA	7×6, 9×6, 10×9
	7-Seneca	China	
	8-Harosoy	USA	
	9-Richland	China	
	10-A80-244036	USA	
Isozyme	11-Magna	USA	12×11, 13×12, 14×11
	12-Mukden	China	15×12, 15×13, 15×14
	13-Funman	USA	
	14-Bansei	Japan	
	15-Beeson 80	USA	
Geographic origin	16-Tastee	Japan	17×18, 17×16, 18×21
	17-PI30.594	Manchuria-China	19×20, 19×16, 19×21
	18-PI69.501	China	
	19-PI80.671	Japan	
	20-PI84.580	Korea	
	21-Richland	USA-China	
Maturuty Group III			
High yield	22-Zane	USA	23×22, 24×23, 25×24
	23 Resnik	USA	25×23, 22×25, 24×22
	24-A86-801024	USA	
	25-Sherman	USA	
RFLP	26-A81-356022	USA	29×26, 27×26, 26×28
	27-Illini	China	28×27, 29×27, 29×28
	28-Dunfield	China	
	29-PI437.477B	USSR	
Isozyme	30-Cloud	China	31×30, 32×30, 35×30
	31-Cumberland	USA	34×36, 33×30, 34×30
	32-Will	USA	
	33-Williams 82	USA	
	34-Mandell	USA	
	35-BSR 301	USA	
	36-Shelby	USA	
Geographic origin	37-PI80.470	Japan	38×37, 41×40, 39×42
	38-PI104.708	Poland	42×40, 39×37, 40×37
	39-PI61.940	China	
	40-PI82.235	Korea	
	41-Manchuria	China	
	42-PI54.592	China	

environments (i.e., 2 years and two locations). Years, locations, and replications were considered random effects, and selection criteria and genotypes were considered fixed effects. The sum of squares due to genotypes was apportioned among and within selection criteria (Table 2).

Seed yield data were collected on an individual-plot basis expressed in g m^{-2} as the weight of the seed artificially dried at 38°C during 72 h. For each hybrid, midparent heterosis (MPH) and high-parent heterosis (HPH) were calculated for seed yield as:

$$MPH = F_1 - MP$$

$$HPH = F_1 - HP$$

where:

MPH = midparent heterosis;
MP = midparent value;
HPH = high-parent heterosis; and
HP = high-parent value.

Single degree-of-freedom contrasts were calculated to test differences between the mean performance of the F_1 hybrid and the MP value and between generations within each selection criterion group. To test differences between the mean performance of the F_1 and its respective HP value, we used a LSD test. Average MPH and HPH heterosis estimates were calculated for each group and their statistical significance was also determined.

After results of the first yield test conducted in 1990 were statistically analyzed, 3 genotypes from each of the four selection criteria groups of each MG were chosen to conduct a second RFLP analysis using 129 soybean genomic and cDNA probes. The genotypes were chosen based on seed yield of the F_1 hybrids (Table 6). Within groups 2 genotypes were chosen if their hybrid exhibited significant MPH values, and the third was chosen if the hybrid obtained in crosses to either of the first 2 did not show significant MPH values. In MG II, there were two exceptions: the 6 hybrids of the ISO group had significant MPH effects, and in the RFLP group none of the hybrids had significant MPH effects (Table 3). In these groups 2 of

Table 2 Mean squares for seed yield from the analysis of variance combined over 2 years and two locations (*HY* high yield. *ISO* isozyme. *GO* geographic origin)

Source of variation	Maturity Group II		Maturity Group III	
	df	Mean squares	*df*	Mean squares
Year (Y)	1	130 419	1	87 397
Location (L)	1	82 358	1	117 975
Y*L	1	62 586	1	15 316
Replication/YL	4	1 686	4	688
Genotype (G)	44	16 567	44	27 361*
Among criteria	3	108 525	3	230 674*
Within HY	9	9 166	9	5 965
Within RFLP	11	10 429	9	18 991
Within ISO	10	12 918	12	11 909
Within GO	11	6 998	11	13 121*
Y*G	44	2 927	44	3 067
L*G	44	1 500	44	1 988
Y*L*G	44	1 656	44	1 894
Error	176	1 139	176	1 586

* Significant at the 0.05 probability level

the genotypes were chosen for having the largest absolute MPH estimate in hybrid combinations. The third was chosen because it had the smallest absolute MPH value in hybrid combination with either of the other 2.

DNA was extracted from 3 g of lyophilized leaf tissue obtained from 10 greenhouse-grown plants (Murray and Thompson 1980). For each genotype, a 5-μg sample of genomic DNA was digested individually with the restriction enzymes *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, and *Pst*I, as well as with *Hpa*II and *Msp*I. Blots containing parental DNA digested with *Bc1*I, *Bg1*II, *Cfo*I, *Dra*I, *Msp*I, *Sca*I, *Ssp*I, and *Xba*I were also used with cloned gene probes. Agarose gel electrophoresis and Southern blotting were done under standard conditions (Maniatis et al. 1982). DNA was transferred to GeneScreen (DuPont) or Hybond-N (Amershan) uncharged nylon membranes. DNA probes were radiolabelled to approximately 10^9 cpm/μg DNA by random-primer synthesis of isolated inserts (Feinberg and Vogelstein 1983). Hybridization occurred at 65°C, during 20–24 h with gentle mixing, inside an air incubator. Blots were placed in polypropylene project folders (C-line Products) and autoradiographed using Kodak X-Omat AR film for 18–100 h. Blots were re-used up to ten times, after stripping in 0.4 *N* NaOH (30 min at 42°C) and washing in 0.2 *M* TRIS-HCl pH 7.5, 0.1× SSPE, 0.1% SDS (30 min, 42°C).

The soybean genomic probes were derived from genomic soybean DNA digested with the methylation-sensitive restriction enzyme *Pst*I. Only low-copy-number genomic clones were used as probes. With the exception of the genes *SAC* (from R. Meagher, University of Georgia), *PD1* (from P. Scolnik, DuPont de Nemours Company), and *DS1*, now called *RPS24.1* (from N. Yadav, DuPont de Nemours Company), probes were anonymous *Pst*I genomic clones derived by the DuPont de Nemours Company.

Bands for each parent profile from the autoradiograms were coded 1 or 0 for presence or absence of the band, respectively. Genetic distance (GD_M) between all pairs of parents were calculated as $1 - S_{XY}$, applying the method developed by Nei and Li (1979) [i.e., $S_{XY} = 2N_{XY}/(N_X + N_Y)$]

where:

S_{XY} = measure of genetic similarity between a pair of parents;
N_{XY} = number of bands common to parents X and Y;
N_X and N_Y = number of bands for parents X and Y, respectively.

Correlation coefficients of GD_M with average seed yield performance and heterotic effects of the F_1 hybrids were calculated (SAS 1985).

Results and discussion

For both MGs, the main effects of year and location were not significant for seed yield (Table 2). The interaction of year × location, however, was significant due to a change in the relative ranking of the genotypes. In 1990, the Ames location had higher average yield than Burkey, whereas the opposite occurred in 1991. Significant differences were observed among genotypes and selection criteria groups, and also within each selection group, except for the HY criterion. For MG II the year × genotype interaction was significant due to changes in the relative ranking and in the size of the differences among genotypes.

For MG II, the ISO and GO selection criteria groups had average MPH and HPH that were positive and significant in each year and in the combined analysis over environments (Table 3). For the HY group, MPH averages were significant in 1990 and in combined environments, whereas the HPH estimate was significant only in 1990. For the RFLP group, significant MPH averages were observed in each of the 2 years and in the combined analysis. For MG III, significant MPH averages were observed for each of the four selection criteria groups in 1990 and in the combined analysis except for the RFLP group. A significant HPH average was observed only for the GO selection criterion group in 1990. For both MGs the largest average estimate of MPH combined over environments was observed for the ISO group, followed by the GO group. For HPH, results were similar but only observed for MG II. Relative sizes of the estimates in each year varied among groups. In general, the smallest heterosis estimates were observed for the RFLP group.

In the combined analysis over environments and across all selection criteria groups, seed yield of each of 17 hybrids in MG II and 15 in MG III was significantly different from their respective MP values (Table 3). These hybrids were observed in each selection criterion group; however, the highest frequencies were in the ISO and GO groups of both MGs. Ten hybrids in MG II (5 in ISO, 4 in GO, and 1 in HY groups), and 2 in MG III (ISO selection criterion group) were significantly superior to their respective HP values. For both MGs and for each selection criterion group, frequencies varied for individual years.

The GD_M estimates from the second RFLP analysis conducted on the 12 genotypes of each MG ranged from 15% between 'Williams 82' and 'BSR 301' to 46% between 'Zane' and 'PI 82.235' when all possible hybrid combinations among the 24 genotypes were considered (Table 4). The smallest GD_M estimates were always observed between genotypes released from the US soybean germplasm pool, and the largest were observed between a US-released genotype and a plant introduction. Among the hybrids obtained for the study, the smallest GD_M was 18%, between 'Kenwood'

Table 3 Average mid- and high-parent heterosis for seed yield of F_1 hybrids of Maturity Groups II and II and number of F_1 hybrids with each selection criterion group with significant ($P \leq 0.05$) mid- and high-parent heterosis for 1990, 1991, and combined over years (*MPH* mid-parent heterosis. *HPH* high parent heterosis)

Selection criterion	Maturity Group II						Maturity Group III					
	1990		1991		Combined		1990		1991		Combined	
	MPH	HPH	MPH	HPH	MPH	HPH	MPH	HPH	MPH	HPH	MPH	HPH
	Seed yield (g m^{-2})											
High yield	55*	44*	25	17	40*	31	55*	36	14	− 2	34*	16
RFLP	34*	− 1	44*	17	39	8	45*	8	8	− 26	25	− 7
Isozyme	66*	56*	72*	49*	69*	52*	54*	25	50	7	52	16
Geographic origin	37*	29*	65*	49*	51*	39*	66*	46*	35	− 4	49*	22
	Number of hybrids											
High yield	2	0	2	1	2	1	2	2	2	1	3	0
RFLP	0	0	3	0	3	0	2	0	2	0	2	0
Isozyme	6	4	5	2	6	5	3	1	3	0	5	2
Geographic origin	5	2	5	4	6	4	5	3	3	0	5	0

* Significantly higher than mid- or high-parent values at the 0.05 probability level

Table 4 Genetic distance matrix in percentage for 24 genotypes estimated on the basis of 129 RFLP

Genetoype	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. Kenwood	0																							
2. Asgrow25AF	28	0																						
3. Beeson 80	28	14	0																					
4. BSR 301	22	34	32	0																				
5. Cloud	37	38	44	36	0																			
6. Conrad	25	23	29	26	37	0																		
7. Dunfield	32	32	33	35	36	34	0																	
8. Elgin 87	18	22	19	29	38	20	31	0																
9. Funman	39	36	33	36	30	34	29	27	0															
10. Harosoy	28	27	27	32	45	23	32	19	31	0														
11. Illini	30	26	28	30	36	29	29	30	36	30	0													
12. PI54.592	37	24	26	35	33	30	25	27	26	26	28	0												
13. PI61.940	35	29	23	31	35	34	27	28	24	33	29	24	0											
14. PI80.671	31	33	26	27	38	31	34	27	28	26	29	30	25	0										
15. PI82.235	33	35	33	43	37	37	37	36	41	34	38	32	36	34	0									
16. Resnik	29	27	34	28	34	15	34	24	28	30	33	32	34	34	40	0								
17. Seneca	37	29	26	40	36	35	26	31	30	32	29	27	28	28	34	40	0							
18. Sherman	26	25	24	21	36	15	35	20	34	27	27	34	32	27	37	23	27	0						
19. Tastee	31	34	29	40	39	34	36	31	34	33	35	32	31	29	19	38	27	36	0					
20. Zane	32	36	31	28	38	27	31	24	31	35	31	35	31	29	46	31	29	22	44	0				
21. Williams 82	27	34	31	15	35	27	30	28	30	31	31	36	32	26	42	27	31	18	39	20	0			
22. PI437.477B	24	28	25	27	34	30	27	26	32	33	24	26	26	29	35	31	34	33	33	29	28	0		
23. Mukden	34	22	22	41	35	36	33	28	30	31	30	28	24	31	28	35	21	30	28	37	39	27	0	
24. Richland	33	26	25	32	40	33	32	30	37	31	30	27	29	35	35	36	33	29	29	44	32	32	28	0

and 'Elgin 87' (HY criterion group of MG II) (Table 5). Both cultivars have a common immediate ancestor 'Elgin', which may explain this result (Cianzio et al. 1990; Fehr et al. 1988). The largest GD_Ms were observed in the ISO groups of both MGs (i.e., 'BSR 301' × 'Cloud, 'Williams 82' × 'Cloud', 'Beeson 80' × 'Funman') and in the GO group of MG II (i.e., 'PI 80.671' × 'Richland').

The GD_M estimate between 2 genotypes was not related to the heterosis expression of yield in the hybrid (Table 5). For instance, in the GO selection criterion group of MG III, the hybrid between 'PI 54.940' and 'PI 54.595' had a significant MPH estimate and a GD_M of 24%. The hybrid between 'PI 54.592' and 'PI 82. 235', which did not show heterosis, had a GD_M of 32%. Similar results were also observed in other hybrids; i.e., 'PI 80.671' × 'Tastee' had a significant MPH estimate and a GD_M of 30%, while the same PI crossed to 'Richland' produced a hybrid that was not heterotic for seed yield and had a GD_M of 34%. The

Table 5 Mid- and high-parent heterosis for seed yield of 2 F_1 hybrids chosen in each selection criterion group of Maturity Groups II and III having one common parent and the RFLP-based genetic distance between the parents of each hybrid (*MPH* mid-parent heterosis, *HPH* high-parent heterosis)

Selection criterion	Parentage of hybrid	GD_M	MPH		HPH	
			1990	Combined	1990	Combined
	Maturity Group II	%	$g\ m^{-2}$			
High yield	Kenwood × Conrad	26	102*	93*	97	85*
	Kenwood × Elgin 87	18	26	17	10	7
RFLP[b]	Harosoy × Seneca	32	59	45*	57	45
	Asgrow25AF × Seneca	29	7	30	−17	5
Isozyme	Beeson 80 × Mukden	22	72*	65*	72*	44*
	Beeson 80 × Funman	33	46*	52*	30	45*
Geographic origin	PI80.671 × Tastee	30	50*	66*	48*	61*
	PI80.671 × Richland	34	14	38*	4	28*
	Maturity Group III					
High yield	Sherman × Resink	23	72*	43*	56*	27
	Resnik × Zane	31	30	44*	30	39
RFLP[b]	Dunfield × Illini	29	57*	51*	46	32
	PI437.477B × Dunfield	27	24	20	2	18
Isozyme	BSR 301 × Cloud	36	83*	54*	66*	31
	Williams 82 × Cloud	35	57	36	16	−18
Geographic origin	PI54.940 × PI54.592	24	60*	60*	51*	51
	PI54.592 × PI82.235	32	38	13	18	−2

* Significantly ($P \leq 0.05$) superior in yield to the mid-parent value and the high-parent of the cross
[a] GD_M = Genetic distance estimate
[b] Selected on the basis of smallest and largest absolute values for mid-and high-parent heterosis

correlation coefficients on an entry-mean basis between GD_M and the heterosis estimates for all hybrids as a group were low and not significant for MPH($r = 0.08$) and for HPH ($r = -0.18$). The correlation between GD_M and seed yield also was low and not significant ($r = -0.29$).

Our results indicated that RFLP marker diversity was not associated with the heterosis expression of seed yield in soybean. Although significant average MPH values were observed in hybrids of the RFLP groups, the frequency of heterotic hybrids was the least or next to the least of all selection criteria groups, and there were no significant HPH estimates. In absolute values, these estimates were also generally the smallest. Further evidence was obtained from the second RFLP analysis conducted on the set of 12 parents of each MG intentionally selected to maximize differences in heterosis. Correlations between GD_M estimates with heterosis effects and seed yield were both negligible and not significantly different from zero. In this second RFLP analysis, 129 markers were used which represented, for the genotypes originally included in the RFLP criterion group of both MGs, a 7.5-fold increase from the number of markers used by Keim et al. (1989). Melchinger (1993), in a review of the published literature in maize, also concluded that GD_M estimates based on a random set of RFLP markers are of no value in predicting hybrid performance.

The results also suggested that allelic differences in the six isozymes could be associated with the heterosis observed in seed yield. In both MGs, the ISO groups had the greatest frequency of hybrids with significant MPH and HPH estimates, and the estimates were also the largest. Selection of parents on the basis of their isozyme loci may be of limited use, however, because the number of assayable isozyme loci in soybean is limited. Griffin and Palmer (1987) reported that the six isozymes in which the genotypes of the ISO group differed are independent of each other; therefore, 6 of the 20 linkage groups of soybean were represented in the genotypes. Stuber (cited by Melchinger 1993) reviewed the literature on the use of isozymes as possible predictors of hybrid performance in maize and other crops. In all studies, allelic differences at enzyme loci were positively associated with hybrid performance for grain yield, but correlations generally were too low to be of predictive value. Because relatively few (<20) polymorphic isozyme loci were available in these investigations, it remained unclear if results were due to poor marker coverage of the genome or if other causes were involved (Melchinger 1993).

These findings suggest that, in soybean, parent selection to exploit heterosis in seed yield based on genetic divergence estimated by RFLP analysis and by isozyme loci may not be a feasible approach. Bernardo (1992) indicates that molecular marker heterozygosity would be most useful for predicting hybrid performance in crop species where (1) dominance effects are strong, (2) heterotic groups are complementary and allele frequencies at individual loci in the parental inbreds are negatively correlated, (3) trait heritability is high, (4) average parental allele frequencies vary only

within a narrow range, (5) at least 30–50% of the quantitative trait loci (QTLs) are linked to molecular markers, and (6) not more than 20–30% of the molecular markers are randomly dispersed or unlinked to QTL. In soybean, some of these conditions may not be met, and current knowledge may not yet be sufficient to establish the usefulness of molecular markers as predictors for the heterosis expression of seed yield. For instance, most gene action reported in soybean for economically important traits is additive, and heritability estimates are low (Brim 1973). In other instances, even though relationships between QTLS and molecular markers have been reported (Diers et al. 1992; Graef et al. 1989; Keim et al. 1990; Suarez et al. 1991), no relationships have yet been determined between RFLP markers and QTL for seed yield.

An assessment of the usefulness of RFLP markers in breeding soybeans for yield improvement may therefore need further consideration. The evaluation of the non-association between RFLP marker diversity and the expression of heterosis we found in our study would require more genotypes in hybrid combinations, and also that the RFLP markers used to estimate genetic divergence be linked to QTL for seed yield. At the time we conducted the study, mapping of the soybean genome had just begun, and only a small number of markers had been identified. Currently, more than 600 markers have been placed on the soybean map (R. Shoemaker. personal communication, USDA-ARS, Iowa State University). As more markers are identified and the map becomes more saturated, there will be opportunities to detect linkage associations between markers and QTLs for seed yield. An assessment of the relative advantage of using isozyme loci as a measure of genetic divergence versus RFLPs will require that both selection criteria be applied on the same set of genotypes, which was not possible in the study we conducted. Any differences observed between the two selection criteria groups could then be attributed to the molecular marker used rather than to differences in genotypes.

References

Bernardo R (1992) Relationship between single-cross performance and molecular marker heterozygosity. Theor Appl Genet 83:628–634

Brim CA (1973) Quantitative genetics and breeding. In: Caldwell BE (ed) Soybeans: improvement, production, and uses. ASA CSSA SSSA, Madison, Wis., pp 155–186

Burr B, Evola SV, Burr FA, Beckman JS (1983) The application of restriction fragment length polymorphisms to plant breeding. In: Setlow JK, Hollaender A (eds) Genetic engineering principles and methods, vol 5. Plenum Press, London, pp 45–59

Burton JW (1987) Quantitative genetics: results relevant to soybean breeding. In: Wilcox JR (ed) Soybeans: improvement, production, and uses. ASA CSSA SSSA, Madison, Wis PP 211–242

Castle WE (1946) Genes which divide species or produce hybrid vigor. Proc Natl Acad Sci USA 32:145–149

Chauhan VS, Singh BB (1982) Heterosis and genetic variability in relation to genetic divergence in soybean. Indian J Genet 42:324–328

Cianzio SR, Schultz SP, Voss BK, Fehr WR (1990) Registration of 'Kenwood' soybean. Crop sci 30:1162

Diers BW, Cianzio SR, Shoemaker RC (1992) Possible identification of quantitative trait loci affecting iron efficiency in soybean. J Plant Nutr 15:2127–2136

Dudley JW, Saghai Maroof MA, Rufener GK (1991) Molecular markers and grouping of parents in maize breeding programs. Crop Sci 31:718–723

Fehr WR, Walker AK, Schmitthenner AF, Cianzio SR, Voss BK (1988) Registration of Elgin 87 soybean. Crop Sci 28:1025

Feinberg AP, Vogelstein B (1983) A technique for labelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Godshalk EB, Lee M, Lamkey KR (1990) Analysis of the relationship of restriction fragment length polymorphisms to maize single-cross hybrid performance. Theor Appl Genet 80:273–280

Graef GL, Fehr WR, Cianzio SR (1989) Relation of isozyme genotypes to quantitative characters in soybean. Crop Sci 29:683–688

Griffin JD, Palmer RG (1987) Inheritance and linkage studies with five isozyme loci in soybean. Crop Sci 27:885–892

Hull FH (1945) Recurrent selection for specific combining ability in corn. J Am Soc Agron 37:134–145

Jones DF (1945) Heterosis resulting from degenerative changes. Genetics 30:527–542

Keim P, Shoemaker RC, Palmer RG (1989) RFLP diversity in soybean. Theor Appl Genet 77:786–792

Keim P, Diers BW, Olson TC, Shoemaker, RC (1990) RFLP mapping in soybeans: association between marker loci and variation in quantitative traits. Genetics 126:735–742

Kiang YT, Gorman MB (1983) Soybean. In: Tanksley SD, Orton TF (eds) Isozymes in plant genetics and breeding, part B. Elsevier Science Publ, Amsterdam, pp 295–328

Lamkey KR, Hallauer AR, Kahler AL (1987) Allelic differences at enzyme loci and hybrid performance in maize. J Hered 78:231–234

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–189

Lee M, Godshalk EB, Lamkey KR, Woodman WL (1989) Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. Crop Sci 29:1067–1071

Maniatis T, Fristch EF, Sambrook J (1982) Molecular cloning: a laboratory manual, 1st edn. Cold Spring Harbor Laboratory Press, New York

Metha SK, Lal MS, Beohar ABL (1984) Heterosis is soybean crosses. Indian J Agric Sci 54:682–684

Melchinger AE (1993) Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. In: Buxton DR, Shibles R, Forsberg RA, Blad BL, Asay KH, Paulsen GM, Wilson RF (eds) Int Crop Sci I. ASA CSSA SSSA, Madison Wis, pp 621–628

Melchinger AE, Lee M, Lamkey KR, Hallauer AR, Woodman WL (1990) Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. Theor Appl Genet 80:488–496

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5256–5273

Nelson RL, Bernard RL (1984) Production performance of hybrid soybeans. Crop Sci 24:549–553

Paschal EH, Wilcox JR (1975) Heterosis and combining ability in exotic soybean germplasm. Crop Sci 15:344–349

Peng JY, Glaszmann JC, Virmani SS (1988) Heterosis and isozyme divergence in indica rice. Crop Sci 28 : 561–563

Price SC, Kahler AL, Hallauer AR, Charmley P, Giegel P (1986) Relationships between performance and multilocus heterozygosity at enzyme loci in single-cross hybrids of maize. J Hered 77 : 341–344

SAS Institute (1985) SAS user's guide: statistics. SAS Institute, Cary N.C.

Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield, grain yield, heterosis, and RFLPs. Theor Appl Genet 80 : 833–840

Suarez JC, Graef GL, Fehr WR, Cianzio SR (1991) Association of isozyme genotypes with agronomic and seed composition traits in soybean. Euphytica 52 : 137–146

Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Biol Rep 1 : 3–8

Veatch C (1930) Vigor in soybeans as affected by hybridity. J Am Soc Agron 22 : 289–310