

# Quantitative trait loci in Two Soybean Recombinant Inbred Line Populations Segregating for Yield and Disease Resistance

J. Yuan, V. N. Njiti, K. Meksem, M. J. Iqbal, K. Triwitayakorn, My. A. Kassem, G. T. Davis, M. E. Schmidt, and D. A. Lightfoot\*

## ABSTRACT

Molecular markers linked to quantitative trait loci (QTL) can assist soybean [*Glycine max* (L.) Merr.] breeders to combine traits of low heritability, such as yield, with disease resistance. The objective of this study was to identify markers linked to yield QTL in two recombinant inbred line (RIL) populations ['Essex' × 'Forrest' (E×F;  $n = 100$ ) and 'Flyer' × 'Hartwig' (F×H;  $n = 94$ )] that also segregate for soybean cyst nematode (SCN) resistance genes (*rhg1* and *Rhg4*). Each population was yield tested in four environments between 1996 and 1999. The resistant parents produced lower yields. Heritability of yield across four environments was 47% for E×F and 57% for F×H. Yield was normally distributed in both populations. High yielding, SCN resistant transgressive segregants were not observed. In the E×F RIL population, 134 microsatellite markers were compared against yield by ANOVA and MAPMAKER QTL. Regions associated with yield were identified by SATT294 on linkage group (LG.) C1 ( $P = 0.006$ ,  $R^2 = 10\%$ ), SATT440 on LG. I ( $P = 0.007$ ,  $R^2 = 10\%$ ), and SATT337 on LG. K ( $P = 0.004$ ,  $R^2 = 10\%$ ). Essex provided the beneficial allele at SATT337. Mean yields among F×H RILs were compared against 33 microsatellite markers from LG. K. In addition 136 markers from randomly selected LGs were compared with extreme phenotypes by bulk segregant analysis. Two regions on LG. K (20 cM apart) associated with yield were identified by SATT326 ( $P = 0.0004$ ,  $R^2 = 15\%$ ) and SATT539 ( $P = 0.0008$ ,  $R^2 = 14\%$ ). Flyer provided both beneficial alleles. Both populations revealed a yield QTL in the interval (5 cM) between SATT337 and SATT326. These populations may share a common allele for yield in this region, given that about 40% of Flyer genome derived from Essex.

SELECTION FOR increased yield potential is the main goal of plant breeding (Fehr, 1987). Much of the yield increases over the past 60 years have been due to genetic advances by intercrossing existing varieties (Specht et al., 1999). However, yield is a multigenic trait and therefore the yield potential of lines derived by inter-crossing is difficult to predict without extensive field tests.

Crop improvement has relied on phenotypic selection for soybean yield that involves carrying a large number of lines with high yield potential to later stages of breeding programs (Stuber et al., 1992). Phenotypic selection for soybean yield is complicated by significant genotype × environment interactions (G×E) that influence yield and other quantitative traits. Hence, selection for high and stable yield requires evaluation in multiple environ-

ments over several years which is expensive, time consuming, and labor intensive (Maughan et al., 1996).

Components of yield are often identifiable which aid the selection of yield (Fehr, 1987; Specht et al., 1999). In soybeans, the basis of yield improvement is unclear, but maturity and growth habit have major effects (Mansur et al., 1996; Orf et al., 1999; Specht et al., 1999). Resistance to disease is usually a strong component of yield in disease infested environments (Njiti et al., 1998).

Disease resistance in cultivars (particularly SCN resistance) has consistently been associated with a 1–2% decrease in yield when disease was absent (Concibido et al., 1997). In addition, many SCN resistant cultivars appear to display poor combining ability during intercrossing (Concibido et al., 1997). Sudden death syndrome (SDS) resistance has also been associated with low yield potential (Rupe et al., 1993).

Genetic maps have been useful for soybean genome analysis. Maps have allowed the identification of many economically important soybean genes conditioning quantitative trait loci (QTL), including those for disease resistance (Webb et al., 1995; Chang et al., 1996; Concibido et al., 1997; Meksem et al., 1999), yield (Mansur et al., 1996; Njiti et al., 1997), or yield conditioned by disease resistance (Hnetkovsky et al., 1996). Molecular markers can also be employed in exotic germplasm to discover new genes and alleles that can be introgressed into elite germplasm to increase yield potential as well as genetic diversity (Lande and Thompson, 1990; Fray et al., 2000).

Molecular markers linked to loci conditioning disease resistance in soybeans have mostly been detected in populations in which one parent is not adapted. Some of these loci have been reported to be associated with low yield in non-infested locations (Mansur et al., 1993; Mudge et al., 1996). However, disease resistance mapping populations are rarely suitable for mapping QTL underlying yield in elite cultivars and their derived intercross populations.

The objective of this study was to identify QTL among 100 RILs from the cross Essex (Smith and Camper, 1973) × Forrest (Hartwig and Epps, 1973) and 94 RILs from the cross Flyer (McBlain et al., 1990) × Hartwig (Anand, 1992). The pair of cultivars involved in each cross are adapted, yet contrast for yield potential. Progeny of both crosses segregate for the SCN (*rhg1* and *Rhg4*) and SDS resistance (*Rfs1*, *rfs2*, and *rft*; Meksem et al., 1999, 2001). Both of these diseases have been associated with low yield potential (Rupe et al., 1993; Concibido et al., 1997). This study examined the association between disease resistance (SCN and SDS) and yield among RILs populations derived from intercross of adapted cultivars.

J. Yuan, V.N. Njiti, K. Meksem, M.J. Iqbal, K. Triwitayakorn, My. A. Kassem, M.E. Schmidt, and D.A. Lightfoot, Dep. of Plant, Soil and General Agriculture, Southern Illinois University at Carbondale, Carbondale, IL 62901-4415; G.T. Davis, Dep. of Biological and Allied Health Science, Bloomsburg University, Bloomsburg, PA 17815. This research was supported by Illinois Soybean Program Operating Board (ISPOB) grants 95-222-24-3 and 98-222-25-2. Received 25 Jan. 2001. \*Corresponding author (ga4082@siu.edu).

## MATERIALS AND METHODS

### Plant Material

This study involved two F<sub>5</sub>-derived recombinant inbred line (RILs) populations from Essex × Forrest (E×F, *n* = 100) and Flyer × Hartwig (F×H, *n* = 94). The parents for each cross contrast for yield potential, SCN and SDS resistances (Smith and Camper, 1973; Hartwig and Epps, 1973; McBlain et al., 1990; Anand, 1992; Hnetkovsky et al., 1996). The F×H RILs (F<sub>5:11</sub>) also segregate for growth habit. The E×F derived RILs (F<sub>5:16</sub>) have been described extensively (Hnetkovsky et al., 1996; Chang et al., 1996, 1997). The F×H derived RILs were described by Prabhu et al. (1999).

In 1996, the E×F population was planted in three southern Illinois locations (Carbondale, Ridgway, and Desoto). Soil types were Stoy (Fine-silty, mixed, mesic, Aquic, Hapludalfs) in Carbondale, Camden (Fine-silty, mixed, mesic, typic Hapludalfs) in Desoto, and Reesville silt loam (Fine-silty, mixed, superactive, mesic, aeric, Epiaqualfs) in Ridgway. Some water deficit stress occurred in late July and August. In 1997, the E×F RILs were planted in Desoto on Camden soil (Fine-silty, mixed, mesic, typic Hapludalfs). These environments were free from visible disease symptoms and had nondetectable cyst nematode counts.

Forty F<sub>5:10</sub>-derived near-isogenic lines from one E×F recombinant inbred line (E×F11) that was heterozygous in the BARC\_SATT309 region (Njiti et al., 1998; Meksem et al., 2001) were evaluated in two environments (Desoto and Ridgway) in 1997 in order to evaluate the association between yield and the SCN resistance in a homogenous genetic background. The NILs were previously screened for SCN resistance (Meksem et al., 1999).

In 1998, the F×H population was grown in the field at Ridgway on Reesville silt loam soil (Fine silty, mixed, superactive, mesic, aeric, Epiaqualfs) and Nashville, IL, on Bonnie silt loam soil (fine-silty mix, acid, mesic Typic Fluvaquents). In 1999, the RILs were planted in Harrisburg, IL, on Patton silty clay loam soil (Fine-silty, mixed, superactive, mesic Typic Endoaquolls), and Nashville on Bonnie silt loam soil (fine-silty mix, acid, mesic Typic Fluvaquents). There was some water deficit stress in 1999. All test locations were free of disease and had nondetectable soybean cyst nematode count.

A pre-plant herbicide combination of 1.12 kg a.i. ha<sup>-1</sup> glyphosate and 0.92 kg a.i. ha<sup>-1</sup> pendimethalin was used to treat the plots. Post-emergence grasses were controlled with a herbicide combination of 0.52 kg a.i. ha<sup>-1</sup> clethodim and 0.28 kg a.i. ha<sup>-1</sup> fomesafen. Broadleaf weeds were controlled by hand hoeing.

Experiments were planted with a four-row cone planter in no-till plots between 22 and 31 May in 1996 and 1997 and between 1 May and 31 May in 1998 and 1999. Plots consisted of four rows, 6.1 m long, 0.75 m apart, and were arranged in a randomized complete block design in three replications (Carbondale had two replications). The middle two rows of each plot were harvested with a two-row combine. The seeds were cleaned without losing broken seeds. Yield (kg ha<sup>-1</sup>) was calculated for each plot at 13% (w/v) moisture content. Before harvest, data were also collected on plant height, lodging, growth habit, and days to maturity.

### DNA Isolation

The RILs were grown in the greenhouse, 3 grams of leaves were collected from 5–6 two week old seedlings and immediately frozen in liquid nitrogen. The leaves were ground in liquid nitrogen into a very fine powder and DNA was extracted after Paterson et al. (1991). DNA concentration was measured

by a fluorometer and diluted to 15 ng/μl for further use in PCR reactions.

### Microsatellite Amplifications

Microsatellites markers from all 20 linkage groups were selected at 25 cM intervals from the soybean genetic map (Cregan et al., 1999). The primer pairs were purchased from Research Genetics, Inc. (Huntsville, AL). Amplifications were carried out in a PE 9600 thermal cycler as described by Akkaya et al. (1995). Two negative controls (with no template DNA) along with the two parents DNA as positive controls were run in all the amplifications. After PCR, the amplification products were electrophoresed on 5% (w/v) acrylamide gel that was used to expose X-ray film. The recombinant inbred lines were classified by parental genotypes for each marker.

### Construction of Pools

In the F×H RIL population, DNA from the 10 highest yielding RILs was pooled to form the high yield pool while DNA from the 10 lowest yielding RILs was pooled to form the low yield pool (Meksem et al., 2001). The pools were used for bulk segregant analysis (Mansur et al., 1996) of yield using microsatellite markers that have been found to be associated with yield in the E×F RIL population. Markers that were polymorphic between high and low yield pools were used to score the parental alleles among individual RILs.

### Data Analysis

Analysis of variance (ANOVA; SAS Institute Inc., Cary, NC) for yield was conducted across locations. Variance components were estimated. The heritability of yield was determined from variance components (Nyquist, 1991) as

$$h^2 = \sigma_g^2 / (\sigma_g^2 / r + \sigma_e^2 / l + \sigma_{gl}^2)$$

where  $h^2$  = heritability;  $\sigma_g^2$  = genotypic variance;  $\sigma_e^2$  = error variance;  $\sigma_{gl}^2$  = genotype × location variance;  $r$  = number of replications, and  $l$  = number of locations. To detect genomic regions associated with yield, marker data were compared with trait data by one way analysis of variance. The probability of association of each marker with yield was determined and a significant association was declared if  $P \leq 0.009$ . Loci influencing yield were examined for association with plant height, lodging, days to maturity, and disease (SCN and SDS) resistance. The relationship between yield and growth habit was determined by a one-way ANOVA. The univariate procedure was used to test for normality of yield distribution among RILs.

### Mapping Quantitative Trait Loci

A linkage map was created using MAPMAKER-EXP 3.0 (Lander et al., 1987). Map distances between linked markers were calculated in centimorgans (cM) to construct a linkage map (heterogenous lines were excluded). The recombinant inbred line (ri-self) genetic model was used. The log<sub>10</sub> of the odds ratio (LOD) for grouping markers was set at 2.0, maximum distance was 30 cM. Conflicts were resolved in favor of the highest LOD score after checking the raw data for errors. Marker order within groups was determined by comparing the likelihood of many map orders. A maximum likelihood map was computed with error detection. Groups were assigned to linkage groups by anchored microsatellite markers (Shoemaker and Specht, 1995; Cregan et al., 1999).

The map and yield data were simultaneously analyzed with Mapmaker/QTL 1.1 (Paterson et al., 1991) using the F<sub>2</sub>-back-

cross genetic model for trait segregation (Webb et al., 1995; Hnetkovsky et al., 1996; Chang et al., 1996, 1997). Quantitative trait loci were inferred when LOD scores exceeded 2.0 at some point in each interval since this was found empirically to be equivalent to a single marker  $P < 0.005$ , the criterion used in one-way ANOVA. The positions of the QTL were inferred from the interval peak LOD score.

The microsatellite markers used in this study have been mapped (Cregan et al. 1999) in other soybean populations. Therefore, markers were anchored on the linkage groups on the basis of their known locations. Anchored markers are important for the comparison of both known QTL locations from several populations (Mansur et al., 1996) and positions of dominant markers from earlier studies (Njiti et al., 1997).

## RESULTS

### Essex × Forrest Population

The mean yield across four environments among 100 recombinant inbred lines had a distribution that was normal ( $P > 0.05$ ), continuous and unimodal (Fig. 1a). The heritability of yield on a line mean basis among the recombinant inbred lines was 47% across four environments and ranged from 25% to 50% within specific environment. About 48% of SCN resistant progeny had better yield than the SCN resistant parent. However, none yielded as well as or better than the SCN susceptible parent (Fig. 1a). This population did not segregate for growth habit.

### Polymorphism and Linkage

A total of 500 microsatellite markers covering all the 20 linkage groups of soybean were tested for polymorphism between Essex and Forrest. One hundred thirty-five microsatellite markers were polymorphic between the two parents and scored in the RIL population. One hundred seven markers were found to be linked, representing 18 known linkage groups. On average 6 markers were placed on each linkage group. The actual number of markers ranged from 3 for linkage groups C1, D1a, and J to 16 on G. The total map encompassed 2823.1 cM with an average of 26.4 cM between loci excluding the 27 unlinked markers. This coverage is comparable to the known recombination distance of about 3000 cM encompassing 20 linkage groups of soybean (Cregan et al., 1999).

### Genomic Regions Associated with Yield

Three chromosomal regions on three different linkage groups were found to contain quantitative trait loci (QTL) that influence soybean yield in this population. One region derived the beneficial allele from Essex (high yielding parent) and two regions derived the beneficial allele from Forrest (low yielding parent). The regions jointly explained about 22% of the total variation in mean soybean yield in this population.

A region on linkage group K identified by BARC\_SATT337 was significantly ( $P = 0.0042$ ,  $R^2 = 10\%$ ) associated with mean yield across four environments and derived the beneficial allele from Essex. The linked marker BARC\_SATT167 also was significantly ( $P =$

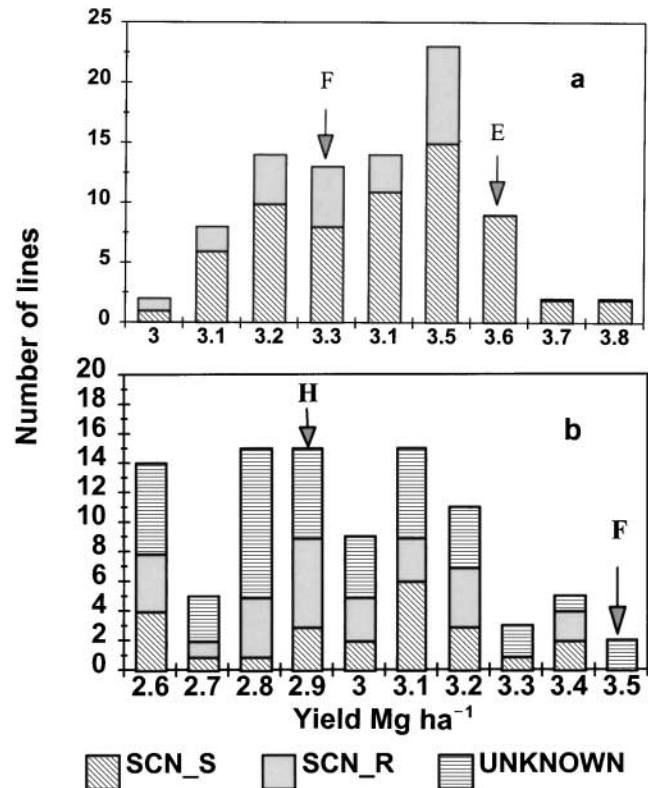


Fig. 1. (a) Frequency distribution of mean yield among Essex × Forrest recombinant inbred lines grouped by SCN resistance. F and E directed by arrows indicate where the mean yield of Forrest and Essex respectively fall within the distribution. (b) Frequency distribution of mean yield among Essex × Hartwig recombinant inbred lines grouped by SCN resistance. Letters H and F directed by arrows indicate where the mean yield of Hartwig and Essex respectively fall within the distribution.

0.0023,  $R^2 = 12\%$ ) associated with mean yield. The interval containing the QTL spanned about 5 cM and had a peak LOD score of 2.2 (Fig. 2a) and explained about 12% of variation in mean yield (Table 1).

A region on linkage group C1 identified by BARC\_SATT294 was significantly ( $P = 0.0056$ ,  $R^2 = 10\%$ ) associated with mean yield across four environments and derived the beneficial allele from Forrest. The linked markers in this region (BARC-SATT476, 2 cM away and BARC-195, 15 cM away) were not associated with yield in this population. The interval containing the QTL spanned about 12 cM had a peak LOD score of 2.3 (Fig. 2b) and explained about 12% of variation in mean yield (Table 1). This region was only associated with yield in one of four environments.

A region on linkage group I identified by BARC\_SATT440 was significantly ( $P = 0.0071$ ,  $R^2 = 10\%$ ) associated with mean yield across four environments and derived the beneficial allele from Forrest. The linked marker in this region (SAT\_105, 15 cM away) was not associated with yield in this population. The interval containing the QTL spanned about 10 cM, had a peak LOD score of 2.3 (Fig. 2b) and explained about 10% of variation in mean yield (Table 1). This region was only associated with yield in one of four environments.

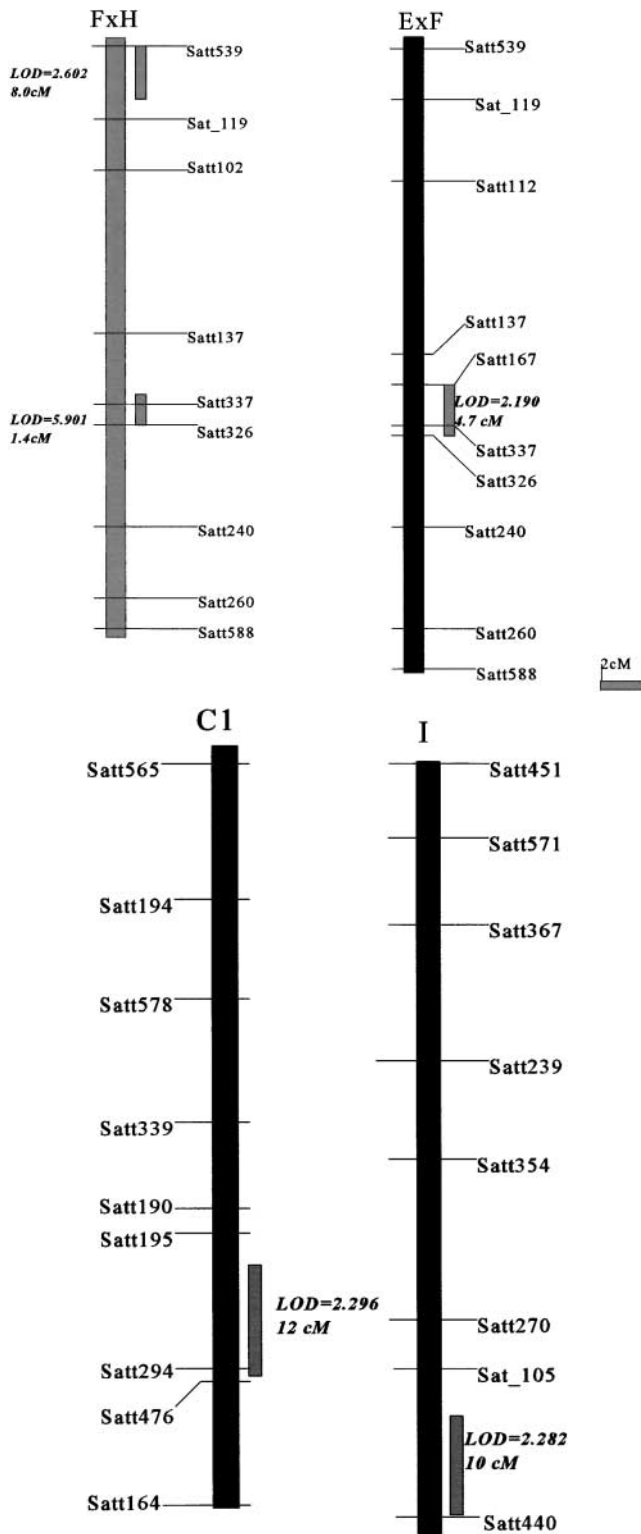


Fig. 2. (a) Flyer  $\times$  Hartwig and Essex  $\times$  Forrest maps for linkage group K. Yield QTL position, LOD scores and map distances are indicated. (b) Essex  $\times$  Forrest map for linkage groups C1 and I showing yield QTL position, LOD scores and map distances.

The SCN resistance loci reported (Meksem et al., 2001) were not significantly ( $P \leq 0.009$ ) associated with yield in environments that were not infested by *Heterodera glycine*. Molecular marker BARC\_SATT309 that

maps about 0.5 cM from *rhg1* was not significantly associated with yield among the RILs (Table 1). The Forrest (beneficial) allele for SCN resistance had a slightly higher mean yield than the Essex allele. Among near-isogenic lines ( $F_{5:10}$ -derived from E $\times$ F #11;  $n = 40$ ; Njiti et al., 1998) that contrasted for SCN resistance, the SCN resistance locus identified by BARC-Satt309 did not influence yield ( $P = 0.12$ ,  $R^2 = 7\%$ ). Yield means were  $3.2 \pm 0.05$  and  $3.06 \pm 0.08$  Mg ha $^{-1}$  for lines with alleles from Forrest ( $n = 25$ ) and Essex ( $n = 10$ ), respectively. The test was not sufficiently sensitive to detect small differences ( $<0.2$  Mg/ha).

The six SDS resistance loci previously reported (Hnetkovsky et al., 1996; Chang et al., 1996, 1997; Iqbal et al., 2001) were not significantly ( $P \leq 0.009$ ) associated with yield in environments that were not infested by *Fusarium solani* f. sp. *glycines* (Roy, 1997).

A QTL for plant height was detected at the BARC-Satt440 locus on LG. I. The QTL derived the beneficial allele from Essex. No other QTL for agronomic traits were detected at the loci influencing yield in this population.

### Flyer $\times$ Hartwig Population

The mean yield across four environments among 94 recombinant inbred lines had a distribution that was normal ( $P > 0.05$ ), continuous and unimodal (Fig. 1b). The heritability of yield on a line mean basis among the recombinant inbred lines was 57% across four environments and ranged from 72 to 89% within specific environment. About 45% of SCN resistant progeny had higher yields than the SCN resistant parent. However, none yielded as well as or better than the SCN susceptible parent (Fig. 1b).

There was no significant association ( $P = 0.06$ ,  $R^2 = 4.0\%$ ) between yield and growth habit. Yield means were  $2.92 \pm 0.03$  and  $2.79 \pm 0.07$  Mg ha $^{-1}$  for indeterminate ( $n = 72$ ) and determinate ( $n = 22$ ) lines, respectively.

### Polymorphism and Linkage

The 136 DNA markers polymorphic in E $\times$ F were used in bulk segregant analysis of yield in F $\times$ H. A marker (BARC\_SATT326) on linkage group K was polymorphic between high and low yield pools ( $n = 10$ ) (Fig. 3). Thirty-two DNA markers from linkage group K were then tested for polymorphism between the parents. Nine were found that were polymorphic between the parents and scored in the RIL population. Linkage group K has previously been shown to contain a major QTL for yield in the E $\times$ F RIL population (Njiti et al., 1997).

### Genomic Regions Associated with Yield

Two regions on linkage group K were found to contain QTL that influence soybean yield in this population. One region derived the beneficial allele from Flyer (high yielding parent) and the other region derived the beneficial allele from Hartwig (low yielding parent). The re-

**Table 1.** Markers associated with yield among recombinant inbred lines in two soybean populations. Essex  $\times$  Forrest was evaluated in Carbondale, IL (C96) and Ridgway, IL (R96) in 1996 and in Desoto, IL (D97) in 1997. Flyer  $\times$  Hartwig was evaluated in Nashville, IL (N98 and N99) in 1998 and 1999; Ridgway, IL (R98) in 1998 and Harrisburg, IL (H99) in 1999.

Marker (LG)	Location	P value	R <sup>2</sup>	LOD	QTL var	Yield mean $\pm$ SEM (Mg ha <sup>-1</sup> ) (alleles from)	
						P1	P2
<b>Essex <math>\times</math> Forrest Population</b>						<b>Essex</b>	<b>Forrest</b>
SATT337 (K)	C96	0.0328	5.8	1.9	13.7	2.88 $\pm$ 0.06	2.71 $\pm$ 0.06
	R96	0.0269	6.2	1.3	7.5	4.51 $\pm$ 0.05	4.38 $\pm$ 0.04
	D97	0.0197	6.9	1.6	10.0	3.02 $\pm$ 0.03	2.91 $\pm$ 0.03
	Mean <sup>†</sup>	0.0042	10.1	2.1	10.7	3.40 $\pm$ 0.03	3.29 $\pm$ 0.03
SATT167 (K)	D97	0.0047	10.0	1.5	7.7	3.01 $\pm$ 0.03	2.87 $\pm$ 0.04
	Mean	0.0023	11.6	1.9	9.6	3.39 $\pm$ 0.03	3.27 $\pm$ 0.03
SATT294 (C1)	D97	0.0002	16.4	3.7	23.7	2.88 $\pm$ 0.03	3.06 $\pm$ 0.03
	Mean	0.0056	9.5	2.0	12.1	3.29 $\pm$ 0.02	3.41 $\pm$ 0.03
SATT440 (I)	D97	0.0010	14.8	2.5	14.7	2.86 $\pm$ 0.03	3.02 $\pm$ 0.04
	Mean	0.0071	10.2	1.6	10.3	3.28 $\pm$ 0.03	3.39 $\pm$ 0.03
SATT309 <sup>‡</sup>	Mean	0.2964				3.32 $\pm$ 0.03	3.36 $\pm$ 0.03
<b>Flyer <math>\times</math> Hartwig population</b>						<b>Hartwig</b>	<b>Flyer</b>
SATT337 (K)	N98	0.0042	10.1	5.9	26.9	3.20 $\pm$ 0.06	2.69 $\pm$ 0.08
	N99	0.0001	27.1	1.2	6.4	2.53 $\pm$ 0.05	2.36 $\pm$ 0.04
	Mean	0.0006	13.8	2.7	13.7	2.98 $\pm$ 0.03	2.77 $\pm$ 0.05
SATT326 (K)	N98	0.0001	26.3	5.4	25.5	3.20 $\pm$ 0.06	2.69 $\pm$ 0.08
	N99	0.0082	8.6	1.6	7.8	2.52 $\pm$ 0.06	2.33 $\pm$ 0.04
	Mean	0.0004	15.0	3.0	14.6	2.98 $\pm$ 0.04	2.76 $\pm$ 0.05
SATT539 (K)	R98	0.0490	5.0	0.8	4.5	3.39 $\pm$ 0.09	3.62 $\pm$ 0.07
	H99	0.0006	14.7	2.6	13.4	2.45 $\pm$ 0.07	2.74 $\pm$ 0.04
	N99	0.0120	8.1	1.5	7.7	2.35 $\pm$ 0.05	2.54 $\pm$ 0.05
	Mean	0.0008	14.0	2.5	13.0	2.77 $\pm$ 0.06	2.99 $\pm$ 0.03
SATT038 <sup>‡</sup>	Mean	0.8216				2.89 $\pm$ 0.07	2.91 $\pm$ 0.05

<sup>†</sup> Mean involves all locations used for evaluation within each population.

<sup>‡</sup> Significantly associated with SCN resistance (close to *rhg1*).

gions jointly explained about 27% of the total variation in mean soybean yield in this population.

A region on linkage group K identified by BARC\_SATT337 was significantly ( $P = 0.0006$ ,  $R^2 = 14\%$ ) associated with mean yield across four environments and derived the beneficial allele from Flyer. The linked marker BARC\_SATT326 was also significantly ( $P = 0.0004$ ,  $R^2 = 15\%$ ) associated with mean yield. The interval containing the QTL spanned about 1.5 cM had a peak LOD score of 5.9 (Fig. 2a) and explained about 14% of variation in mean yield (Table 1). The region was associated with yield in two of four environments in this population.

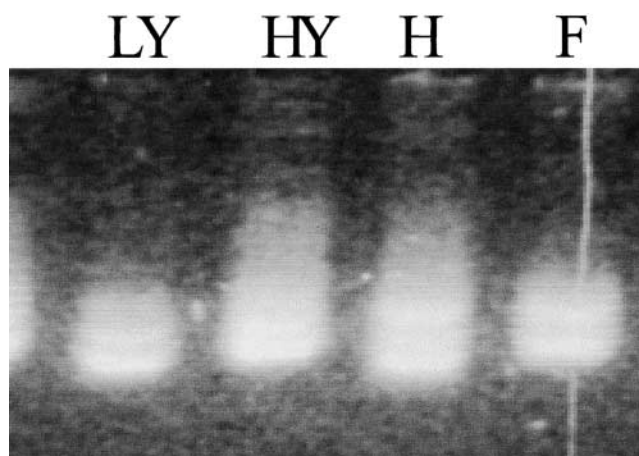
A region on linkage group K identified by BARC\_SATT539 was significantly ( $P = 0.0008$ ,  $R^2 = 14\%$ ) associated with mean yield across four environments and derived the beneficial allele from Hartwig. The linked marker (SAT\_119, 10 cM away) in this region was not significantly associated with mean yield. The interval containing the QTL spanned about 8 cM and had a peak LOD score of 2.6 (Fig. 2a) and explained about 13% of variation in mean yield (Table 1). The region was associated with yield in one of four environments in this population.

Yield in environments that were not infested by *Heterodera glycine* and *Fusarium solani* f. sp. *glycines*, was not significantly ( $P \leq 0.009$ ) associated with the six loci reported to be associated with SCN and SDS resistances (Hnetkovsky et al., 1996; Chang et al., 1996, 1997; Prabhu et al., 1999; Iqbal et al., 2001; Meksem et al., 2001). The molecular marker BARC-SATT038 that maps about 2.4 cM from *rhg1* was not significantly associated with yield in this population (Table 1).

A QTL for plant height ( $P = 0.007$ ,  $R^2 = 15\%$ ) was detected at the BARC-Satt337 locus on LG. K. The QTL derived the beneficial allele from Flyer, the high yielding parent. No other QTL for agronomic trait were detected at any of the loci influencing yield.

## DISCUSSION

Three QTL underlying mean yield were detected in the E $\times$ F and two in the F $\times$ H RIL populations. In each of the mapping populations, the total amount of variation in yield explained by the QTL was less than 50%



**Fig. 3.** Microsatellite marker Satt326 amplified from genomic DNA, ran on a 4% agarose gel and stained with ethidium Bromide. Bulk segregant analysis of high yield pool (HY) vs. low yield pool (LY). Flyer (F) is the high yield parent and Hartwig (H) is the low yield parent.

of the heritability of yield among RILs within the population. Therefore, additional QTL for yield may yet be detected in these populations.

In this study, BARC\_SATT337 identified a major QTL for yield in both E×F and F×H populations on molecular linkage group K. Interestingly, no other studies had detected yield QTL in this chromosomal location (Mansur et al., 1996; Orf et al., 1999). However, date of flowering, maturity, and protein content have been mapped close to this locus (Soybase). The appearance of the same QTL in two mapping populations suggests that the QTL may be consistent. While this QTL was effective in two of four environments in F×H, it was effective in means across four environments in each of the populations. Mean yield in both populations was enhanced, by 0.11 Mg/ha in E×F and 0.21 Mg/ha in F×H, indicating the value of the QTL. The mapping populations share a significant amount of their genome considering that Flyer is a BC<sub>3</sub>F<sub>2</sub> plant selection from the cross 'A3127'<sup>4</sup> × L24. The cultivar A3127 has Essex as a parent (Asgrow Seed Company, Des Moines, IA). Hartwig was derived from 'Forrest'<sup>3</sup> × PI437654. This is an indication that the QTL identified by BARC\_SATT337 may have identical alleles in these two populations. The same alleles may be over-represented in elite cultivars given the extent to which A3127 is used in breeding programs. Preliminary analysis of parents in the Southern Illinois University at Carbondale soybean breeding program indicated that alleles for the three markers linked to the yield QTL on LG. K, identified by BARC-SATT337 are segregating (unpublished data of the authors, 2001). It is not yet clear whether there is allelic disequilibrium (Meksem et al., 2001). The QTL will be useful in genome recovery during backcrossing.

In cultivated varieties of soybean, the basis of yield improvement is unclear. However, maturity, plant height, and growth habit have major effects on yield (Mansur et al., 1996; Orf et al., 1999). While the F×H population segregates for maturity, plant height, and growth habit, the E×F population segregates for plant height, has only a seven day spread in maturity and does not segregate for growth habit. In these populations, yield did not cosegregate with maturity or lodging. However, a QTL for maturity has been mapped close to BARC\_Satt337 in other populations (Soybase). Alleles in E×F and F×H may not be identical as epistasis may mask the effects of maturity QTL.

The absence of a detectable QTL for plant height at locus BARC-Satt337 in one of the two populations indicated that the QTL for plant height may be independent of the yield QTL rather than the QTL for plant height having a pleiotropic effect on yield. Although it is also possible that the test was not sufficiently powerful to pick up small but real differences or the difference may be masked by epistasis.

Soybean cyst nematode and SDS resistances have been reported to be associated with low yields (Rupe et al., 1993; Mudge et al., 1996; Concibido et al., 1997). It is not clear whether the yield reduction is due to gene(s) linked to disease resistance loci, a pleiotropic effect of disease resistance genes or genome wide re-

combination. Soybean yield depression by SCN resistance in the absence of SCN has been reported to probably be a result of a linked gene (Mansur et al., 1996; Concibido et al., 1997) within 10 cM of *rhg1*. However, in QTL analysis, this level of resolution does not exclude pleiotropy. In the present study, yield QTL were not detected at loci influencing disease resistance. However, the SCN resistance genes in the present population were derived from different sources (Peking and PI437 654) compared with the source of SCN resistance (PI209 332) in the population that was used to reveal the yield associated with SCN resistance (Mudge et al., 1996). Unlike the previous studies, the population used in this study was derived by intercrossing adapted susceptible × adapted resistance source. The adapted susceptible × non-adapted resistance source used in previous studies exhibited a lower combining ability for yield probably due to a large number of genes for low yield from the non-adapted parent. In addition, the introduction of a disease resistance gene to an adapted cultivar may have the effect of altering physiological processes that can influence yield. The lack of yield depression associated with disease resistance in the present populations may be attributed to selection against the yield depression genes or increased combining ability. Increased combining ability may be due to both selection against genes for low yield and adaptation of the genome to the disease resistance genes. Studies have shown that antibiotic resistance in bacteria resulted in a genome wide adaptation (Levin et al., 2000). Detailed analysis of yield and SCN resistance with recombinant near-isogenic lines will be necessary to distinguish these possibilities.

The proportion of SDS resistant RILs among the highest yielding RILs was high (>50%) in both populations suggesting that this yield QTL had a good combining ability with SDS resistance. There were more lines that were considered to be SDS resistant than SCN resistant. The SDS resistant high yielding lines lacked SCN resistance suggesting a yield drag associated with SCN resistance.

These results suggest that any yield depression associated with disease resistance (SCN and SDS) is more likely due to several genes unlinked to disease resistance loci than to a pleiotropic effect of disease resistance genes.

This study demonstrates specific loci influencing soybean yield can be identified via molecular markers. While few QTL were effective in a wide range of environments, some QTL had major effects on yield in specific environment and can be used in marker-assisted selection programs to develop cultivars in Southern Illinois. The applicability of these results may prove to be limited by both the sample size and number of test environments. Further analysis will test the ability of the QTL to predict yield in a wide selection of soybean germplasm.

#### ACKNOWLEDGMENTS

The authors acknowledge Dr. Paul Gibson and Dr. Oval Myers, Jr., Dep. of Plant, Soil, and General Agriculture

(PSGA), Southern Illinois University, for developing the recombinant inbred line populations used in this study. Special thanks to J.H. Klein, III and the rest of the soybean field crew in the Dep. of PSGA for participating in planting and managing plots. This research was supported by Illinois Soybean Program Operating Board (ISPOB) grants 95-222-24-3, 98-222-25-2, and NSF 9872635.

## REFERENCES

- Anand, S.C. 1992. Registration of 'Hartwig' soybean. *Crop Sci.* 32: 1069–1070.
- Akkaya, M.S., R. Shoemaker, J.E. Specht, A.A. Bhagwat, and P.B. Cregan. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Sci.* 35:1439–1445.
- Chang, S.J.C., T.W. Doubler, V. Kilo, R.J. Suttner, M.E. Schmidt, P.T. Gibson, and D.A. Lightfoot. 1996. Two additional loci underlying durable field resistance to soybean sudden death syndrome (SDS). *Crop Sci.* 36:1684–1688.
- Chang, S.J.C., T.W. Doubler, V. Kilo, R. Suttner, M.E. Schmidt, P.T. Gibson, and D.A. Lightfoot. 1997. Association of field resistance to soybean sudden death syndrome (SDS) and cyst nematode (SCN). *Crop Sci.* 37:965–971.
- Concibido, V., S. Boutin, R. Denny, R. Hautea, J. Orf, and N.D. Young. 1997. Genome mapping of a soybean cyst nematode resistance gene in Peking, PI 91763, and PI 88788 using DNA markers. *Crop Sci.* 37:258–264.
- Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean. *Crop Sci.* 39:1464–1490.
- Fehr, R.W. 1987. Principles of cultivar development. Vol. 1: Theory and technique. Macmillan, New York.
- Frary, A., T.C. Nesbitt, A. Frary, S. Grandillo, E. Van der Knaap, B. Cong, J. Liu, J. Meller, R. Elbert, K.B. Alpert, and S.D. Tanksley. 2000. A quantitative trait locus key to the evolution of tomato fruit size. *Science* 289:85.
- Hartwig, E.E., and J.M. Epps. 1973. Registration of 'Forrest' soybeans. *Crop Sci.* 13:287.
- Hnetkovsky, N., S.J.C. Chang, T.W. Doubler, P.T. Gibson, and D.A. Lightfoot. 1996. Genetic mapping of loci underlying field resistance to soybean sudden death syndrome (SDS). *Crop Sci.* 36:393–400.
- Iqbal, M.J., K. Meksem, V.N. Njiti, M. Kassem, and D.A. Lightfoot. 2000. Microsatellite markers identify three additional quantitative trait loci for resistance to soybean sudden death syndrome (SDS) in Essex × Forrest RILs. *Theor. Appl. Genet.* 102:187–192.
- Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124: 743–756.
- Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daley, S. Lincoln, and L. Newburg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181.
- Levin, B.R., V. Perrot, and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. *Genetics* 154:985–997.
- Mansur, L.M., K.G. Lark, H. Kross, and A. Oliveira. 1993. Interval mapping of quantitative trait loci for reproductive, morphological, and seed traits of soybean [*Glycine max* (L.) Merr.]. *Theor. Appl. Genet.* 86:907–913.
- Mansur, L.M., J.H. Orf, K. Chase, T. Jarvik, P.B. Cregan, and K.G. Lark. 1996. Genetic mapping of agronomic trait using recombinant inbred lines of soybean. *Crop Sci.* 36:1327–1336.
- Maughan, P.J., M.A. Saghai Maroof, and G.R. Buss. 1996. Molecular markers analysis of seed-weight: genomic locations, gene action, and evidence for orthologous evolution among three legume species. *Theor. Appl. Genet.* 93:574–579.
- McBlain, B.A., R.J. Fioritto, S.K. St. Martin, A. Calip-Dubois, A.F. Schmitthenner, R.L. Cooper, and R.J. Martin. 1990. Registration of 'Flyer' soybean. *Crop Sci.* 30:425.
- Meksem, K., T.W. Doubler, K. Chanchaoenchai, V. Njiti, S.J.C. Chang, A.P. Rao-Arelli, P. Cregan, L.E. Gray, P.T. Gibson, and D.A. Lightfoot. 1999. Clustering among loci underlying soybean resistance to *Fusarium solani*, SDS and SCN in near-isogenic lines. *Theor. Appl. Genet.* 99:1131–1142.
- Meksem, K., P. Pantazopoulos, V.N. Njiti, L.D. Hyten, P.R. Arelli, and D.A. Lightfoot. 2001. 'Forrest' resistance to the soybean cyst nematode is bigenic: Saturation mapping of the *rhg1* and *Rhg4* loci. *Theor. Appl. Genet.* 103:710–717.
- Mudge, J., V. Concibido, R. Denny, N. Young, and J. Orf. 1996. Genetic mapping of a yield depression locus near a major gene for soybean cyst nematode resistance. *Soybean Genetics Newsletter* 23:175–178.
- Njiti, V.N., M.A. Shenaut, R.J. Suttner, M.E. Schmidt, and P.T. Gibson. 1998. Relationship between soybean sudden death syndrome (SDS) and yield components in  $F_6$ -derived lines. *Crop Sci.* 38: 673–678.
- Njiti, V.N., C.A. Schmidt, M.E. Schmidt, and D.A. Lightfoot. 1997. Mapping loci underlying yield in Illinois. *Soybean Genetics Newsletter* 24:136–138.
- Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10(3): 235–322.
- Orf, J.H., K. Chase, F.R. Alder, L.M. Mansur, and K.G. Lark. 1999. Genetics of soybean agronomic traits: II. Interaction between yield quantitative trait loci in soybean. *Crop Sci.* 39:1652–1657.
- Paterson, A.H., S. Damon, D.J. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.S. Lander, and S.D. Tanksley. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics* 127:181–197.
- Prabhu, R.R., V.N. Njiti, B. Bell-Johnson, M.E. Schmidt, J.H. Klein, and D.A. Lightfoot. 1999. Selecting soybean cultivars for dual resistance to soybean cyst nematode and sudden death syndrome using two DNA markers. *Crop Sci.* 39:982–987.
- Roy, K.W. 1997. *Fusarium solani* on soybean roots: Nomenclature of causal agent of sudden death syndrome and identity and relevance of *F. solani* form B. *Plant Dis.* 81:259–266.
- Rupe, J.C.D., W.E. Sable, R.T. Robbins, and E.E. Gbur. 1993. Soil and plant factors associated with sudden death syndrome of soybean. *J. Prod. Agric.* 6:218–221.
- Shoemaker, R.C., and J.E. Specht. 1995. Integration of the soybean molecular and classical genetic linkage groups. *Crop Sci.* 35:436–446.
- Smith, T.J., and H.M. Camper. 1973. Registration of 'Essex' soybeans. *Crop Sci.* 13:495.
- Specht, J.E., D.J. Hume, and S.V. Kumudini. 1999. Soybean yield potential—A genetic and physiological perspective. *Crop Sci.* 39: 1560–1570.
- Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132:832–839.
- Webb, D.M., B.M. Baltazar, A.P. Rao-Arelli, J. Schupp, K. Clayton, P. Keim, and W.D. Beavis. 1995. Genetic mapping of soybean cyst nematode race-3 resistance loci in the soybean PI 437.654. *Theor. Appl. Genet.* 91:574–581.