Theor Appl Genet (1993) 86:914–918

Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean (*Glycine max* L. Merr.)

L. M. Mansur[1], J. Orf[2], K. G. Lark[1]

[1] Department of Biology, University of Utah, Salt Lake City, UT 84112, USA
[2] Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA

Received: 21 September 1992 / Accepted: 28 January 1993

Abstract. An experimental test is described for linkages between RFLP markers and quantitative trait loci (QTL). Two hundred and eighty-four F_7-derived recombinant inbred lines (RIL) obtained from crossing the soybean cultivars (*Glycine max* L. Merr.) 'Minsoy' and 'Noir 1' were evaluated for maturity, plant height, lodging, and seed yield. RIL exhibiting an extreme phenotype for each trait (earliest and latest plants for maturity, etc.) were selected, and two bulked DNA samples were prepared for each trait. A Southern transfer of the digested bulked DNA was hybridized with restriction fragement length polymorphism (RFLP) probes, and linkages with QTL were established by quantitating the amount of radioactive probe that bound to fragments defining alternative parental RFLP alleles. When an RFLP marker was linked to a QTL, one parental allele predominated in the bulked DNA from a particular phenotype; the other allele was associated with the opposite phenotype. When linkage was absent, radioactivity was associated equally with both alleles for a given phenotype (or with both phenotypes for a given allele). These results confirmed RFLP-QTL associations previously discovered by interval mapping on a smaller segregating population from the same cross. New linkages to QTL were also verified.

Key words: DNA – RFLP – QTL – Soybean –Recombinant inbreds

Communicated by A. L. Kahler
Correspondence to: K. G. Lark

Introduction

Recombinant inbred lines (RIL) represent a stable source of germ plasm, comprising different combinations of genetic material derived from their parents (Bailey 1971; Burr et al. 1988). In previous papers (Lark et al. 1993; Mansur et al. 1993), we presented evidence that the two soybean (*Glycine max* L. Merr.) cultivars 'Minsoy' and 'Noir 1' contain different alleles of quantitative trait loci (QTL) that produce transgressive variation in their segregating progeny. We have prepared 284 RIL from this cross that are now in the F_{11} generation and have used this RIL population to test a technique that significantly improves the efficiency of QTL mapping with restriction fragment length polymorphism (RFLP) markers.

F_7-derived RILs in the F_9 amd F_{10} generations were evaluated for quantitative traits of agronomic interest, and those exhibiting the most extreme phenotypes were selected (e.g., shortest and tallest for height, early and late for maturity, etc.). These RILs were used to determine whether RFLP markers can by linked to QTL using the concept of selective genotyping (Lander and Botstein 1989) coupled to a method similar to that employed by Michelmore et al. (1991) for qualitative traits. In this approach, DNA from individuals of extreme phenotype (e.g., tallest) are combined and compared to similarly bulked DNA from the opposite phenotype (e.g., shortest). If an RFLP marker is not linked to a QTL, both parental RFLP alleles should be represented in the bulked DNA. If an RFLP marker is linked, then the parental allele linked to the selected phenotype should predominate. We have confirmed that this methodology works using data for maturity, plant height, lodging, and seed yield.

Materials and methods

Genetic materials

Near-inbred recombinant inbred lines (284 in total) were developed from the intraspecific cross of two soybean cultivars: PI 27.890 ('Minsoy') and PI 290.136 ('Noir 1'). Characteristics of these cultivars which suggest that they are genetically distant have been discussed in previous papers (Lark et al. 1993; Mansur et al. 1993). These RIL were developed as follows: a single seed was chosen at random from each plant in each generation, starting with the F_2, and carried forward, allowing plants to self pollinate in a greenhouse. At the F_7 generation each plant was threshed individually and maintained as a separate RIL. Seed of the F_7 RILs were multiplied in the field in Chile during the US winter of 1990. An F_9 bulk sample of seeds was used for field experiments in Minnesota during the summer of 1991, and F_{10} progeny seed were used in Chile during the winter of 1992.

Field experiments

The 284 RILs, each parent, and 12 check cultivars were arranged in a randomized complete block design with three replicates at each of two locations: the University of Minnesota field research station in Rosemount, Minnesota, and at La Platina experiment station of the National Institute for Agricultural Research (INIA) near Santiago, Chile. Each entry was planted in paired-row plots 2 m long, at 25 seeds per meter, with a spacing between rows of 67 cm at Rosemount and of 1 m at La Platina. The seeds were, planted May 25, 1991 and December 13, 1991 in Rosemount and La Platina, respectively. The traits measured included: (a) R_8, or maturity (Fehr and Caviness 1977); the number of days after July 31 at Minnesota and after February 29 at La Platina when 95% of the normal pods on the main stem had reached their mature pod color. Maturity values were expressed as the deviation in days from the mean of the maturity of the 12 check cultivars at each location; (b) plant height, the height in centimeters of the main stem measured from the soil surface to the stem tip; (c) lodging score, the average angle between the plants and the soil surface (with 1 being fully erect and 5 all plants prostrate); and (d) seed yield in gm/m^2.

Selection of extreme phenotypes

Means across locations were computed for each RIL. Those falling in the extreme 7% of the upper and lower end of the range of values (20 plants of each extreme) were chosen for selective genotyping. Table 1 lists the means of each group of extreme phenotypic plants, together with the mean of the entire RIL population.

Selective genotyping using bulk DNA of extreme phenotypes

Following the method described by (Michelmore et al. 1991), DNA from extreme plants was bulked according to each phenotype. The techniques used were essentially the same as those described previously (Lark et al. 1993). In brief: 0.5 µg DNA from each of the 20 extreme high and low plants was pooled to form two samples, 10 µg per sample, for each trait. Ten micrograms of DNA of 'Minsoy' and of 'Noir 1' were used as controls. Each sample was digested with restriction enzyme (according to the specification of the supplier), and the resulting fragments were separated using agarose gel electrophoresis and transferred to a nylon membrane. The resulting blot was hybridized with a radioactive probe and exposed to produce autoradiographs and also scanned in the phosphoimager (Fig. 1). The probes R79, A397, and A60 were chosen because previous experiments (Mansur et al. 1993) had shown they were linked to QTL for at least one of the four traits used in this experiment.

Hybridized radioactivity was quantitated using a gel blot analyzer (Phosphoimager, Molecular Dynamics model 400). To compensate for possible differences in the concentration of DNA in each lane, the radioactivity of bands corresponding to alleles was divided by the radioactivity of a common fragment, present in all of the lanes. Figure 1 presents a picture of the phosphoimager screen for probe A60. The analyzed image is shown on the left; the entire image at the upper right. Numbers have been added to each lane of the image for clarity. Lanes 5 and 6 (plant height) are analyzed in the profiles to the right. Intensity profiles from left to right correspond to radioactive bands from the top (large fragments) to the bottom (small fragments) of the gel.

On the extreme right of the two profiles are two bands common to both 'Minsoy' and 'Noir 1'. Immediately to the left of these are a pair of bands representing the polymorphism, A60b, in which the 'Minsoy' allele is bisected by a line. (Polymorphisms corresponding to A60a are seen to the left of the profiles.) The profiles as shown in Fig. 1 are for purposes of illustration. In practice, the profiles were extended to the width of the screen when separating the peaks to facilitate analysis.

A rectangular portion of each lane was analyzed and the radioactivity profile prepared. Radioactivity in each band was

Table 1. Mean and range of four traits for parents and for the 20 (7%) extreme phenotypes (high and low) selected from recombinant inbred lines

		Trait			
		Maturity (days)	Plant height (cm)	Lodging (score 1–5)	Yield (gm/m^2)
Minsoy	Mean	−9	56	2.9	227
	Maximum	−5	58	3.5	257
	Minimum	−10	53	2.0	196
Noir 1	Mean	−11	82	2.5	241
	Maximum	−7	92	3.0	306
	Minimum	−13	77	2.0	181
Extreme high	Mean	7	108	4.0	304
	Maximum	11	118	4.8	325
	Minimum	5	100	3.7	294
Extreme low	Mean	−15	36	1.3	102
	Maximum	−13	43	1.4	133
	Minimum	−20	24	1.0	41

computed as the area under a peak. Values of polymorphic bands were divided by the value of a designated common band, and these ratios were compared as the percentage radioactivity in each sample of bulked DNA. For example, in the profile in Fig. 1, the peaks designated by the line were compared as A60b 'Minsoy' alleles for the extremes of plant height. Each area was divided by the area of the common band to the right (↑). Subsequently, the two corrected areas were summed, giving the total for both extremes, and each extreme was expressed as the percentage of this total.

Results

The extreme phenotypes, selected from a population of 284 recombinant inbred lines, represent substantial deviations from the mean of the total population in their trait values (Table 1). For example, 'Minsoy' and 'Noir 1' differ in maturity by 2 days (−9 versus −11 days, Table 1), whereas the late maturity extreme is 22 days later than the early (+7 versus −15). Bulked DNA from these extreme phenotype populations is thus pooled from a highly selected group of plants that were grown at two distinct environments.

R79 is a locus which shows a high log of likelihood (LOD) score for linkage to QTL for maturity and yield and low LOD scores for linkage to QTL for lodging and plant height (Mansur et al. 1993 and Table 2). Figure 2 presents an autoradiogram of a Southern transfer of restriction fragments from extreme phenotype DNA hybridized to this probe. Maturity and yield show patterns typically expected of linkage. That is, early maturity (lane 3) or low yield (lane 9) show fragments corresponding to the 'Minsoy' allele (lane 1), whereas late maturity (lane 4) and high yield (lane 10) show fragments corresponding to the 'Noir 1' allele (lane 2). The quantitative distribution of radioactivities is shown in Table 2 together with the LOD scores from previous linkage data (Mansur et al. 1993).

Plant height shows a similar linkage to this locus, despite previous data in which a lower LOD score (ca. 1.0, Table 2) suggested a lack of such linkage. However, the quantitative radioactive data in Table 2 suggest that this linkage may be quite strong. This is consistent with the previous finding of a linkage between marker R79 and canopy height (see Table 2 of Mansur et al. 1993). Finally, lodging shows that both upright and prostrate plants are associated with both types of restriction fragments – consistent with the lack of linkage observed previously (LOD = 0.037, Table 2).

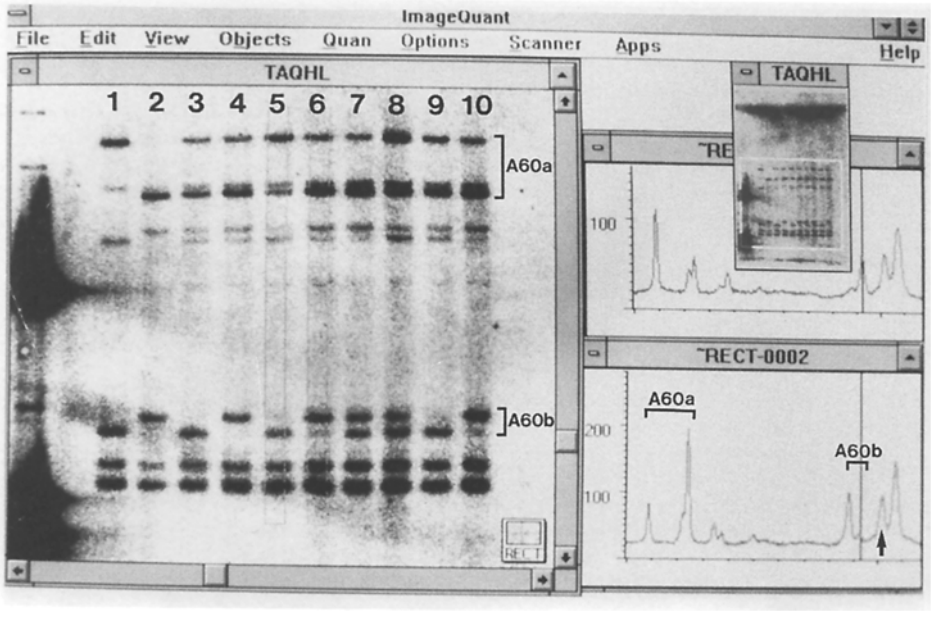

Fig. 1. Phosphoimager analysis of plant height using a *Taq*I restriction digest of bulked DNA from tall and short plants hybridized with the radioactive probe A60. Preparation of the Southern transfers and its hybridization with radioactive probe is described in the text (Materials and methods) or in a previous paper (Lark et al. 1993). The figure presents a photograph of the output on the imager screen: the *small inset* at the *upper right* (*small TAQHL*) is an image of the entire Southern transfer. The *enclosed rectangle* in that image is the area to be analyzed, the image of which is shown on the *left* (*large TAQHL*). In that image, the lanes are from the extreme *left*: a standard (1-kb ladder), *1* 'Minsoy' DNA, *2* 'Noir 1' DNA, *3* early maturity, *4* late maturity, *5* short plants, *6* tall plants, *7* upright plants, *8* prostrate plants, *9* low yield, *10* high yield. DNA was prepared from these extreme phenotypes, digested with *Taq*I, and combined for gel electrophoresis as described in the text. The two *rectrangular areas* in *lanes 5 and 6* have been analyzed to produce the radioactive intensity profiles on the *right* (*lane 5 at top, 6 at bottom*). The cursor line in these profiles has been placed over a fragment corresponding to the Minsoy allele for the A60b RFLP marker. Common fragments are to the *right of this line*. As noted in the text, these profiles can be enlarged to facilitate separation of the peaks

Table 2. Comparison of linkages obtained using extreme phenotypes of RIL plants with previous linkages obtained from interval mapping of an F_2 popuation

Allele	Maturity Early	Late	Plant height Short	Tall	Lodging Upright	Prostrate	Yield Low	High
R79 Minsoy	98	2	3	97	60	40	83	17
R79 Noir 1	12	88	100	0	54	46	5	95
Probability of linkage[a]	LOD = 5.1		LOD = 1.1		LOD = 0.0		LOD = 3.2	
A60a Minsoy	45	55	30	70	32	68	60	40
A60a Noir 1	38	62	74	26	59	41	40	60
Probability of no linkage[b]	$P < 0.01$		$P < 0.004$		$P < 0.28$		$P < 0.2$	
A60b Minsoy	100	0	0	100	55	45	87	13
A60b Noir 1	0	100	100	0	45	55	10	90
Probability of linkage[a]	LOD = 3.3		LOD = 0.58		LOD = 0.0		LOD = 2.4	
A397 Noir 1	100	0	8	92	15	85	73	27
Probability of linkage[a]	LOD = 2.5		LOD = 2.2		LOD = 0.1		LOD = 2.4	
Probability of no linkage[c]	$P < 0.001$		$P < 0.001$		$P < 0.001$		$P < 0.005$	

[a] LOD score from interval mapping of between 56 and 69 segregating F_2 plants (see Table 2, Mansur et al. 1993)
[b] Probability values for lack of linkage obtained from analysis of variance comparing means of 69 $F_{2:5}$ families grouped by RFLP allele genotype of the F_2 plants (see Mansur et al. 1993)
[c] Probability values for lack of linkage obtained from analysis of variance comparing means of 237 F_7-derived recombinant inbred lines from this experiment

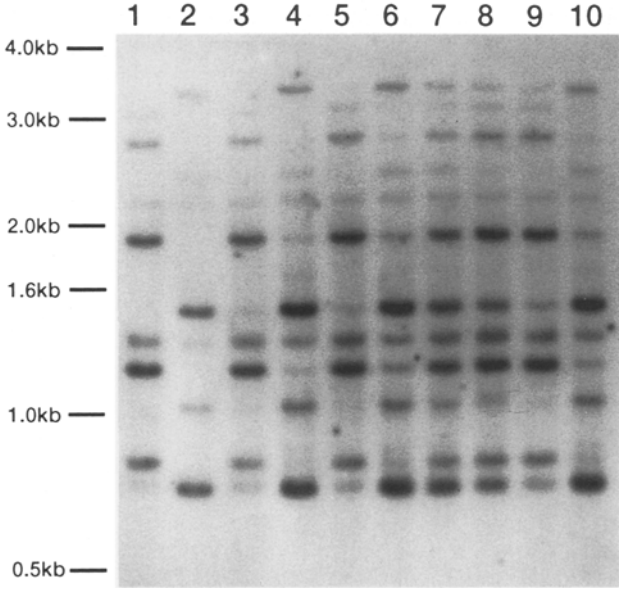

Fig. 2. Autoradiogram of the Southern transfer in Fig. 1 after hybridization with the radiocative probe R79. As in Fig. 1, *from left to right* the lanes are: *1* 'Minsoy' DNA, *2* 'Noir 1' DNA, *3* early maturity, *4* late maturity, *5* short plants, *6* tall plants, *7* upright plants, *8* prostrate plants, *9* low yield, *10* high yield

A60a is unlinked to other qualitative genetic markers and therefore not located on the linkage map (Lark et al. 1993). This marker was previously shown, however, to be linked to a QTL for plant height (Mansur et al. 1993 and Table 2). This is confirmed in Fig. 1. Taller plants are associated with the 'Minsoy' A60a allele, whereas shorter plants are associated with the 'Noir 1' allele (lanes 5 and 6, respectively). The data for lodging (lanes 7 and 8 and Table 2) are not so compelling, quantitative measurements showing a bias for the 'Minsoy' allele but *not* for the 'Noir 1' allele. This is consistent with previous linkage data. Possible linkages to maturity and yield (Table 2) are similarly not compelling, and are also consistent with previous data. It should be noted that, as expected, quantitative measurements with the phosphoimager were more reliable than those estimated from the autoradiograms. The film, which does not respond linearly to the amount of radioactivity present, appeared to indicate a stronger bias for the 'Minsoy' allele (data not shown).

The image in Fig. 1 shows the presence of another RFLP marker previously not detected with probe A60. This marker appears to be unlinked to the original A60a locus based on linkage or lack of linkage to QTL for the four traits. Data for these traits is also presented in Table 2. This locus has now been mapped and is linked to marker R79 in linkage group 15 of Lark et al. (1993) (11.7 cM from R79 distal to A584). This location explains the similarity of QTL linkages to those for R79. The LOD scores for linkage of the four traits to A60b were also determined using the smaller segregant population and are presented in Table 2.

Data for marker A397 (Table 2) further demonstrated the power of this technique. Previous data had shown marginal or low LOD scores for the linkage of the four traits to this locus. Because of the complexity of the autoradiographs, it was only possible to score one of the alleles quantitatively (the 'Noir 1' allele). ['Minsoy' alleles were too close to common fragments (present in both parents) to allow accurate quantitation.] However, the data clearly suggested linkage not only to a QTL for maturity (as predicted from previous data), but also for height and lodging. Previous LOD scores for three loci suggested that height might be

linked, but that lodging was not. Because the RIL population is much larger than the previous population, we could measure the linkage of QTL for these traits against 237 plants that had already been genotyped with this probe. These statistics are also shown in Table 2; they confirm the extreme phenotype data, but not the previous LOD scores based on the smaller plant population.

Discussion

Recombinant inbred lines of soybean provide a variety of homozygous (and therefore replicable) genotypes composed of different mixtures of parental alleles. Genetic diversity is insured by having a large number of lines derived from parents containing different alleles at a large number of loci. In a previous paper, we demonstrated that 'Minsoy' and 'Noir 1' can provide this diversity for traits of agronomic interest (Table 1 in Mansur et al. 1993). From this cross we have derived 284 recombinant inbred lines. Ideally, we would like to determine the genotype of each of these lines, and we are at the present time in the process of preparing a genetic description of each line in order to comprise a framework of suitably spaced RFLP markers.

In this paper we have demonstrated that by focusing on extreme phenotypes (Lander and Bottstein 1989), such a population can be surveyed rapidly for RFLP markers linked to traits of interest. The method clearly confirms previously predicted linkages based on LOD scores greater than 2.5 or $P < 0.005$ (markers A79 and A60a, Table 2). Lower LOD scores (higher P values) were clarified by this approach: there was no evidence for a linkage of maturity, lodging, or yield to A60a or of lodging to R79, whereas plant height was linked to R79, and all four traits were linked to A397. This last observation was surprising in view of the previous LOD scores derived from the smaller segregating population (Mansur et al. 1993), but was confirmed by using this probe to examine the majority (237) of RIL plants.

An unexpected result was the linkage of maturity, plant height, and yield to marker A60b, a polymorphism not detected in our previous work (due to a bad signal to noise ratio in the earlier Southern blots). This data predicted a separate map location for A60b, which was confirmed by analyzing the smaller segregant population for this marker.

There are two distinct advantages to our methodology. The first is that it is very efficient and that the efficiency increases as more traits are surveyed. Thus, for each trait, only two restriction digests (two gel lanes) need to be analyzed. A dozen different traits can be easily analyzed on a single Southern blot. As we have seen, a great deal of information can be obtained from a single blot, such as with probes A60 or A397. A second advantage is that if a phenotype can be defined in such a manner that extremes can be identified, then linkage can be determined. This is a distinct advantage where complex traits, often difficult to quantitate precisely, are concerned (for example, taste of a processed product).

We used the phosphoimager in these studies for two reasons. We felt that the non-linear response of enhanced autoradiograms might lead to misinterpretations of the data. Comparison of the autoradiogram of A60a with its phosphoimage indicated that this supposition might be true in that the autoradiogram indicated a possible linkage for lodging, whereas the phosphoimager showed that this data (Table 2, A60a) was equivocal and certainly no better than the data for yield, which appeared to be unlinked in the autoradiogram. It appears that autoradiograms may be adequate for tightly linked markers but that more quantitation is useful as markers become less tightly linked (i.e., values of 70% versus 30% for the two extreme phenotypes as seen for plant height linked to marker A60a). More importantly, we had hoped that quantitation would yield an approximation of the degree of linkage or of the statistical probability of linkage. We were unable to demonstrate this with the data available from the smaller segregant population (Table 2 and Mansur et al. 1993). However, the results for A397 using the larger RIL population suggest that it may be possible in the future to establish the validity of such a relationship using a larger population.

Acknowledgments. We thank Kevin Chase and Pete Wilhelm for technical assistance. This research was supported by the American Soybean Association, by Grant GM42337 to K. G. L. from the NIH, and funds from the Minnesota Soybean Research and Promotion Council.

References

Bailey DW (1971) Recombinant-inbred strains. Transplantation 11:325–327

Burr B, Burr VK, Thompson H, Albertson MC, Stuber CW (1988) Gene mapping with recombinant inbreds in maize. Genetics 118:519–526

Fehr WR, Caviness CE (1977) Stages of soybean development. Iowa Agric Home Econ Exp Stn, Iowa Coop Ext Serv Spec Rep 80

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

Lark KG, Weissemann JM, Mathews BF, Palmer R, Chase K, Macalma T (1993) A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir 1'. Theor Appl Genet (in press)

Mansur LM, Lark KG, Kross H, and Oliveria A (1993) Interval mapping of quantitative trait loci for reproductive, morphological and seed traits of soybean (*Glycine max* L.). Theor Appl Genet (in press)

Michelmore RW, Paran I, Kesseli V (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci 88:9828–9832