Theor Appl Genet (1993) 86:901–906

A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minosy' and 'Noir 1'

K. G. Lark[1], J. M. Weisemann[2], B. F. Matthews[2], R. Palmer[3,4], K. Chase[1], T. Macalma[1]

[1] Department of Biology, University of Utah, Salt Lake City, UT 84112, USA
[2] USDA-Agricultural Research Service, Beltsville, MD 20705, USA
[3] Department of Agronomy, Iowa State University, Ames, IA 50011, USA
[4] USDA-Agricultural Research Service, Field Crops Research Unit, Iowa State University, Ames, IA 50011, USA

Received: 21 September 1992 / Accepted: 28 January 1993

Abstract. Genetic markers were mapped in segregating progeny from a cross between two soybean (*Glycine max* (L.) Merr.) cultivars: 'Minsoy' (PI 27.890) and 'Noir 1' (PI 290.136). A genetic linkage map was constructed (LOD $\geqslant 3$), consisting of 132 RFLP, isozyme, morphological, and biochemical markers. The map defined 1550 cM of the soybean genome comprising 31 linkage groups. An additional 24 polymorphic markers remained unlinked. A family of RFLP markers, identified by a single probe (hybridizing to an interspersed repeated DNA sequence), extended the map, linking other markers and defining regions for which other markers were not available.

Key words: DNA – RFLP – Soybean – Genetic map

Introduction

The development of restriction fragment length polymorphism (RFLP) molecular markers (Botstein et al. 1980) has facilitated the mapping of plant and animal genomes. In plants, the extensive genetic variation found with RFLP markers has led to the development of many comprehensive genetic linkage maps, including ones in maize (Helentjaris et al. 1986), tomato (Tanksley et al. 1987), lettuce (Landry et al. 1987), soybean [*Glycine max* (L.) Merrill] (Keim et al. 1990; Diers et al. 1992), and common bean (*Phaseolus vulgaris*) (Vallejos et al. 1992).

Prior to the use of RFLP markers, only 17 linkage groups had been identified in soybean, comprising 57 linked markers that covered approximately 420 centiMorgans (cM) (Palmer and Kilen 1987). In contrast, Keim et al. (1990) constructed a genetic map that included 130 RFLP markers in 26 linkage groups that covered approximately 1200 cM; this was later expanded to 252 markers in 31 linkage groups covering 2147 cM (Diers et al. 1992).

Genetic analysis of progeny from parents of similar phenotypes but different genotypes can uncover alternative genetic solutions to the same agronomic problem. For such an analysis, it is necessary to evaluate transgressively segregating progeny. We have crossed two soybean cultivars of similar phenotype, 'Minsoy' (PI 27.890) and 'Noir 1' (PI 290.136) whose progeny show a high degree of transgressive variation (Mansur et al. 1993a) thereby yielding many offspring that surpass either parent for various reproductive, morphological, and seed traits, as well as yield.

In this paper, we present the genetic map obtained from this cross. In subsequent papers we analyze the segregation of quantitative traits in this cross (Mansur et al. 1993a) and the use of recombinant inbred lines to link quantitative traits to RFLP markers (Mansur et al. 1993b).

Materials and methods

An intraspecific cross of two soybean cultivars, 'Minsoy' (PI 27.890) and 'Noir 1' (PI 290.136), was used for this mapping study. These cultivars have been shown to have a high frequency of variation in RFLP and morphological alleles (Keim et al. 1988; Palmer and Kilen 1987), which suggests that they also might be distant with respect to quantitative traits.

Genetic markers

RFLP markers were identified using soybean DNA cloned into plasmids (A, BLT, C, K, L, R, or T); using DNA cloned into

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

phage λ and then subcloned into plasmids; or using DNA cloned into phage M-13. Plasmid cloned DNA was obtained as a *Pst*I library of soybean genomic DNA (Keim and Shoemaker 1988) or, in some cases, as a cDNA library (Kalinski et al. 1990). The M-13 library was constructed from a *Sau*3a digest of genomic DNA and was prescreened for single-copy DNA (Apuya et al. 1988). Probes derived from λ clones of soybean DNA also have been described (Apuya et al. 1988).

A and K clones were a gift from Dr. R. C. Shoemaker, Iowa State University. BLT clones were isolated and identified as being polymorphic in Beltsville, Maryland. Clones C, L, R, T, M, and G were isolated in Utah. Morphological and biochemical markers which were analyzed on F_2 progeny in Iowa have already been described in the literature (Palmer and Kilen 1987). These include: w_1 (flower color), I (seed-coat color), *Pb* (pubescence tip), *Ep* (seed-coat peroxidase), and *R* (hilum color). In addition, three isozyme loci were scored using methods described by Cardy and Beversdorf (1984). These include aconitase (Aco_4), Malic enzyme (*ME*), and isocitrate dehydrogenase (Idh_2).

Measurement of RFLP markers

RFLP markers were revealed by hybridizing radioactive cloned DNA fragments to restriction fragments of genomic DNA previously separated according to size by agarose gel electrophoresis. The techniques used have been described in detail elsewhere (Sambrook et al. 1989; Roth et al. 1989). In brief, genomic DNA was isolated from freeze-dried leaves and digested with restriction enzymes according to the manufacturers' specifications. Fragments were separated by agarose gel electrophoresis. The separated fragments were transferred to membranes following the method of Southern (1975) and, after linking the DNA to the membrane by UV irradiation, the Southern transfers were hybridized with radioactive DNA. Radioactive DNA probes were prepared by labeling inserts from cloned DNA with radioactive [^{32}P] by the method of primer extension (Sambrook et al. 1989) using a kit from Pharmacia and dCT[^{32}P] obtained from New England Nuclear Corporation.

Polymorphisms were identified by comparing the size of radioactive restriction fragments obtained from digested 'Minsoy' DNA with fragments from digested 'Noir 1' DNA.

Segregation of markers

RFLP markers were scored in Utah against segregating progeny using pooled DNA prepared from families of 20 or more F_3 plants, each derived from a single F_2 plant. Morphological, biochemical, and isozyme markers were scored on F_2 plants in Iowa.

Construction of linkage map in Fig. 1

F_3 families from 69 of the F_2 plants were used for constructing the genetic map. The amount of DNA available from 13 of these families was limited. These 13 segregants were only scored for those markers that appeared to be linked to quantitative trait loci (QTL) (see Mansur et al. 1993a) or for distant markers with lower log of likelihood (LOD) scores. In all, about one-half of the markers were screened against all 69 plants, the rest against only 56 of the plants. Markers were screened against equal numbers of segregating progeny from 'Minsoy' × 'Noir 1' and from 'Noir 1' × 'Minsoy' crosses. After scoring the segregation of markers, each locus was tested for deviation from the expected 1:2:1 or 3:1 ratios by a chi-square test. Subsequently, they were analyzed for linkage using the "Mapmaker" computer package obtained from Dr. E. S. Lander (Lander et al. 1987): loci were grouped at a LOD threshold of 3 using tow-point comparisons, ordered using 'three-point' analyses, and mapped.

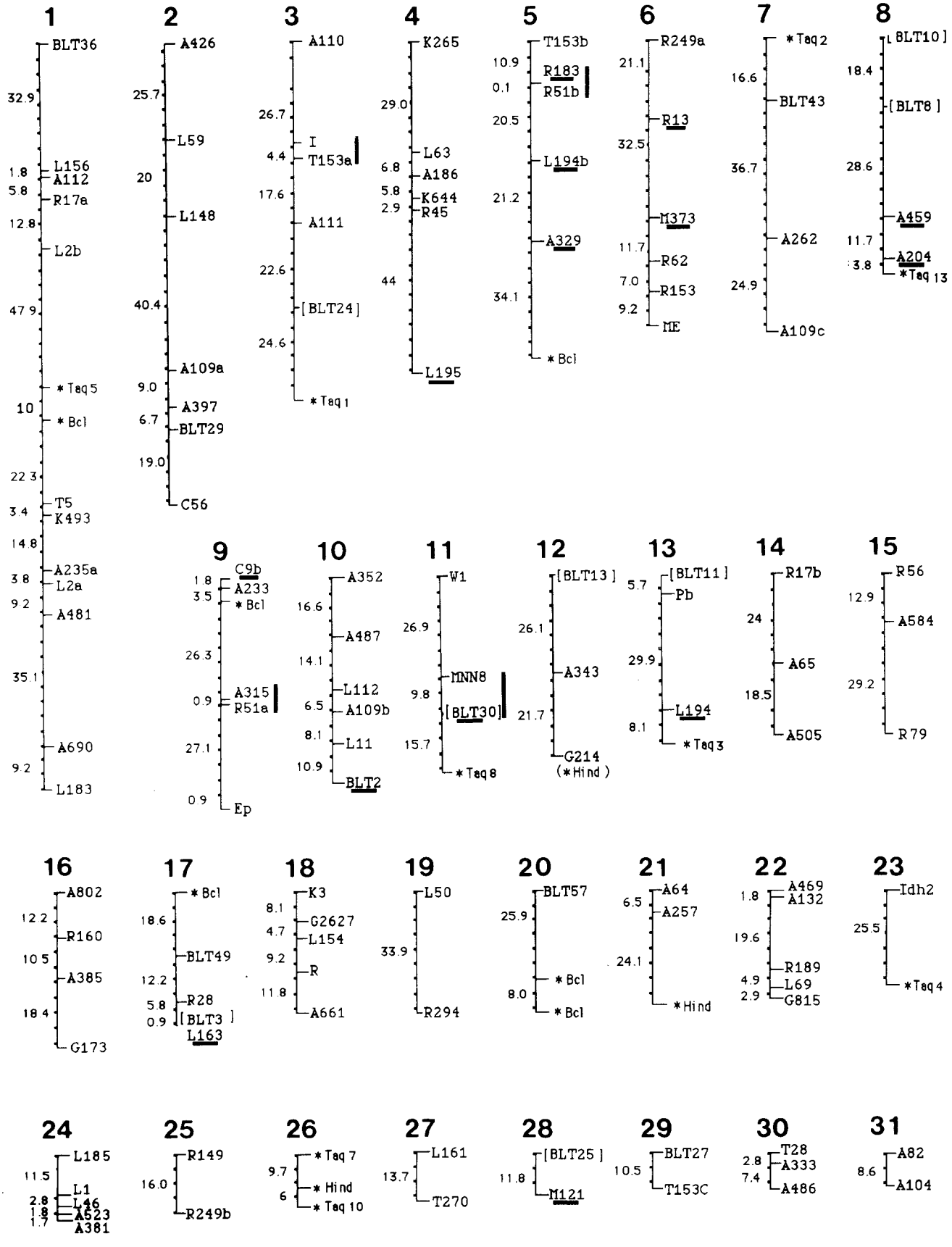
Results

As already reported (Apuya et al. 1988), polymorphisms have been found with a frequency of about one in five probes using random clones of low-copy genomic DNA as probes. For those probes (A or K) already selected to be polymorphic in the *G. soja* by *G. max* cross, the frequency was higher, about one in two.

One hundred and fifty-six markers were analyzed for segregation among progeny obtained by crossing 'Minsoy' and 'Noir 1'. These markers included RFLP loci corresponding to genomic DNA or cDNA probes, and isozyme, morphological, and biochemical markers. Of these, 132 were placed into 31 linkage groups with a LOD score of 3 or greater (Fig. 1). The linkage groups in Fig. 1 comprise 1550 cM of the soybean genome. In all of the linkage groups except 3, 5, 9, and 11 the order of loci appeared to be unambiguous. In linkage groups 3, 5, 9, and 11 the relative orders of closely linked loci (|) remained indeterminate. Map positions for one of two possible orientations are shown. The order of markers *I* and pT153a in linkage group 3 is consistent with data from the interspecific, *G. soja* – *G. max* map (Keim et al. 1989a; Diers et al. 1992). Marker pBLT24 is part of a cDNA clone for thiol-protease (Kalinski et al. 1990). It co-segregated with linkage group 3 at a LOD score greater than 4. However, its position was not well-defined on the basis of linkage to other qualitative markers. The order shown in Fig. 1 is consistent with a regression analysis of seed oil assuming a single major QTL for that trait (Mansur et al. 1993a).

RFLP markers corresponding to cDNA (brackets, [ ], in Fig. 1) were linked with about the same fre-

Fig. 1. A genetic linkage map derived from an analysis of segregating progeny obtained from a cross between 'Minsoy' (PI 27.890) and 'Noir 1' (PI 290.136). Each linkage group is subdivided into units of 4 cM. Markers are noted on the *right*, distances between markers on the *left*. See Materials and methods for a description of the probes used to identify RFLP markers and for a description of the isozyme, biochemical, or morphological markers. A repeated sequence present in G214 identified fragments (see Fig. 2) found in many different linkage groups (*). The enzyme used to produce the fragment is shown (*Bcl*, *Hin*d, or *Taq*). Markers identified by cDNA probes are enclosed in square brackets. Markers defined by the enzyme *Bcl*I are *underlined* (–). Pairs of markers for which the order is ambiguous are denoted by a *vertical line* (|)


1
BLT36
32.9
L156
1.8
A112
5.8
R17a
12.8
L2b
47.9
*Taq5
10
*Bcl
22.3
T5
3.4
K493
14.8
A235a
3.8
L2a
9.2
A481
35.1
A690
9.2
L183
2
A426
25.7
L59
20
L148
40.4
A109a
9.0
A397
6.7
BLT29
19.0
C56
3
A110
26.7
I
4.4
T153a
17.6
A111
22.6
[BLT24]
24.6
*Taq1
4
K265
29.0
L63
6.8
A186
5.8
K644
2.9
R45
44
L195
5
T153b
10.9
R183
0.1
R51b
20.5
L194b
21.2
A329
34.1
*Bcl
6
R249a
21.1
R13
32.5
M373
11.7
R62
7.0
R153
9.2
ME
7
*Taq2
16.6
BLT43
36.7
A262
24.9
A109c
8
[BLT10]
18.4
[BLT8]
28.6
A459
11.7
A204
3.8
*Taq13
9
C9b
1.8
A233
3.5
*Bcl
26.3
A315
0.9
R51a
27.1
0.9
Ep
10
A352
16.6
A487
14.1
L112
6.5
A109b
8.1
L11
10.9
BLT2
11
W1
26.9
MNN8
9.8
[BLT30]
15.7
*Taq8
12
[BLT13]
26.1
A343
21.7
G214
(*Hind)
13
[BLT11]
5.7
Pb
29.9
L194
8.1
*Taq3
14
R17b
24
A65
18.5
A505
15
R56
12.9
A584
29.2
R79
16
A802
12.2
R160
10.5
A385
18.4
G173
17
*Bcl
18.6
BLT49
12.2
R28
5.8
[BLT3]
0.9
L163
18
K3
8.1
G2627
4.7
L154
9.2
R
11.8
A661
19
L50
33.9
R294
20
BLT57
25.9
*Bcl
8.0
*Bcl
21
A64
6.5
A257
24.1
*Hind
22
A469
1.8
A132
19.6
R189
4.9
L69
2.9
G815
23
Idh2
25.5
*Taq4
24
L185
11.5
L1
2.8
L46
1.8
A523
1.7
A381
25
R149
16.0
R249b
26
*Taq7
9.7
*Hind
6
*Taq10
27
L161
13.7
T270
28
[BLT25]
11.8
M121
29
BLT27
10.5
T153C
30
T28
2.8
A333
7.4
A486
31
A82
8.6
A104

quency (8 out of 11) as other markers and were interspersed throughout the map.

The RFLP markers used in this map were detected using the enzymes *Bcl*I, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV, *Hin*dIII, and *Taq*I. Among markers which could be placed in linkage groups, 14 were detected with *Bcl*I and 11 of these appeared to be well-separated from the 96 detected with other enzymes. An additional 5 *Bcl*I markers remained unlinked in contrast to only 19 unlinked markers detected with all of the other enzymes. The 14 linked markers detected with *Bcl*I were located on average 19.5 cM from markers detected with other enzymes, in contrast to an average separation between other linked RFLP markers of 11.3 cM. This broad distribution of *Bcl*I markers made this enzyme particularly useful for RFLP mapping in this soybean cross.

An unusual set of RFLP markers was obtained using the probe G214. This probe hybridized to a number of restriction fragments that segregated from each other (Fig. 2). The map positions of these markers have been designated in Fig. 1 by asterisks (*). Almost all of these fragments behaved as independent markers segregating as alleles in which a fragment was either present or absent (not replaced by an allele corresponding to another fragment). These markers were mapped as dominant markers using mapmaker. A characteristic of these fragments was that most of them appeared to hybridize with a similar amount of radioactive probe. As can be seen, most of these markers were located at a distance from other linked markers, thus extending the linkage map or combining markers into a single large linkage group. Unpublished data (C. Lark, K. G. Lark and T. Macalma) have related this family of markers to an approximately 500-bp region of the G214 probe and shown that the co-dominant G214 marker in linkage group 13 is identified by a low-copy sequence flanking this 500-bp region.

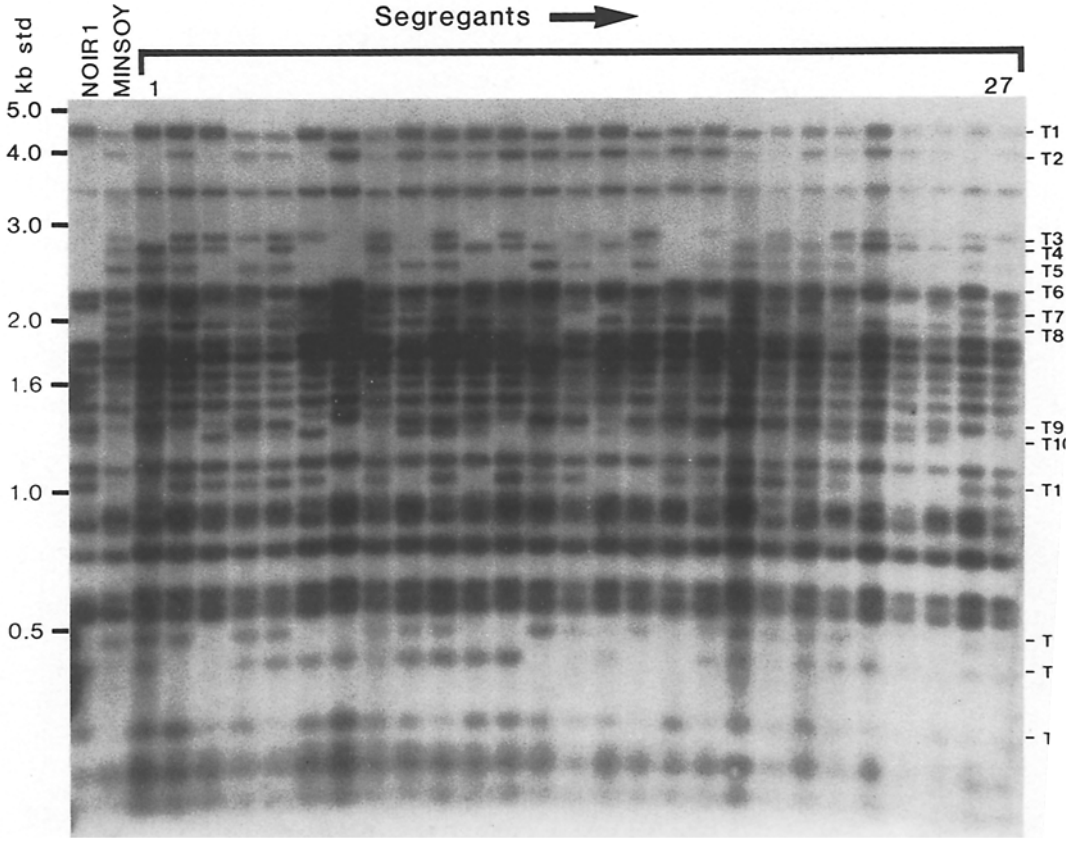


Fig. 2. Segregation of restriction fragments corresponding to a repeated sequence identified by the probe G214. DNA was prepared from parental 'Minsoy' and 'Noir 1' plants and from segregants from the cross between these two parents. After restriction with the enzyme *Taq*I, a Southern blot was prepared and hybridized with a radioactive G214 probe. Fourteen segregating fragments were identified in the autoradiogram. The genetic location of 9 of these fragments are shown in Fig. 1 (**Taq*). Similar results were obtained with the enzymes *Hin*dIII or *Bcl*I; and map positions of an additional 8 fragments determined from autoradiograms using DNA digested with either *Hin*dIII or *Bcl*I are also shown in Fig. 1 (**Bcl* or **Hin*d)

Discussion

An intraspecific genetic map for soybean has been constructed using cvs 'Minsoy' and 'Noir 1'. The 1550-cM map in Fig. 1 comprises 132 genetic markers contained in 31 linkage groups. While the map is clearly not complete – soybean has 20 chromosomes ($2n = 40$) – it has provided an initial genetic structure for the study of QTL that control agronomic traits (Mansur et al. 1993a,b).

As noted, the frequency of polymorphisms was much higher using probes already proven to be polymorphic in an interspecific cross between *G. max* and *G. soja*. It was therefore expected that most of these probes would yield linkage groups similar to linkages found in that cross. However, this was only true for 7 linkage groups (1, 2, 3, 4, 6, 7, and 31 in Fig. 1). In all of these the order of the markers corresponds although linkage distances are different. In other linkage groups where the same probes were used in both crosses, linked markers in our cross did not correspond with any single linkage group in the interspecific cross. This could indicate that a given probe identifies different polymorphic fragments in the two crosses.

An interspersed, repeated sequence that hybridized with a 500-bp portion of the probe G214 defined 17 loci [Fig. 1(*)] with an average contribution of 16 cM per locus. Several of these loci were closely linked to each other, as seen in linkage groups 1, 20, and 26 (Fig. 1). [Another 7 such loci, which were linked more closely (< 5 cM) are not shown]. Unpublished data of K. G. Lark and C. Lark suggest that the sequence involved is a transposable element. It is found in different restriction fragments in different soybean lines, and it seems likely that marker locations also will vary between different crosses. For this reason, we have not assigned specific names to these markers beyond their association with a specific restriction enzyme (as presented in Fig. 1 in relation to the data in Fig. 2).

Within our cross, however, these markers are extremely useful. Recombinant inbred lines (RIL) developed from this cross (Mansur et al. 1993b) show that these fragments have remained sufficiently stable through the F_{10} generation to be useful in mapping studies across the genome (K. G. Lark, L. M. Mansur and J. Orf, unpublished data). For example, a QTL for carbon isotope discrimination (Mansur et al. 1993a) was defined by the marker G214 (**Taq5*), and the map position of this RFLP marker has remained unchanged in RIL after eight additional generations. These markers also should be particularly useful in structuring the map, i.e., linking together clusters of other markers (as in linkage group 2 in Fig. 1). Finally, the broad distribution of such markers over the genome should make them useful in analyzing genotypes of backcross progeny to determine which are most like the recurrent parent used in the cross.

The distribution of the repeated element suggests that at least two mechanisms may be responsible for genetic variation in soybean. In one, sequence changes leading to restriction fragment site polymorphisms for the enzymes *Eco*RI, *Bgl*II, *Eco*RV, *Hin*dIII or *Taq*I (but not *Bcl*I, see below) are clustered in certain regions of the genome (one could imagine the modification of regions of DNA, possibly by methylation, eventually resulting in sequence changes, due perhaps to recombinational events). In the other, sequence changes occur adjacent to a repeated sequence, possibly due to the insertion or excision of a transposable element. In the case of the element studied here, it appears to be preferentially located in regions lacking RFLP markers detected by enzymes such as *Eco*RI, *Hin*dIII, or *Taq*I. Whatever the mechanism, the complementary location of the repeated G214 sequences as opposed to most other RFLP markers suggests the possibility of at least two structures in the soybean genome that could control different types of marker variation.

The data in Fig. 1 suggest that the polymorphisms we have detected with *Bcl*I are not compartmentalized. We have already noted the distance of many of these markers from other RFLP markers. In addition, markers determined by *Bcl*I are much closer to sites occupied by the G214 repeated sequence (an average of 17.4 cM as opposed to 31 cM for other markers). The small number of *Bcl*I markers analyzed here may be misleading, and further analysis of more markers in this and other crosses will be needed to confirm the generality of this phenomenon. However, if confirmed, such data suggest that restriction sites may not be distributed randomly over the soybean genome.

We have presented a first version of an intraspecific genetic map of soybean. Future comparisons between the data from this cross and that from crosses between *G. max* and *G. soja* should help to discriminate between those components of the soybean genetic map that are general and those that are specific to individual crosses. The lower frequency of polymorphisms in the intraspecific cross has made it more difficult to construct a genetic map. However, this is offset by the advantages that can be derived from its use in the genetic analysis of agronomic traits such as yield, plant height, and lodging, which are difficult to evaluate in interspecific crosses. In subsequent papers, we present a genetic analysis of several important agronomic traits and demonstrate that this cross is admirably suited to the genetic analysis of quantitative trait loci (Mansur et al. 1993a, b).

Acknowledgments. This research was supported by the American Soybean Association, funds from the USDA and Grant GM42337 to K. G. L. from the NIH.

References

Apuya NR, Frazier BL, Keim P, Jill E, Lark KG (1988) Restriction fragment polymorphisms as genetic markers in soybean, *Glycine max* (L.) Merr. Theor Appl Genet 75:889–901

Botstein D, White R, Skolnick M, Davis R (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Genet 32:314–331

Cardy BJ, Beversdorf WD (1984) A procedure for the starch gel electrophoretic detection of isozymes of soybean. Univ Guelph Tech Bull 119/8401

Diers BW, Keim P, Fehr WR, Shoemaker RC (1992) RFLP analysis of soybean seed protein and oil content. Theor Appl Genet 83:608–612

Helentjaris T, Slocum M, Wright S, Shaefer A, Nienhuis J (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor Appl Genet 72:761–769

Kalinski A, Weisemann JM, Matthew BF, Herman EM (1990) Molecular cloning of a protein associated with soybean seed oil boides that is similar to thiol proteases of the papain family. J Biol Chem 264:13843–13848

Keim P, Shoemaker RC (1988) Construction of a random recombinant DNA library that is primarily single-copy sequences. Soybean Gene Newsl 15:147–148

Keim P, Diers BW, Palmer RG, Shoemaker RC, Macalma T, Lark KG (1989a) Mapping the soybean genome with RFLP markers. In: Pascae AJ (ed) Realization Orients Graphica Editoria Srl, Bucnos Aires Proc IV World Soybean Res Conf vol 3, pp 1246–1251

Keim P, Shoemaker RC, Palmer RG (1989b) Restriction fragment length diversity in soybean. Theor Appl Genet 77: 786–797

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

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

Landry B, Kesseli R, Farrara B, Michelmore R (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331–337

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

Mansur LM, Lark KG, Orf J (1993b) Determining linkage of quantitative trait loci to RFLP markers using bulked DNA from extreme phenotypes of recombinant inbreds of soybean (*Glycine max* L. Merr.). Theor Appl Genet (in press)

Palmer RG, Kilen TC (1987) Qualitative genetics and cytogenetics. In: Wilcox JR (ed) Soybean: improvement, production and uses, 2nd edn. Agronomy 16:135–199

Roth EJ, Frazier, BL, Apuya NR, Lark KG (1989) Genetic variation in an Inbred plant: Variation in tissue cultures of soybean [*Glycine max* (L.) Merrill]. Genetics 121:359–368

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Tanksley SD, Mutschler MA, Rick CM (1987) Linkage map of the tomato (*Lycorpersicum esculentum*) (2N = 24). In: O'Brien SJ (ed) Genetic maps 1987, vol 4. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 654–669

Vallejos CE, Sakiyama NS, Chase CD (1992) A molecular marker-based linkage map of *Phaseolus vulgaris* L. Genetics 131:733–740