BIOLOGIA PLANTARUM 43 (3): 337-346, 2000

Linkage mapping of DNA markers generated with specific and non-specific gene primers in soybean

K.M. SOLIMAN*, M. PALADUGU* and T.E. DEVINE**

*Department of Plant and Soil Science, Alabama A&M University, Normal, AL 35762, USA**
*Weed Science Laboratory, PSI, BLDG 001, BARC-W, USDA/ARS, Beltsville, MD 20705, USA***

Abstract

Polymerase chain reaction (PCR) has been used extensively in the construction of linkage maps for many cultivated crops including soybean, [*Glycine max* (L.) Merr]. In this study, four sets of oligonucleotide primer pairs of known genes (pearl millet *Adh* 1, nodule specific proline-rich protein, *Drosophila* homeobox, heat shock protein), several different combinations from kits A, D, E, and J of arbitrary primers and five primer pairs of soybean simple sequence repeats of varying length (Satt 9, Satt 20, Satt 42, Satt 64, and Satt 30) were utilized in PCR to identify molecular markers which were then used to construct a genetic linkage map. DNA for the PCR reactions was isolated from 65 recombinant inbred soybean lines resulting from crossing PI 290,136 and BARC-2 (Rj_4), followed by self-pollination for seven generations without selection. Mapmaker 3.0, a computer package, was used for construction of the linkage map. A total of 43 polymorphic markers were identified; 30 markers were linked and distributed among 5 linkage groups while 13 markers were unlinked. Arbitrary primers revealed more polymorphisms than specific primers. A combination of arbitrary primers A5 and A18 revealed the maximum number of polymorphic bands. Five observed linkage groups can be expanded in future soybean research by using additional markers.

Additional key words: *Glycine max*, heat shock protein, homeobox, Polymerase Chain Reaction (PCR), simple sequence repeat

Introduction

DNA markers linked to important agronomic traits have become critical components of plant breeding programs to help in the identification and selection of desirable genotypes in segregating populations (Lander *et al.* 1987, Beckman *et al.* 1990, Mesfin *et al.* 1990, Kolchinsky *et al.* 1993, Skroch and Niienhuis 1995, Garvin *et al.* 1998, Yu *et al.* 1998, Berger and Minor 1999). However, the general utilization of this approach by breeders is hampered by the technical inconveniences of RFLP analysis, an approach used to detect DNA variations. RFLP marker-assisted selection is time consuming, requires access to a laboratory fully equipped for DNA manipulation, the use of radioactive isotopes, and personnel trained in molecular genetics and radioactive waste handling and disposal. These requirements are generally not met by breeding teams. In addition, the RFLP technique necessitates large samples of plant tissue often available only from mature plants, rending impossible early and nondestructive testing of young seedlings with RFLP markers (Landry and Michelmore 1987).

Linkage maps have been or are being constructed for genomes of various major crop plants (Helentjaris *et al.* 1986, Bonierbale *et al.* 1988, Lark *et al.* 1993, Mansur *et al.* 1993, McCouch *et al.* 1998). First correlation between RFLP markers and qualitative and quantitative traits have been reported (Regulski *et al.* 1985, Sarfatti *et al.* 1989, Barone *et al.* 1990, Klein-Lankhorst *et al.* 1991, Yu *et al.* 1991, 1998, Berger and Minor 1999, Mesfin *et al.* 1999).

Polymorphisms generated by Arbitrarily Primed Polymerase Chain Reaction (AP-PCR) have been applied to the identification of inbred parents in maize (Welsh

Received 1 September 1999, *accepted* 10 January 2000.
Acknowledgements: Contributed by the Agricultural Experiment Station, Alabama A&M University. This research was supported in part by the USDA/CSREES Grant 93-38814-8888. The assistance of Mr. Gerry Mon in implementing the *Mapmaker* program is specially acknowledged.
Correspondence should be addressed: K.M. Soliman, fax: (+256) 851 5429, e-mail: ksoliman@aamu.edu

et al. 1991) and a PCR-based co-dominant marker which is tightly linked to the nematode resistance gene *Mi* was developed in tomato (Williamson *et al.* 1993). Ovidio *et al.* (1990) used PCR to amplify genomic DNA of several wheat and maize genotypes with the terminal sequence of a γ-gliadin gene and the specific amplification revealed both inter- and intra-specific genetic polymorphisms among the examined genotypes.

Mansur *et al.* (1993) mapped quantitative trait loci (QTL) in segregating progeny from a cross between two soybean [(*Glycine max* (L.) Merr)] cultivars: Minsoy (PI 27,890) and Noir 1 (PI 290,136). The 15 characters analyzed included reproductive, morphological, and seed traits, seed yield and carbon isotype discrimination ratios ($^{13}C/^{12}C$). Genetic variation was detected for all of the traits and transgressive segregation was common. 132 linked genetic markers and 24 additional unlinked markers were used to locate QTL by interval mapping and one-way analysis of variance, respectively. Byrum *et al.* (1993) identified a RAPD marker 8.9 ± 4.6 map units from *RPS*$_1$ in soybean using a primer OPM-11.

Kolchinsky and Gresshoff (1993) used soybean telomeres as molecular markers after a method for direct labeling of the ends of chromosomes was developed. Telomeres of soybean, cv. Bragg and *Glycine soja* (Sieb and Zucc., PI 468,397) were labeled, the DNA digested with *Hinf* I and analyzed by electrophoresis. To verify the telomeric nature of the labeled fragments, identical non-labeled samples were run in the same gel and hybridized to the plant telomeric probe obtained from "STAGGERED" PCR.

Zhang *et al.* (1997) demonstrated that arbitrary and synthetic gene primers could be utilized in PCR to identify molecular markers which can be used to differentiate species and genotypes. Taylor-Grant and Soliman (1999) characterized the wild relatives of soybean (*Glycine glycine* L.) via tissue culture and PCR. Their results indicated that specific oligonucleotide primers revealed far more amplified products than did the arbitrary primers. However, both methods were effective in constructing a phylogenetic tree of all the species studied.

Skroch and Nienhuis (1995) reported the results of a survey of 400 RAPD primers for polymorphisms among a set of ten snap bean (*Phaseolus vulgaris* L.) genotypes. They were able to identify a set of primers that could be used to more efficiently sample RAPD marker loci and could potentially reduce the number of primers required for the analysis of genetic relationships among snap bean genotypes by over 60 %. These selected RAPD markers may be useful for the unique identification of bean varieties, the organization of bean germplasm, and applications of molecular markers to bean breeding.

Microsatellites or Simple Sequence Repeats (SSRs) are ideal genetic markers because they: *1*) are highly abundant, *2*) appear to be evenly distributed throughout the genome, *3*) are highly polymorphic, *4*) can be typed rapidly via PCR and *5*) are disseminated easily among laboratories by publishing primer sequences (Zhao and Kochert 1993). Akkaya *et al.* (1992) identified an SSR marker closely linked to the Soybean Mosaic Virus resistance gene and were able to detect SSR variation with agarose gel electrophoresis without the use of radiochemicals. Provan *et al.* (1996) reported the amplification of polymorphic microsatellite markers from potato (*Solanum tuberosum*) using 22 sets of primers derived from potato and four sets of primer derived from tomato (*Lycopersicon esculentum*) genomic DNA sequences. Rongwen *et al.* (1995) used a set of microsatellite markers to provide a positive assessment of the ability of SSR markers to produce unique DNA profiles of 96 diverse soybean genotypes. The gene diversity for the seven markers averaged 0.87 for all 96 genotypes and 0.74 for a subset of 26 North American culitvars.

Condit and Hubbell (1991) found that dinucleotide repeat regions are abundant in plant genomes, being numerous enough to be isolated from libraries of small inserts. The clones produced could be completely sequenced without constructing new sequencing primers. Calculations by Condit and Hubbell (1991) suggested that in plants there were approximately 104 to 105 two-based repeats or one in every 60 - 600 inserts produced by a four-base restriction endonuclease. Akkaya *et al.* (1992) found that for three soybean loci, *i.e.* SOYHSP176 $(AT)_n$, SOYSC514 $(AT)_n$, and SOYPRP1 $(ATT)_n$, an average of seven alleles per locus were detected. They reported two findings for length polymorphisms of SSRs from their experiments, first, the presence of polymorphic length SSRs in soybean demonstrated that this type of genetic marker was useful in genetic analysis of higher plant species. The rapidity of PCR, combined with the informativeness of SSR length polymorphisms, should make SSR markers valuable tools for molecular map development, genotype identification, and other uses (Zhao and Kochert 1993). The second result of interest relates to the lack of $(CA/GT)_n$ sequences in soybean comparable to that found in humans. The highly polymorphic nature of SSR markers make them particularly advantageous in a species such as soybean in which RFLPs have been somewhat difficult to detect.

The objectives of our study were to utilize arbitrary, specific and simple sequence repeat primers to generate molecular markers to be used in the generation of a genetic linkage map.

Materials and methods

Plants: Sixty-five F_8 recombinant inbred lines, resulting from crossing PI 290,136 and BARC-2 (*Rj4*), were used in this study. BARC-2 (*Rj4*) contains the dominant (*Rj4*) allele which conditions the ineffective nodulation response characterized by the development of cortical proliferation or rudimentary nodules when plants are inoculated with *Rhizobium* strain USDA 61. The two parents also differ in other traits such as flower color (*W*), root fluorescence (*Fr2*), inhibitor gene causing inhibition of the seed coat color, early flowering (*E3*), and bacterial postule resistance (*Rxp*).

DNA isolation: Total cellular DNA of soybean was isolated from fresh young leaves of one plant from each recombinant inbred line as described by Saghai-Maroof *et al.* (1984). The plants were grown in individual pots in the greenhouse for three to four weeks before fully expanded leaves were harvested for DNA extraction.

Synthesis of oligonucleotide primers: Specific gene oligonucleotide primers were synthesized by standard phosphoramidate chemistry on a *Pharmacia LKB Gene Assembler Plus* (Uppsala, Sweden) DNA synthesizer. The specific primers were all oligodeoxynucleotide 20 to 24-mers containing no internal repeats. The protecting groups of the synthesized primers were removed by placing them in 35 % ammonium hydroxide at room temperature for 24 h and then desalted with *Pharmacia NAP* columns containing *Sephadex G-25*.

The gene-specific primers were designed using published cDNA sequences for *1*) the nodule-specific (hydroxyl) proline-rich protein *LENOD2* gene of yellow lupine (5'-TCTCCGCCTGTCACCCCCA-3') (Szczyglowski and Legocki 1990), *2*) the alcohol dehydrogenase gene (*Adh1*) of pearl millet (5'-GTGATCAAGTGCAAA GCCGCC-3', 5'-CCCCTCCCGTA GGCGACGTAG-3') (Ha *et al.* 1990), *3*) the heat shock protein (*Hsp*) gene of *Pisum sativum* (5'-CGTGCGACAAACACAAAATC ATCC-3', 5'-TTCTTTCTTTTACAATGTAAATG-3') (Lauzon *et al.* 1990), and *4*) the homeobox sequence from the *Drosophila* deformed gene (*Dfd*) consisting of 24-mers encoding for amino acids 1 - 8 and 54 - 61, respectively (5'-CCAAAACGCCAACGCACCGC CTAC-3', 5'-TACTTCACCTTCTTCCTGTTGTTC-3') (Regulski *et al.* 1985).

The root nodule gene was selected because nodules are highly organized structures formed on leguminous plants as a result of interaction with *Rhizobium*. Formation of root nodules involves differential expression of nodule-specific plant genes, including the proline-rich protein genes which are expressed at the early stages of root nodule development (Fuller *et al.* 1983). Alcohol dehydrogenase is a cytosolic enzyme that breaks down alcohol. In most plant species, including maize and pearl millet, two genes (*Adh-1* and *Adh-2*) encode alcohol dehydrogenase. Each gene has two different allelic forms (F and C) in cultivated species. Genes of heat shock proteins (*Hsp*), induced during high temperature stress, belong to a low molecular mass gene family, and whose function is to provide thermotolerance. It has been reported that high temperature induces the synthesis of a novel set of Hsp in soybean (Key *et al.* 1981).

Using *Drosophila* homeobox sequences as probes, previously unknown homeotic genes have been isolated from other higher organisms, including vertebrates and mammals, indicating that homeobox sequences have been highly conserved during evolution. The gene *Dfd* used in this study is part of the antennapedia complex and is homologous to mouse (*Hox 2.6*) and human (*Hox 4.2*). An examination of these genes revealed that 9 amino acids in the homeodomanin are invariant in 16 *Drosophila* and 22 other domains from sea urchins (*Arabacia punctulata*) to humans (Regulski *et al.* 1985). As homeobox sequences influence development in both animals and plants, we tried to answer the question of whether or not soybean exhibits diversity for DNA sequence when primed by a homeobox motif. Furthermore, what is the linkage relationship between the markers generated with the homeobox motif and those generated with other primers?

A total of five pairs of simple sequence repeat primers were used in this study. They were designated Satt 1, Satt 9, Satt 20, Satt 33, and Satt 42 (Rongwen *et al.* 1995). These primer pairs were selected based on their high gene diversity values of 0.84, 0.84, 0.90, 0.83, and 0.88, respectively. All of these primers have repeated AT sequences varying in length from $(AT)_{17}$ to $(AT)_{33}$ and were synthesized by *Research Genetics* (Huntsville, USA). A total of 30 arbitrary primers, in two-primer combinations, were also used in this study and were selected randomly from 4 different kits (A, D, E, J) purchased from the *Operon Technologies* (Alameda, USA).

PCR amplification conditions: PCR amplifications were performed in reaction volumes of 0.05 cm^3 in a 0.5 cm^3 microcentrifuge tube (*Perkin-Elmer Cetus*, Foster City, USA) containing the following: sterile distilled water, 1X Taq DNA polymerase buffer, 200 μM dNTPs, 0.2 μM of each primer, 1.25 mM $MgCl_2$, 70 ng of total genomic DNA of soybean, and 2.5 units of Taq DNA polymerase (*Promega*, Madison, USA). Amplification was carried out in a *Perkin-Elmer Cetus* DNA Thermal Cycler programmed for 44 cycles of 1 min at 94 °C for DNA denaturation, 1 min each at 55 °C and 36 °C for primer annealing of specific and arbitrary primers, respectively, and 2 min at 72 °C for primer extension. At

the end of the cycle program, fragments were extended for 10 min at 72 °C. Amplified PCR products were resolved by electrophoresis on 1.5 % agarose gels.

The amplification conditions for the SSR primers were carried out in 0.025 cm^3 volume essentially as described above except for 47 cycles with a 1 °C decrease in the annealing temperature from 60 °C to 47 °C. A 100 bp DNA ladder (*Gibco BRL*, Gaitherburg, USA) was used as a size marker throughout. Each amplified product was identified by staining gels with ethidium bromide for 10 min (specific and arbitrary primers), destained with double distilled water (ddH_2O) and photographed under UV light with *Polaroid* type *55* film. The amplified products generated with the SSR primer pairs were resolved by staining gels in silver nitrate as follows: gels were kept in 10 % acetic acid for 10 min and then rinsed twice in ddH_2O followed by soaking the gels in silver reagents for 15 min and rinsing two times in ddH_2O, finally fresh developer was added several times until clear bands appeared followed by fixing in 5 % acetic acid. Vertical gels were prepared by mixing 1.25 cm^3 *GeneAmp* detection gel (5× concentrate) at 19 to 30 °C (*Perkin Elmer*), 0.25 cm^3 of 10× TBE buffer, 3.5 cm^3 of H_2O, 0.005 cm^3 of TEMED and 0.06 cm^3 of 10 % ammonium persulfate (APS). The gels were run in a *Bio-Rad Mini-Protean II* (Hercules, USA) apparatus for 1 h at 75 V in 1× TBE buffer. The amplified products were visualized under incandescent light and photographed with *Polaroid* type *55* film.

Data collection: The DNA patterns generated with each set of primers were evaluated with the assistance of a computer-integrated laser densitometer (*Ultrascan XL, LKB Producta AB*, Bromma, Sweden) by using *Gelscan XL* version *2.0* software (*Pharmacia LKB Biotechnology*, Uppsala, Sweden). Each pattern was examined for size (base pairs) and were recorded as "+" and "-" for presence and absence of a fragment, respectively. χ^2-test was used to evaluate the frequency distribution of each fragment among the sixty-five F_8 recombinant inbred lines used in this study. χ^2-tests were carried out to determine the number of genes controlling the respective markers.

The hypothesis that each polymorphic fragment generated by specific gene primer was controlled by a single gene was tested. For example, if the fragment 550 bp generated by the homeobox gene primers is controlled by a single gene, it is expected that ~ 49.6 % of the recombinant inbred lines will have the fragment and 49.6 % of the lines will not have this fragment. These proportions of homozygous lines are calculated based on the formula $(2^m\text{-}1)/2^m$ where m is the number of segregating generations which is equal to 7 in this example. Applying χ^2 formula to the 550 bp fragment $\chi^2[1df]$ = (observed-expected)^2/expected = $(22\text{-}32)^2/32 + (42\text{-}32)^2/32 = 6.25^*$.

The probability of obtaining a χ^2 as large as $(3.125 + 3.125) = 6.25^*$ is far less than 5 in 100, so the hypothesis of single locus is rejected. Similarly, χ^2 values were calculated for all polymorphic fragments amplified by all the primers used in this study.

The linkage relationships among all the markers generated were evaluated using *Mapmaker*, a program for genetic linkage analysis (Lander *et al.* 1987). *Mapmaker 3.0* performs full multipoint linkage analysis (simultaneous estimation of all recombination fractions from the primary data) for dominant, recessive, and co-dominant markers.

Results and discussion

Specific gene primers: Four sets of gene specific primers were synthesized and used in this experiment. All of these gene primers produced amplified products. Only the *Hsp* and Homeobox gene primers produced polymorphic fragments for both parental lines and their recombinant inbred lines.

The specific gene primer synthesized from the *Drosophila* Homeobox revealed three polymorphic bands ranging in size from 550 to 990 bp (Fig. 1). The fragments 870 and 550 bp were expressed in PI 290,136 but were absent in BARC-2 (Rj_4). Fragment 990 was amplified in BARC-2 (Rj_4) and not in PI 290,136, whereas, nine identical fragments (350, 440, 510, 640, 710, 780, 1090, 1340, 1450 bp) were amplified in both parents and their recombinant inbred lines. In addition, a 580 bp fragment was detected in four recombinant inbred lines (41, 62, 107, 109) but was absent in both parents and the remaining recombinant inbred lines examined.

The χ^2 values obtained for the 3 polymorphic fragments generated by Homeobox primers were significant (6.25^*, 22.56^{**}, 36.0^{**}) indicating that these fragments were not controlled by a single locus.

The specific gene primer synthesized from heat shock protein revealed three polymorphic bands ranging in size from 250 to 1540 bp. The fragments 1070 and 1170 bp were present in PI 290,136 and were absent in BARC-2 (Rj_4). Fragment 1290 bp was amplified in BARC-2 (Rj_4) and not in PI 290,136, whereas 13 identical fragments (250, 310, 380, 420, 470, 570, 600, 720, 850, 890, 930, 1440, 1540 bp) were present in both parents and their recombinant inbred lines. A 520 bp fragment was detected in five recombinant inbred lines (20, 45, 52, 99, 107) but was absent in both parents and the remaining inbred lines.

The χ^2 values obtained for the polymorphic fragments 1170 and 1290 bp were not significant indicating that

each was controlled by a single locus. On the other hand, the χ^2 value obtained for the polymorphic fragment 1070 bp was highly significant (38.1**) indicating that this fragment may be controlled by more than one locus.

Simple sequence repeats: As described earlier, this experiment utilized five pairs of simple sequence repeats selected based on their gene diversity values. Only Satt 20 produced polymorphic fragments for both parental lines and their recombinant inbred lines (Fig. 2). On the other hand, Satt 1, Satt 9, Satt 33, and Satt 42 were not polymorphic, and therefore, will not be discussed further.

The SSR primer pair Satt 20 revealed seven polymorphic bands in the recombinant lines from both

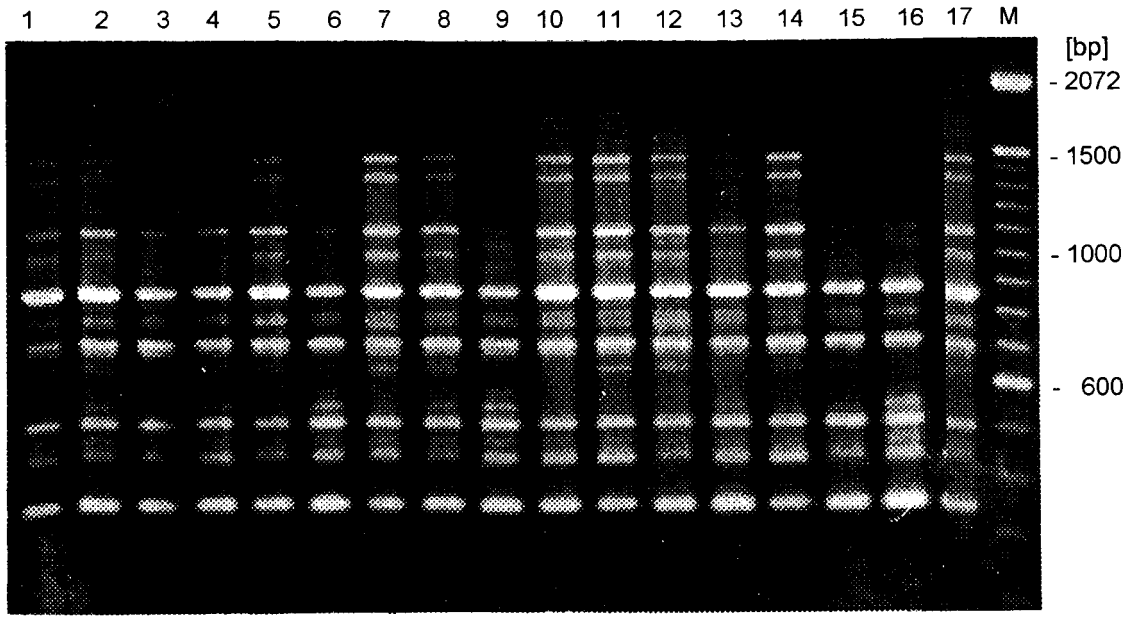

Fig 1. DNA amplification with *Drosophila* homeobox gene primers. *M* - 100 bp DNA ladder. *Lanes 1* to *17* - F_8 recombinant inbred lines. Polymorphisms were observed at 550, 870, 990 bp.

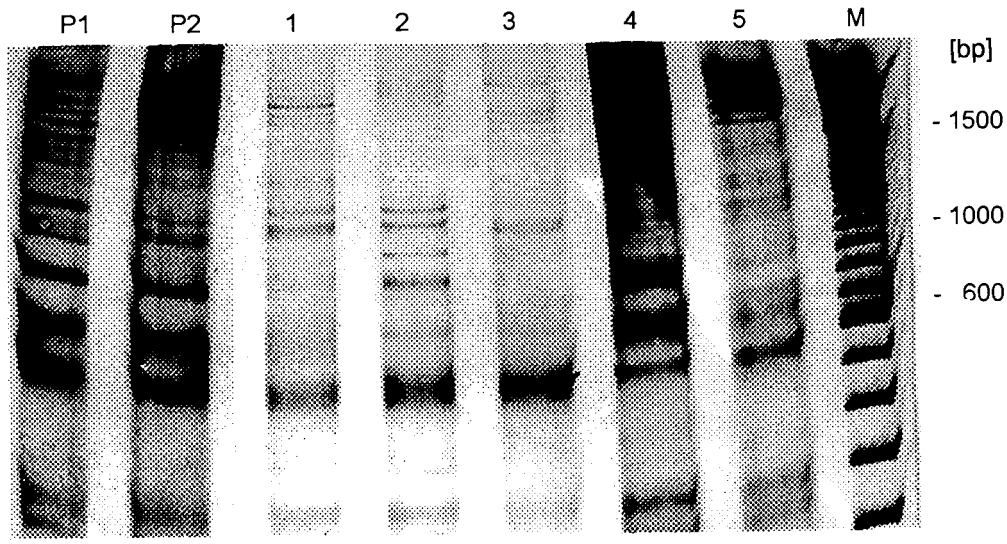

Fig 2. DNA amplification with simple sequence repeat (Satt 20) gene primers P1 (BARC 2), P2 (PI 290,136), *lanes 1* to *5* - F_8 recombinant inbred lines, *M* - 100 bp DNA ladder. Polymorphisms were observed at 1000, 1070, 1100, 1200, 1300, 2190. 2370 bp.

parents. Both parents (PI 290,136 & BARC-2) (*Rj*$_4$) had 14 amplified bands ranging in size from 80 to 2970 bp. The polymorphic fragments 1070, 1100, 1200 and 2370 bp were present in PI 290,136 but not in BARC-2 (*Rj*$_4$).

Fragments 1000, 1300 and 2190 bp, were amplified in BARC-2 (*Rj*$_4$) but not in PI 290,136, while seven identical fragments (80, 90, 190, 290, 470, 2600, 2970 bp) were amplified in the parents and their recombinant inbred lines.

The χ^2 values obtained for the polymorphic fragments (1000, 1300, 2190 and 2370 bp) were significant (28.44**, 28.44**, 31.00** and 8.12**, respectively) indicating departure from single gene segregation expectation. The χ^2 values obtained for the polymorphic fragments 1070, 1100, and 1200 bp were not significant implying that each of these fragments was controlled by a single locus.

Arbitrary primers: This experiment utilized 30 different two-primer combinations of arbitrary primers for genomic DNA amplification. The two-primer combinations of OPA 5 & OPA 18, OPA 20 & OPA 18, and OPE 4 & OPE 10 produced polymorphic fragments for both parental lines and their 65 recombinant inbred lines. The remaining 27 two-primer combinations were monomorphic, and therefore, will not be discussed further. The OPA 5 & OPA 18 (Fig. 3) combination revealed more polymorphisms than the other two combinations.

The primer combination of OPA 5 & OPA 18 revealed 24 bands ranging in size from 100 to 2560 bp. The fragments 840, 990, 1610, 2560 bp were present in BARC-2 (Rj4) and were absent in PI 290,136 while fragments 1670, 1120, 930, 870, 760, 720, 250, 160, 110 bp were amplified in PI 290,136 but not in the other parent. With this primer combination a total of 13 polymorphic bands were amplified in the recombinant inbred lines from the two parents above, whereas 11 identical fragments 130, 270, 330, 410, 450, 500, 540, 600, 1090, 1180, 1450 bp were amplified in the parents and their recombinant inbred lines. The fragments 250, 740, 1240, and 1380 bp were detected in some recombinant inbred lines but not in the parents or the remaining lines.

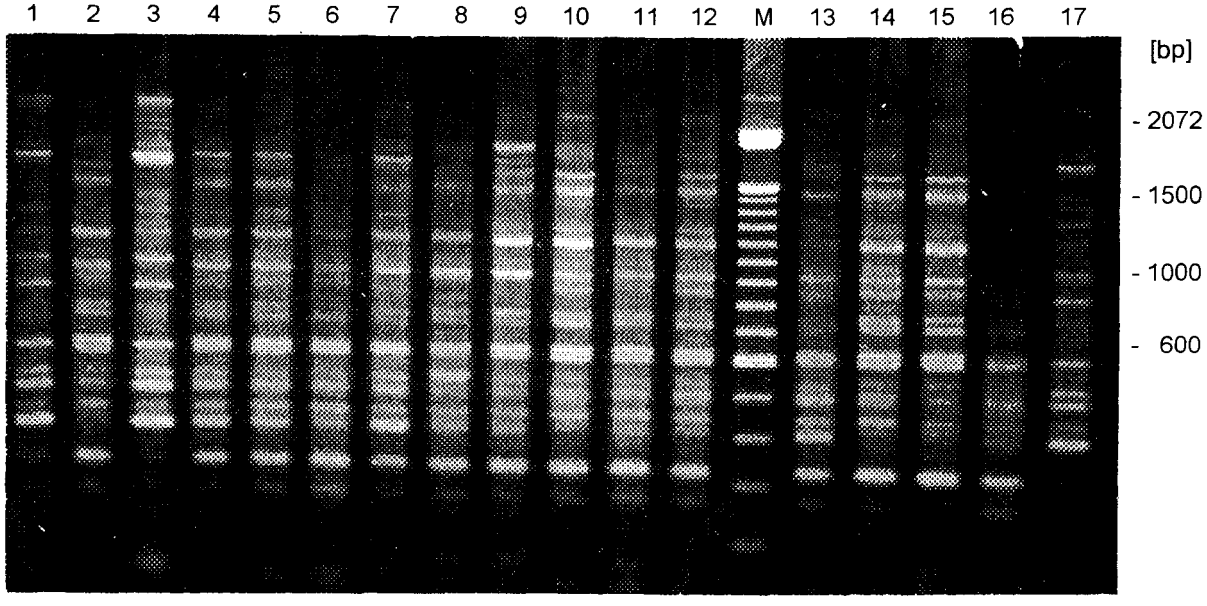

Fig 3. DNA amplification with arbitrary primers A5 & A18. M: 100 bp DNA ladder, *lanes 1* to *17* are F_8 recombinant inbred lines. Polymorphism observed at 250, 720, 760, 840, 870, 930, 990, 1120, 1610, 1670, 2560 bp.

The χ^2 values obtained for the generated polymorphic fragments 250, 720, 760, 930, and 2560 bp were significant at the 1 % level (16.0**, 36.0**, 36.0**, 10.56**, 12.25**). The significant χ^2 values indicated that more than one gene was involved. The χ^2 values were not significant for fragments 110, 160, 840, 870, 990, 1120, 1610, and 1670 bp, indicating that each of these fragments is controlled by a single locus.

The primer combination of OPA 18 & OPA 20 revealed two polymorphic bands of size 660 and 730 bp that were present in BARC-2 (*Rj*$_4$) and were absent in PI 290,136, while seven identical fragments 150, 450, 510, 560, 610, 1130, 2100 bp were amplified in the parents and their recombinant inbred lines. Furthermore, a 375 bp fragment was detected in recombinant inbred lines 3, 55, 62, 106, 116, 135 while neither parent nor the remaining recombinant inbred lines examined showed the 375 bp fragment. The χ^2 values obtained for the polymorphic fragments 660 and 730 were not significant indicating that each was controlled by a single locus.

The primer combination of OPE 4 & OPE 10 resulted in 19 amplified bands ranging in size from 310 to 1790 bp. The fragments 670, 730, 1240, and 1790 bp were present only in BARC-2 (*Rj*$_4$) while fragments 310, 600,

640, 760, and 800 bp were only amplified in PI 290,136 and were absent in BARC-2 (Rj_4), whereas 9 identical fragments 440, 570, 700, 790, 1000, 1160, 1350, 1520 bp were amplified in the parents and their recombinant inbred lines. Furthermore, an 880 bp fragment was detected in few recombinant inbred lines while neither the parents nor the remaining recombinant inbred lines examined showed the 880 bp fragment.

The χ^2 values obtained for fragments 310, 1240 and 1790 bp were not significant, indicating that each of these fragments is controlled by a single locus while values obtained for the other polymorphic fragments 600, 640, 670, 730, 760, and 800 bp were significant (11.56^{**}, 24.14^{**}, 37.67^{**}, 11.56^{**}, 11.56^{**}, 19.44^{**}) implying that more than one gene was involved.

Linkage map: A total of 37 markers generated from all primers utilized in this study were analyzed by *Mapmaker 3.0* (Lander *et al.* 1987). The linkage was determined using the maximum likelihood estimate with a minimum logarithm of difference (LOD) of 4.0 and a maximum distance of 30 cM. The results indicated that 30 of the 37 molecular markers were linked and were distributed among five linkage groups (LGs), covering 264.9 cM (Fig. 4). Linkage group 1 contained 15 markers while LG 2 contained 9 markers and the remaining 3 LGs contained 2 markers each. All Satt markers were mapped to LG 1 whereas the RAPD markers were distributed among the five LGs. The map distance between these markers varied from ~ 0 to 33.5 cM with LG 1 spanning 147.7 cM while LG 2 covered a distance of 127.8 cM and the remaining three linkage groups (3, 4, 5) covered 16.7 cM. It is interesting to note that the markers generated with the Satt 20 primer pairs clustered together on LG 1 and showed very tight linkage. In Akkaya *et al.* (1992) the Satt 20 locus was mapped to the LG 18. So it appears that LG 1 in this study corresponds to LG 18 of Akkaya *et al.* (1995).

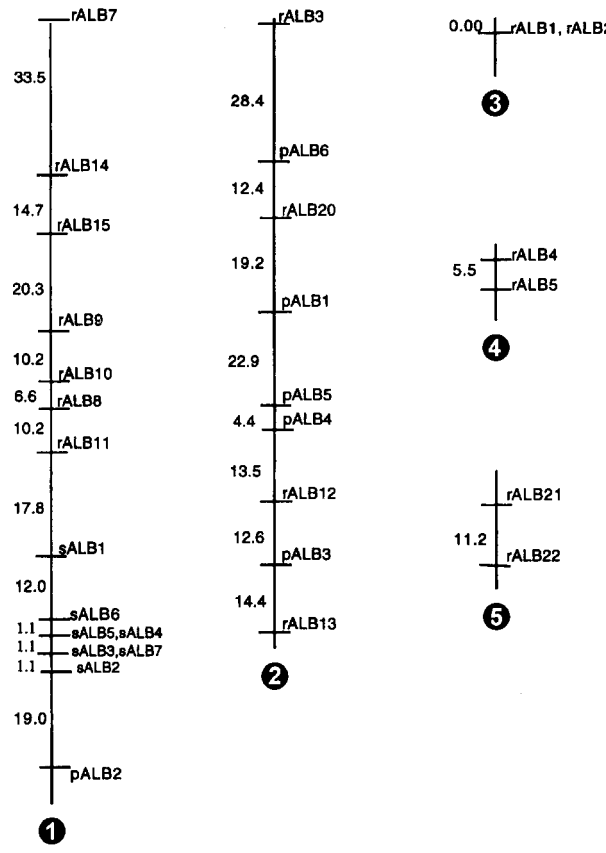

Fig. 4. A genetic linkage map of soybean. The five linkage groups are arranged on the basis of their length. Map distance in centi-Morgans (cM) is listed in the left and loci to the right of each linkage group. The distance between the markers was determined based on minimum LOD 4.00 and maximum distance of 30 cM.

The results obtained from the first experiment showed that primers synthesized from known gene sequences have the ability to detect soybean genomic DNA variation, suggesting that homologous sequences exist between soybean and these species. Although a large number of amplified fragments were generated by these specific gene primers, only six polymorphic bands were observed in this study. The low level of polymorphism

generated with these specific gene primers indicated that these sequences may be conserved or that the parental lines used to generate the mapping population were not highly polymorphic. One can accurately determine the level of genetic polymorphism in soybean by studying soybean germplasm collections using known sequence primers.

The results obtained with the SSR primer pairs indicated that only one pair of primers (Satt 20) was polymorphic. Although these primers had a gene diversity value of at least 0.8 in other studies, they failed to produce significant polymorphisms in our investigation. The lack of polymorphisms generated can be explained, either as a result of narrow genetic base between the parental lines or that the method used to detect polymorphisms was not sensitive enough to produce the expected level of polymorphism. Since silver staining was used in this study, it is likely that the low level of polymorphisms observed is due in part, to the lower sensitivity of this method.

The results obtained with RAPD primers produced the most polymorphisms. This was not surprising, since the RAPD assay targets the entire genome whereas specific and SSR primers target specific DNA sequences. Owing to their technical simplicity, low cost per data unit, and high levels of polymorphisms, RAPDs are the obvious choice of molecular markers for genetic maps. The reproducibility of RAPD markers was examined by re-amplifying and rescoring all 65 recombinant inbred lines and their parents BARC-2 (Rj_4) & PI 290,136. The results indicated that the RAPD assay is highly effective and convenient. Klein-Lankhorst *et al.* (1991) reported that more polymorphic fragments can be generated by using combinations of two primers. In this study, three two-primer combinations OPA 5 & OPA 18, OPA 18 & OPA 20, AND OPE 4 & OPE 10 produced 24 polymorphic bands.

A linkage map was constructed by *Mapmaker 3.0* (Lander *et al.* 1987) with 43 polymorphic markers, 37 of which were generated by the primers used, and the remaining six markers were morphological markers. The two-primer combination OPA 5 & OPA 18 produced 13 polymorphic markers that were distributed among only 3 linkage groups (LG 1, LG 2, LG 4). With the exception of the RAPD markers, it was apparent that the markers from the same primer tended to map to the same linkage group. For example all seven polymorphic markers (SALB1, SALB2, SALB3, SALB4, SALB5, SALB6, SALB7) generated from Satt 20 were included in LG 1, all three polymorphic markers (pALB4, pALB5, and pALB6) from Homeobox were included in LG 2 with 2 of these pALB 4 & 5 closely linked (4.4 cM), while pALB6 was more distantly located and unlinked to pALB 4 & 5. Two of *Hsp* markers were located on the same linkage group but sufficiently distant to be unlinked to each other, while a third *Hsp* marker was on LG 1. Both polymorphic markers (rALB1, rALB2) from OPA 18 & OPA 20 were included in LG 3, two markers (rALB4, rALB5) from OPA 5 & OPA 18 were included in LG 4 and two markers (rALB21, rALB22) from OPE 4 & OPE 10 were included in LG 5. All five linkage groups covered a total distance of 264.9 cM.

No linkage was detected between the molecular markers and six morphological traits. The results also showed that regardless of primers used several new bands were detected in some of the recombinant inbred lines. A similar observation was reported in barley (*Hordeum vulgare* L.) (Soliman 1994). The presence of non-parental bands can arise as a result of gene mutation, migration, recombination hot spots or non-specific amplification. However, since the assays were repeated twice, it is unlikely that non-specific amplification was the cause of the observed variation. It is more likely that the new variants observed were due to either gene migration or genetic recombination and not gene mutation since the variants occur at high frequency than can be accounted for by a single gene mutation.

Conclusion: The three types of primers (specific, SSR, and arbitrary primers) generated a total of 37 amplified products of which six polymorphic markers were generated by specific gene primers, seven polymorphic markers were generated by Satt 20, and 24 polymorphic markers were generated by arbitrary primers. When compared to all primers utilized in this study, the RAPD primer combination OPA 5 & OPA 18 contributed the maximum number of polymorphic markers, *i.e.* 13 markers.

A linkage map was developed covering a distance of 264.9 cM and consisting of five linkage groups. The map distance covered by the markers generated in this study represent only a fraction of the entire soybean linkage map (<3000 cM, 20 linkage groups). It should be noted that Satt 20 was mapped to LG 18 suggesting that all the markers mapped to LG 1 in this study may belong to LG 18 in the integrated soybean linkage map. There is a need to integrate all the markers obtained in this study into the soybean genetic linkage map. This integration is necessary before one can utilize this data in practical plant breeding programs.

References

Akkaya, M.S., Bhagwat, A.A., Cregan, P.B.: Length polymorphisms of simple sequence repeat DNA in soybean. - Genetics **132**: 1132-1139, 1992.

Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A., Cregan, P.B.: Integration of simple sequence repeat DNA markers into a soybean linkage map. - Crop Sci. **35**: 1439-

1445, 1995.

Barone, A., Ritter, E., Schachtschabel, U., Debener, T., Salamini, F., Gebhardt, C.: Localization by restriction fragment length polymorphism mapping in potato of a major dominant gene conferring resistance to the potato cyst nematode *Globoderaro stochiensis*. - Mol. gen. Genet. **224**: 177-182, 1990.

Beckman, J.S., Soller, M.: Restriction fragment length polymorphisms and genetic improvement of agricultural species. - Euphytica **35**: 1-11, 1986.

Berger, G.U., Minor, H.C.: An RFLP marker associated with resistance to *Phomopsis* seed decay in soybean PI 417,479. - Crop Sci. **39**: 800-805, 1999.

Bonierbale, M.W., Plaisted, R.L., Tanksley, S.D.: RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. - Genetics **120**: 1095-1103, 1988.

Byrum, J.R., Kimbirauskas, P.M., Shoemaker, R.C., Diers, B.W.: Identification of a RAPD marker linked to the *RPS*$_1$ gene in soybean (*Glycine max* (L.) Merr). - Soybean Genet. Newslett. **20**: 112-117, 1993.

Condit, R., Hubbell, S.P.: Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. - Genomes **34**: 66-71, 1991.

Fuller, F., Kunster, P.W., Nguyen, T., Verma, D.P.S.: Soybean nodulin genes: Analysis of cDNA clones reveals several major tissue-specific sequences in nitrogen fixing root nodules. - Proc. nat. Acad. Sci. USA **80**: 2594-2598, 1983.

Garvin, D.F., Miller-Garvin, J.E., Viccars, J.V., Brown, A.H.D.: Identification of molecular markers linked to *ant*28-484 a mutation that eliminates proanthocyanidin production in barley seeds. - Crop Sci. **38**: 1250-1255, 1998.

Ha, D.B.D, Buffard, D., Berger, F., Breda, C., Esnault, R.: Nucleotide sequence encoding a slow allele of *Adh 1* in pearl millet. - Plant mol. Biol. **14**: 453-455, 1990.

Helentjaris, T., Slocum, M., Schaefer, A., Nienhuis, J.: Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. - Theor. appl. Genet. **72**: 761-769, 1986.

Key, J.L., Lin, C.Y., Chen, Y.M.: Heat shock proteins of higher plants. - Proc. nat. Acad. Sci. USA **78**: 3526-3530, 1981.

Klein-Lankhorst, R.M., Vermunt, A., Weide, R., Liharrska, T., Zabel, P.: Isolation of molecular markers for tomato (*L. esculentum*) using random amplified polymorphic DNA (RAPD). - Theor. appl. Genet. **83**: 108-114, 1991.

Kolchinsky, A., Gresshoff, P.M.: Soybean telomeres as potential molecular markers. - Soybean Genet. Newslett. **20**: 190-196, 1993.

Lander, E,S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., Newburg, L.: Mapmaker: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. - Genomics **1**: 174-181, 1987.

Landry, B.S., Michelmore, R.W.: Methods and applications of restriction fragment length polymorphism analysis to plants. - In: Bruening, G., Harada, J.J., Hollaender, A. (ed.): Tailoring Genes for Crop Improvement. An Agricultural Perspective. Pp. 25. Plenum Press, New York 1987.

Lark, K.G., Weisemann, J.M., Matthews, B.F., Palmer, R., Chase, K., Macalma, T.: A genetic map of soybean (*Glycine max* L.) using an intraspecific cross of two cultivars: 'Minsoy' and 'Noir 1'. - Theor. appl. Genet. **86**: 901-906 1993.

Lauzon, L.M., Helm, K.W., Elizabeth, Y.A.: cDNA clone from *Pisum sativum* encoding a low molecular weight heat shock protein. - Nucl. Acids Res. **8**: 4274, 1990.

Mansur, L.M., Lark, K.G., Kross, H., Oliveira, A.: Interval mapping of quantitative trait loci for reproductive, morphological and seed traits of soybean (*Glycine max* L.). - Theor. appl. Genet. **86**: 907-913, 1993.

McCouch, S.R., Kochert, G., Yu, Z.H., Wang, Z.Y., Khush, G.S., Coffma, W.R., Tanksley, S.D.: Molecular mapping of rice chromosomes. - Theor. appl. Genet. **76**: 815-829, 1988.

Mesfin, A., Frohberg, R.C., Anderson, J.A.: RFLP markers associated with high grain protein from *Triticum turgidum* L. var dicoccoides introgressed into hard red spring wheat. - Crop Sci. **39**: 508-513, 1999.

Ovidio, R.D., Tanzarella, O.A., Porceddu, E.: Rapid DNA amplification by polymerase chain reaction. - Plant mol. Biol. **15**: 169-171, 1990.

Provan, J., Powell, W., Waugh, R.: Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). - Theor. appl. Genet. **92**: 1078-1084, 1996.

Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M., McGinnis, M.: Homeobox genes of the antennapedia and bithorax complexes of *Drosophila*. - Cell **43**: 71-80, 1985.

Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U., Cregan, P.W.: The use of microsatellite DNA markers for soybean genotype identification. - Theor. appl. Genet. **90**: 43-48, 1995.

Saghai-Maroof, M.A., Soliman, K.M., Jorgensen, R.J., Allard, R.W.: Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance chromosomal location and population dynamics. - Proc. nat. Acad. Sci. USA **81**: 8014-8017, 1984.

Sarfatti, M., Katan, J., Fluhr, R., Zamir, D.: An RFLP marker in tomato linked to the *Fusarium oxysporum* resistance gene 12. - Theor. appl. Genet. **78**: 755-759, 1989.

Skroch, P.W., Nienhuis, J.: Qualitative and quantitative characterization of RAPD variation among snap bean (*Phaseolus vulgaris*) genotypes. - Theor. appl. Genet. **91**: 1078-1085, 1995.

Soliman, K.M.: Multilocus structure in barley Composite Cross II populations. - Egypt. J. Genet. Cytol. **23**: 147-161, 1994.

Szczyglowski, K., Legocki, A.B.: Isolation and nucleotide sequence of a cDNA clone encoding nodule-specific (hydroxyl) proline-rich protein LENOD2 from yellow lupin. - Plant mol. Biol. **15**: 361-363, 1990.

Taylor-Grant, N., Soliman, K.M.: Detection of polymorphic DNA and taxonomic relationships among 10 wild perennial soybean species using specific and arbitrary nucleotide primers. - Biol. Plant. **42**: 25-37, 1999.

Welsh, J., Honeycutt, R.J., McClelland, M., Sobral, B.W.S.: Parentage determination in maize hybrids using the arbitrarily primed polymerase chain reaction (AP-PCR). - Theor. appl. Genet. **82**: 473-476, 1991.

Williamson, V.M., Ho, J.Y., Wu, F.F., Miller, N., Kaloshian, I.: A PCR based marker tightly linked to the nematode resistance gene *Mi*, in tomato. - Theor. appl. Genet. **87**: 757-763, 1993.

Yu, Z.H., Mackill, D.J., Bomann, J.M., Tanksley, S.D.: Tagging genes for blast resistance in rice via linkage to RFLP markers. - Theor. appl. Genet. **81**: 471-476, 1991.

Yu, Z.H., Stall, R.E., Vallejas, C.E.: Detection of genes for resistance to common bacterial blight of beans. - Crop Sci. **38**: 1290-1296, 1998.

Zhang, K., Soliman, K.M., Sapra, V.: Polymerase chain reaction from known function genes detects polymorphisms and trait associations in soybean genetic resources. - Biol. Plant. **40**: 43-53, 1997.

Zhao, X, Kochert, G.: Phylogenetic distribution and genetic mapping of a $(GGC)_n$ microsatellite from rice (*Oryza sativa* L.). - Plant mol. Biol. **21**: 607-614, 1993.