Theor Appl Genet (1998) 96: 1086–1096

G. Wang · R. Mahalingam · H. T. Knap

(C-A) and (G-A) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, *Glycine max* (L.) Merr.

Received: 24 July 1997 / Accepted: 17 November 1997

Abstract We used thirty simple sequence repeats (SSRs) with a variable 2–4 base 'anchor' at their 5′ ends (ASSRs) independently or with arbitrary primers in analysis of soybean germplasm and the intercross of 'Essex' and PI 437654. $(AG)_6$, $(GA)_6$ or $(CT)_6$ and $(GT)_6$, $(TG)_6$ or $(CA)_6$ were efficient in the detection of (G-A) and (C-A) ASSR-generated polymorphisms, respectively. DNA sequence analysis of the ASSR-amplified fragments confirmed the presence of SSR sequences. (A-T) ASSRs failed to give amplification or generated fewer number of fragments. Only one of the four tested decamer primers altered ASSR banding patterns in the soybean. All the six long primers (18–20 mer) tested changed ASSR banding profiles. On average, seven polymorphic fragments per ASSR primer were produced in soybean germplasm and four in the intraspecific cross of 'Essex' and PI 437654. The grouping of 48 genotypes in UPGMA analysis using four (C-A) and four (G-A) ASSR primers was consistent with the classification obtained with RFLP markers. Seventy-seven (91%) ASSR markers were dominant, while the remaining 8 (9%) showed codominant segregation. Fifty-eight ASSR markers were mapped onto 18 RAPD/RFLP linkage groups, which covered approximately 50% of the soybean genome. Of the (G-A) ASSR-derived markers 49% remained unlinked compared with 17% of (C-A) ASSR markers at LOD 3.0. Map linkage information showed that the assigned (C-A) polymorphisms had a biased distribution, whereas (G-A) polymorphisms were randomly dispersed.

Key words Anchor · Short sequence repeat · Polymorphism · Linkage · Soybean

Communicated by M. A. Saghai Maroof

G. Wang · R. Mahalingam
Department of Agronomy, E272 Poole Agric. Ctr. Bldg., Clemson University, Clemson, SC 29634, USA

H. T. Knap (✉)
Department of Agronomy and Department of Biological Sciences, E272 Poole Agric. Ctr. Bldg., Clemson University, Clemson, SC 29634, USA
Fax: 864 656 3443
E-mail: hskrpsk@clemson.edu

Introduction

Simple sequence repeats (SSRs) or microsatellites are tandem arrays of short (1–6) nucleotides that are ubiquitous in most eukaryotic genomes. They have received considerable attention as highly informative molecular markers in human, mammalian and plant genetic studies (Tautz 1989; Weber and May 1989; Weber 1990; Lagercrantz et al. 1993; Maughan et al. 1995). In soybean, identification of SSRs has been accomplished by screening genomic libraries and sequence data analysis. As in other plant species, $(AT)_n$, $(ATT)_n$ or $(AAT)_n$ tandem repeats are abundant in the soybean genome and are thus the most frequently detected (Cregan et al. 1994; Morgante et al. 1994; Rongwen et al. 1995; Powell et al. 1996; Yu et al. 1996). Cregan (1997) isolated more than 300 SSRs with $(ATT)_n$ or $(TAA)_n$ repeats. GenBank search for SSR sequences in soybean by Akkaya et al. (1992) identified only 4 $(CA/GT)_n$ and 5 $(TC/AG)_n$ repeats with *n* ranging from 5 to 8, in contrast to 17 $(AT/TA)_n$ with *n* ranging from 5 to 27. They suggested that the length of $(CA/GT)_n$ and $(TC/AG)_n$ repeats might hinder their usefulness as polymorphic markers in soybean.

A difficulty in microsatellite identification is the requirement of sequence information to design locus-specific primers. Wu et al. (1994) developed a methodology using anchored SSR (ASSR) primers which permit the detection of microsatellites without the need for cloning or sequencing. ASSR primers also facilitate identification of less abundant microsatellites in

genomes. The anchor of the ASSR primer is usually a 2- to 4-nucleotide sequence attached to the 5′ end. It is used to prevent slippage and increase the specificity in polymerase chain reaction (PCR) cycles. ASSR primers can also be used in combination with arbitrary primers, which may further increase their versatility and informativeness (Wu et al. 1994). Since the ASSR primer alone is labeled, the amplification products containing SSRs are detected. In the study presented here, we describe detection of $(GT)_n$, $(TG)_n$, $(CA)_n$, $(AG)_n$, $(GA)_n$, and $(CT)_n$ repeats in soybean using ASSR primers. The experiments reported here were designed to: (1) assess the polymorphisms generated in soybeans by ASSR primers solely or in combination with arbitrary primers; (2) confirm the presence of SSRs by sequencing the amplified fragments; (3) investigate the usefulness of ASSR markers for genotype identification; and (4) determine the mode of inheritance of ASSR-generated polymorphisms and linkage assignment in the soybean genome.

Materials and methods

Plant materials

Seven soybean genotypes, a breeding germplasm, J87-233; cvs 'Hutcheson', 'Essex' and 'Peking'; and plant introductions (PIs) PI 437654, PI 90763 and PI 88788, were used for initial tests to examine ASSR-generated polymorphisms. Forty-eight soybean PIs and cultivars listed in Table 1 were used to assess the suitability of ASSR markers in determining genetic relatedness in the soybean. These genotypes were selected based on their differential response to the soybean cyst nematode (SCN), *Heterodera glycines* I., maturity group (MG) and seed coat color. Seeds of the PIs and cultivars were obtained from the USDA Soybean Germplasm Collection (USDA-ARS, Urbana-Champaign, Ill.). The inheritance study and linkage assignment of the polymorphic markers were conducted on a population of 65 F_2 progeny of the 'Essex' and PI 437654 cross.

Oligonucleotide primers and end-labeling

Based on the nucleotide content, three types of ASSRs were used: (A-T) ASSRs were represented by $(AT)_6$, $(ATT)_5$, and $(TAA)_4$ repeats; (G-A) ASSRs by $(GA)_6$, $(AG)_6$, and $(CT)_6$ repeats; and (C-A) ASSRs by $(GT)_6$, $(TG)_6$ and $(CA)_6$ repeats (Table 2). Twenty-eight ASSR primers were designed by adding different anchors (2–4 nucleotides) at the 5′ end of the repeats. Two ASSR primers, K2 and K3 (Table 2), were anchored at the 3′ end. ASSR primers were synthesized by Integrated DNA Technologies (Coralville, Iowa). End-labeling was performed using 250 ng of ASSR primer at 37°C for 30 min in a 50-µl volume containing 5 µl $10 \times T_4$ Kinase buffer, 2.5 µl (92.5×10^4 Bq) ATP [γ-$^{33}P^+$] and 5 units T_4 polynucleotide Kinase. This mixture was used for 50 PCR reactions.

ASSR primers K5, K8, K15, K17, K18 or K22 were used in combination with eleven arbitrary primers of different length and sequence. The arbitrary primers were the decamers B15 (GGAGGGTGTT), C08 (GATGACCGCC), H05 (AGTCGTCCCC), S07 (TCCGATGCTG) (Operon Technologies, Alameda, Calif.); a 14-mer primer CDP2 (GCNAANGGIAATCC) (Grant et al. 1995); five 18- to 20-mer primers (Udvardi et al. 1993; Grant et al. 1995); and an amplified fragment length polymorphism (AFLP) primer MACC (Vos et al. 1995). The sequences of the 18- to 20-mer primers are given in Table 2.

DNA extraction, PCR and gel analysis

DNA was extracted from the leaves of 15-day-old plants using a modified CTAB (Hexadecyltrimethylammonium bromide) method (Keim et al. 1988). PCR was performed in a 10-µl volume containing

Table 1 Soybean plant introductions and cultivars used for anchored simple sequence repeat (ASSR) analysis

Number	Genotype	Origin	Number	Genotype	Origin
1	Hutcheson	USA	25	PI 91138	China
2	Haskell	USA	26	PI 92720	China
3	Dillon	USA	27	PI 200495	Japan
4	Cloud	China	28	PI 303652	China
5	Columbia	China	29	PI 339868B	Korea
6	Ilsoy	Korea	30	PI 437654	Russia
7	Patoka	China	31	PI 437655	Russia
8	Peking	China	32	PI 437679	Russia
9	Sooty	USA	33	PI 437690	Russia
10	PI 54591	China	34	PI 437770	Russia
11	PI 79609	China	35	PI 438183	Russia
12	PI 79693	China	36	PI 438496B	Russia
13	PI 84751	China	37	PI 438498	Russia
14	PI 398682	Korea	38	PI 438503A	Russia
15	PI 404166	Russia	39	Chesapeake	USA
16	PI 407944	Korea	40	Marcus	USA
17	PI 417091	Japan	41	Kenwood	USA
18	PI 417094	Japan	42	BSR101	USA
19	PI 437090	Russia	43	Asgrow 2506	USA
20	PI 437488	Russia	44	Dimon	USA
21	PI 88788	China	45	NKs 19-90	USA
22	PI 89014	China	46	Williams 82	USA
23	PI 89772	China	47	J87-233	USA
24	PI 90763	China	48	Essex	USA

Table 2 Sequences and number of amplification products generated by ASSR primers

ASSR		Sequences (5′ → 3′)[a]	Number of fragments	Polymorphic fragments	Core
Type	Name				
(A-T)	K4	$SG(AT)_6$	0	0	$(AT)_6$
	K11	$CCAG(AT)_6$	0	0	$(AT)_6$
	K12	$SG(ATT)_5$	19	0	$(ATT)_5$
	K14	$SG(TAA)_4$	6	0	$(TAA)_4$
	K34[b]	$GAGC(TAA)_4$	6	0	$(TAA)_4$
(C-A)	K1	$HVH(CA)_6$	31	10	$(CA)_6$
	K10	$CCG(CA)_6$	23	7	$(CA)_6$
	K17	$BDB(CA)_6$	19	6	$(CA)_6$
	K5[b]	$CCAG(GT)_6$	21	7	$(GT)_6$
	K6[b]	$CCC(GT)_6$	26	9	$(GT)_6$
	K16	$VHV(GT)_6$	26	8	$(GT)_6$
	K20	$CAA(GT)_6$	26	6	$(GT)_6$
	K15	$HVH(TG)_6$	28	10	$(TG)_6$
(G-A)	K7[b]	$CAA(CT)_6$	22	4	$(CT)_6$
	K18	$VHV(CT)_6$	23	8	$(CT)_6$
	K21	$CGG(CT)_6$	24	8	$(CT)_6$
	K23	$CCG(CT)_6$	22	7	$(CT)_6$
	K24	$CCAG(CT)_6$	24	7	$(CT)_6$
	K25	$CCA(CT)_6$	23	5	$(CT)_6$
	K26	$RY(CT)_6$	17	5	$(CT)_6$
	K29[b]	$GAAT(CT)_6$	22	7	$(CT)_6$
	K46	$TCGC(CT)_6$	20	5	$(CT)_6$
	K8[b]	$CCT(GA)_6$	28	8	$(GA)_6$
	K9[b]	$CCC(GA)_6$	29	7	$(GA)_6$
	K22	$CCAG(GA)_6$	30	5	$(GA)_6$
	K28[b]	$GCTTG(GA)_6$	30	6	$(GA)_6$
	K30[b]	$GCCA(GA)_6$	20	5	$(GA)_6$
	K19	$CCC(AG)_6$	31	6	$(AG)_6$
ASSR/arbitrary primer	K5/MACC	K5/GATGAGTCCTGAGTAAACC	28	8	
	K8/IDHF	K8/TCGCCAACCCCATCGTCG	29	9	
	K8/P34R	K8/TGGAGCTCTGTACCATAATG	29	8	
	K15/P34F	K15/GGCGAGGGAACATACGGCG	30	7	
	K15/IDHF	K15/TCGCCAACCCCATCGTCG	25	6	
	K17/PLP2	K17/ATGGGNGGNGTNGGIAAGAC	28	11	
	K17/IDHF	K17/TCGCCAACCCCATCGTCG	31	10	
	K17/IDHR	K17/GGCGGGTAACTGTGCCGT	23	5	
	K18/PLP2	K18/ATGGGNGGNGTNGGIAAGAC	30	12	
	K18/P34F	K18/GGCGAGGGAACATACGGCG	26	8	
	K22/PLP2	K22/ATGGGNGGNGTNGGIAAGAC	26	6	
3′ Anchored primers	K2	$(CA)_6RG$	15	2	$(CA)_6$
	K3	$(CA)_6RY$	17	2	$(CA)_6$

[a] B = C, G, T; D = A, G, T; H = A, C, T; S = C, G; R = A, G; V = A, C, G; Y = C, T; N = A, C, G, T; I = Inosine
[b] The ASSR primers were synthesized according to Wu et al. (1994)

50 ng template DNA, 5 ng labeled ASSR primer (1 µl from the 50 µl of labeled mixture), 45 ng non-labeled ASSR primer (1 : 9 ratio of labeled and non-labeled ASSR primer), 50 ng arbitrary primer, 0.2 m*M* of each dNTP, 0.25 units of Ampli-*Taq* DNA Polymerase (Perkin Elmer Cetus), 10 m*M* TRIS-HCl, 50 m*M* KCl and 2.0 m*M* $MgCl_2$. The PCR thermal profile was according to Wu et al. (1994).

Two and a half microliters of each PCR reaction was mixed with 2.5 µl of loading buffer [98% formamide, 10 m*M* EDTA, 0.1% (w/v) bromophenol blue and 0.1% (w/v) xylene cyanol], denatured at 90°C for 3 min, cooled on ice briefly and immediately loaded on a 5% or 6% denaturing polyacrylamide (PA) gel. Electrophoresis was performed at 60 W in 1 × TBE buffer for 2.5 h. The gel was vacuum-dried at 60°C for 1 h and exposed to X-ray film (Bio MAX MR, Eastman Kodak, Rochester, N.Y.) at − 80°C from 5 h to 2 days. The ATP[γ-$^{33}P^{+}$] end-labeled 100-bp DNA ladder (Promega, Madison, Wis.) was used to estimate the size of the amplified fragments.

Cloning and sequencing of amplified fragments

ASSR-generated fragments were cut from dried gels using a scalpel blade, transferred into Eppendorf tubes each containing 40 µl of distilled water, placed at 90°C for 10 min, and used directly in reamplification reaction in a final volume of 50 µl under the same PCR conditions.

Five microliters of the PCR reaction was loaded on a 1% agarose gel to verify the sizes of the reamplified fragments. The remaining mix was purified using Wizard PCR Preps DNA purification kit (Promega, Madison, Wis.) and cloned into pGEM-T Vector (Promega, Madison, Wis.) according to the manufacturer's instructions. DNA sequencing in both forward and reverse directions was done using the ABI PRISM dye terminator cycle sequencing kit (Perkin Elmer Cetus) with T7 and SP6 primers and analyzed in an ABI 373 Automated DNA Sequencer (Perkin Elmer Cetus) at the sequencing facility at Clemson University. Sequencing data were analyzed by MACDNASIS V2 (Hitachi software Engineering Co.) and compared with the sequences of GenBank and EMBL databases.

Data analyses

For an estimate of the genetic relationships among the 48 soybean PIs and cultivars, monomorphic fragments were included, as these shared bands are an indication of genomic similarity. Polymorphic fragments were scored as present or absent. A pairwise similarity matrix was calculated using a simple matching (SM) coefficient. A SM coefficient is defined as the number of identical present and absent alleles shared by two genotypes divided by the total number of comparisons (Sneath and Sokal 1973). Cluster analysis based on the SM coefficient values was done using the unweighted paired group method analysis (UPGMA) in the numerical taxonomy and multivariate analysis system, NTSYS-PC, version 1.80 (Rohlf 1994). Chi-square tests were conducted for segregating ASSR markers. Linkage analysis was performed using MAPMAKER/exp, Version 3.0, (Lander et al. 1987). Genetic distance estimation was based on Kosambi function (Kosambi 1944). A minimum LOD score of 3.0 and maximum distance of 50 centiMorgans (cM) were used in map construction.

Results

Identification of informative ASSR primers

Thirty ASSR primers, containing either different SSRs and an identical anchor sequence or the same SSR with different anchor sequences, were used in amplification reactions. The primers with $(AT)_n$, K4 or K11, failed to give any amplification product (Table 2). Primers with $(ATT)_n$ or $(TAA)_n$ produced weakly amplified fragments and monomorphic banding patterns (Table 2). Therefore, we suggest that (A-T) repeats are not suitable for the ASSR methodology in soybean. We used six dinucleotide repeats: (C-A) ASSRs were $(CA)_6$, $(GT)_6$, $(TG)_6$, and (G-A) ASSRs were $(AG)_6$, $(GA)_6$, and $(CT)_6$ (Table 2). Amplified DNA fragments ranged in size from about 100 to 900 bp. The total number of DNA fragments and polymorphic fragments generated from the ASSR primers are shown in Table 2. Excluding 5 (A-T) ASSR primers that gave no polymorphisms and 2 primers that were 3′ anchored, 23 ASSR primers produced, on average, seven polymorphisms per primer (from K1 to K19, in Table 2).

Differences were observed in the size of the amplified fragments and levels of polymorphisms produced by (G-A) and (C-A) ASSR primers. Generally, (G-A) ASSR primers amplified fragments up to 900 bp, and a 5% PA gel was sufficient for separation of these products. For (C-A) ASSR primers, the amplified fragments ranged from 100 to 700 bp in size and were better resolved on 6% gels. On average, 8 polymorphisms were generated from (C-A) ASSR primers and 6 from (G-A) ASSR primers. Differences between (G-A) and (C-A) ASSR primers were also observed in map assignment (discussed in 'Inheritance and linkage assignment'). A comparison of the polymorphisms produced by (G-A) and (C-A) primers is presented in Table 4.

Importance of the anchor sequences

We have tested several ASSR primers with the same dinucleotide repeat but with one-, two- or three-base change in the 5′ anchor sequence (Table 2). A single-base change in the anchor sequence did not significantly modify the banding pattern produced by the primers (i.e. K21, K23, K24 and K25; K8 and K9). The introduction of a two- or three-base change in the anchor sequence (i.e. K5 and K6; K7 and K21) resulted in a significant alteration of the banding patterns. ASSR primers with four-base 5′ anchors (i.e. K5, K30, K46) produced fewer amplified fragments, though the level of polymorphism was similar to those primers with three-base 5′ anchors. The introduction of degenerate anchor sequences, HVH and VHV, increased both the total number of amplified and polymorphic fragments (Fig. 1A, Table 2).

ASSR primer in combination with arbitrary primers

When 6 ASSR primers were used in combination with 4 decamer primers or a 14-mer primer, the banding patterns were nearly identical with those obtained with the ASSR primer alone. Only in the case of primer combination K5 and decamer H05, were 2 novel polymorphisms revealed (data not shown). However, when the ASSRs were used in combination with longer (18–20 mer) primers, the banding patterns changed and new polymorphisms were identified. ASSR K18 primer produced 8 polymorphic fragments and 12 when used in combination with PLP2 primer (Fig. 1B). In some primer combinations (K17/PLP2, K17/IDHF), the number of polymorphic fragments nearly doubled (Table 2).

DNA sequencing of the amplified fragments

Two amplified ASSR fragments, obtained from the combination of ASSR primer K18 and arbitrary primer PLP2, were selected for sequencing (Fig. 1B). One was a strong monomorphic fragment, and the other was a polymorphic fragment of medium intensity (Fig. 1B).

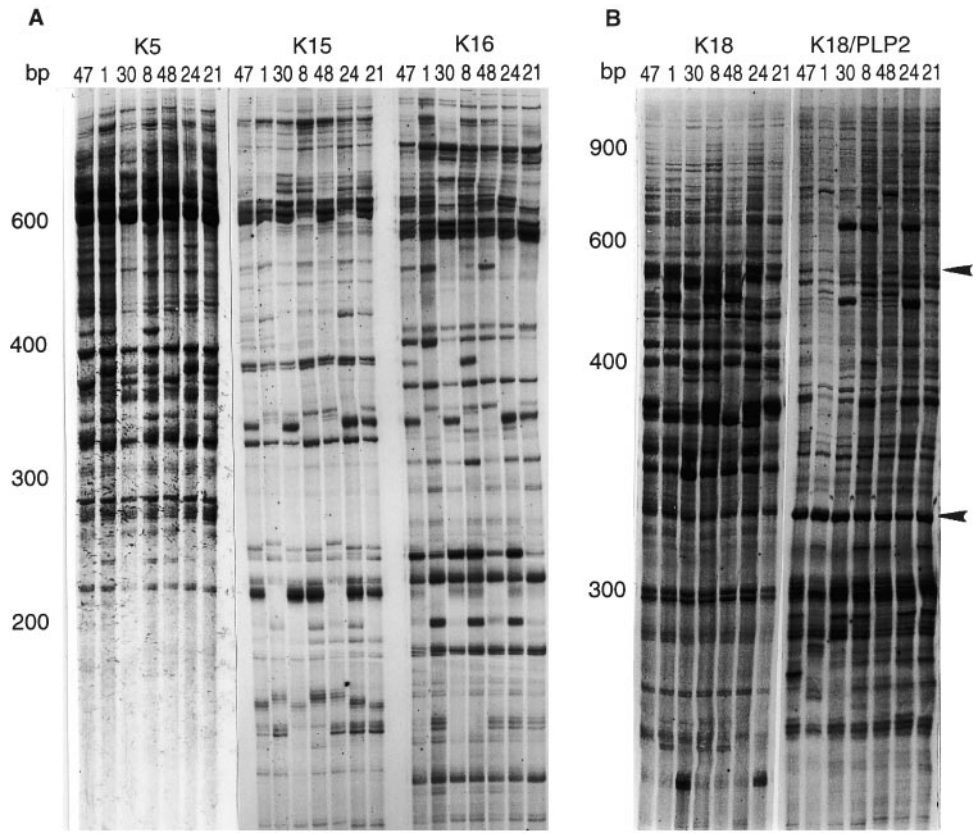

Fig. 1 **A** Banding profiles of soybean DNA obtained with ASSR primers K5, K15 or K16 and resolved on 6% denaturing polyacrylamide gel. **B** Banding profiles of soybean DNA obtained with ASSR primer K18 independently or in combination with primer PLP2 and resolved on 5% denaturing polyacrylamide gel. *Arrows* indicate the two sequenced fragments. *Numbers* on *left* indicate molecular size of the 100-bp DNA ladder (Promega, Madison, Wis.). The soybean genotypes are identified by the numbers in Table 1. The primer sequences are reported in Table 2

The polymorphic fragment was present in all four black seed coat genotypes (i.e. 'Peking', PI 437654, PI 88788 and PI 90763), but absent from the three yellow-seeded genotypes (J87-233, 'Hutcheson' and 'Essex'). The polymorphic fragment was 541 bp, and the monomorphic fragment had 321 bp. Both fragments contained the K18 and PLP2 primer sequences (Fig. 2). No common internal sequences between the two fragments were present. The sequence of the polymorphic fragment lacked homology to any DNA sequences in the GenBank database. The sequence of the monomorphic fragment contained a single open reading frame encoding 106 amino acids. The deduced amino acid sequence showed 50% homology to the receptor-like protein kinase precursor in *Arabidopsis* (Walker 1993), based on the BLAST search and DNAsis maximum matching analysis (Hitachi Software Engineering Co. Japan).

A

K18

GAGCTCTCTCTCTCTCCCACACTCTCCATCTTCTCGTCACCCCCTCCCAAACCCTAACCT 60
CACTTACCAATTCTCTCTCCCAAACCCTAACCTCACACCTCCACTCCACACCCTCCTCCC 120
TCCGCTCCCCTCTCCTCACCATCCCTCACTCCTCTATTGACCAAATCATCCAAAACCACT 180
TCATCACACAAAACCCTAACCCTAACCCTAACCAAGTCCATCTCTACCTCCTCAACCTCC 240
CTCCCGCTTCCTCTTCCAATTCCAAAGCCTATGCTTACTCATACTCCCCCGGCGATTCTT 300
CCGCTGCTGTGACGAAATGCTCCGGCACTTTCTTCACCTCCTCCCACCGCTACTTCTGGA 360
TCGACCTCCGCGCCGGCCCCGTCGACTACGGCCCCGCGATCTCCGGCGACGGAGTCATCC 420
CCCGCGGCGAGTTCCACCCTCTCGCCGCGCTCCACGGCCGTCCGAAGTCCAACAAGGCCT 480
TCGCCGCCGACATCGCCTCCGTCGTCTGGAGCGCCTACCACGTCTTCCCCACTCCCCCCA 540
T PLP2 541

B

K18

GCGCTCTCTCTCTCTAACCCAAGACGGTCTCTTTCTCCTCGAAGCAAGACGGCACCTTTC 60
CGACCCGGAAAATGCCCTCTCCTCCTGGAACCCTGCCGCCACCACCCCATGCCGTTGGCG 120
CAGCGTCACCTGCGACCCCCTCACCGGCGCCGTCACCTCCGTGAGCCTCCCCAACTTTTC 180
TCTCTCCGGCCCCTTCCCCGCCGTCCTCTGTCGAATCGCCTCCCTGACCACCCTCAACCT 240
CGCCTCCAACCTCATCAACTCCACCCTCTCCGCCGTCGCCTTCGCTGCGTGCCGCAACCT 300
CGTCTTCCCAACCCCACCCAT 321
PLP2

Fig. 2A, B Nucleotide sequences of amplification products obtained with ASSR primer K18 and primer PLP2 in soybean. **A** Sequence of the polymorphic fragment, 541 nt; **B** sequence of the monomorphic fragment, 321 nt. ASSR primer K18 and arbitrary primer PLP2 sequences are *underlined*

ASSR-generated polymorphism for soybean genotypes' relatedness

Analysis of the 48 soybean genotypes with 4 (C-A) and 4 (G-A) ASSR primers (K1, K15, K16, K17, and K9, K18, K23, K28, respectively) identified 82 polymorphic fragments. An example of molecular polymorphisms in soybean germplasm revealed with one of the primers is shown on Fig. 3. Two of the ASSR primers, K17 and K18, were used in combination with a 20-mer PLP2, because of the high number of polymorphic fragments produced. Genomic similarity coefficients (based on SM values) among the 48 soybean PIs and cultivars were between 0.75 and 0.98 as estimated by numerical analysis of both the monomorphic

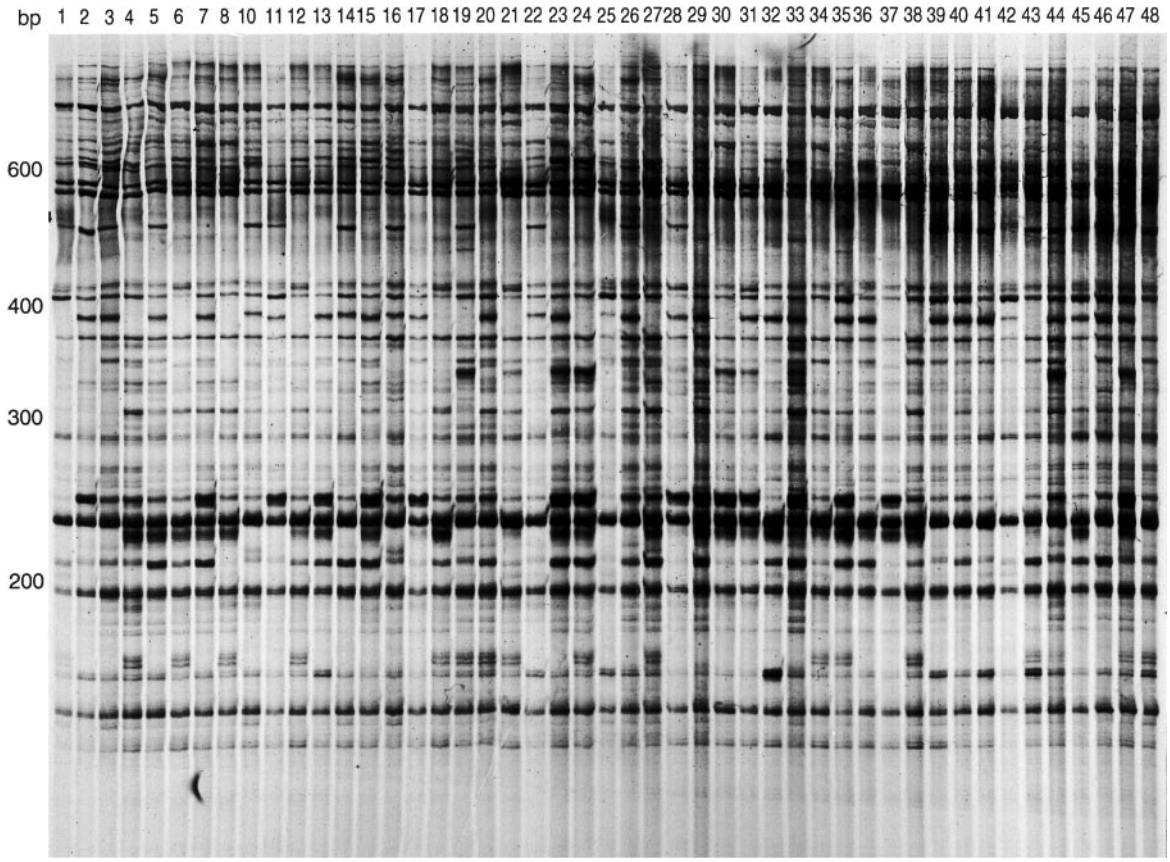

Fig. 3 Banding profiles of soybean DNA generated by ASSR primer K16 resolved on 6% denaturing polyacrylamide gel. The soybean genotypes are identified by the numbers in Table 1. *Numbers* on the *left* indicate molecular size of the 100-bp DNA ladder (Promega, Madison, Wis.)

and polymorphic fragments. Even within this high level of similarity, all genotypes could be differentiated by ASSRs. In the UPGMA, two major clusters were found: one from 'Hutcheson' to PI 91138 containing 19 genotypes and the second from 'Cloud' to J87-233 containing 29 genotypes (Fig. 4). Genotypes in the top cluster of the dendrogram were mostly susceptible (S) or moderately susceptible (MS) to several soybean cyst nematode races, *H. glycine* I., and were predominantly yellow seeded. Genotypes in the second large cluster were mostly resistant (R) or moderately resistant (MR) and predominantly black seeded. Association between clusters and maturity group was observed only in the group from 'Ilsoy' to PI 438503A, in which 5 of the 6 genotypes were MG III.

Inheritance and map assignment of ASSR markers

Twelve ASSR primers and 10 ASSR/arbitrary primer combinations were tested on the F_2 progeny of the 'Essex' and PI 437654 cross. A total of 85 polymorphic ASSR markers were used in inheritance and linkage tests. Seventy-seven markers (91%) were dominant and 8 (9%) were codominant. PI 437654 and 'Essex' contributed nearly equal number of dominant markers; 39 dominant markers were from 'Essex' and 38 markers were from PI 437654. The segregation pattern of 83 markers showed a good fit to a 3:1 ratio for dominant inheritance or 1:2:1 ratio for codominant inheritance based on the chi-square test. One marker, K8/PLP2$_{360}$, segregated in a 1:1 ratio. Segregation patterns were also confirmed in the F_2 progenies from the cross J87-233 × 'Hutcheson' (data not shown). Chi-square test results for the 58 markers, for which linkage assignments have been established, are in Table 3.

Linkage groups (LGs) constructed for the 'Essex' and PI 437654 F_2 population contain restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD) and ASSR marker loci (Fig. 5). Due to an insufficient level of RFLP polymorphism of the 170 genomic clones tested, only 5 linkage groups (LGs A, C, E, G and K) were associated with the soybean molecular linkage map (Shoemaker and Specht 1995). The total genetic distance of the ASSR/RAPD/RFLP map was 1377 cM, which covers approximately 50% of the soybean genome. Fifty-eight ASSR markers (68% of the total 85 analyzed) were mapped to 18 linkage groups (Fig. 5). Seven ASSR

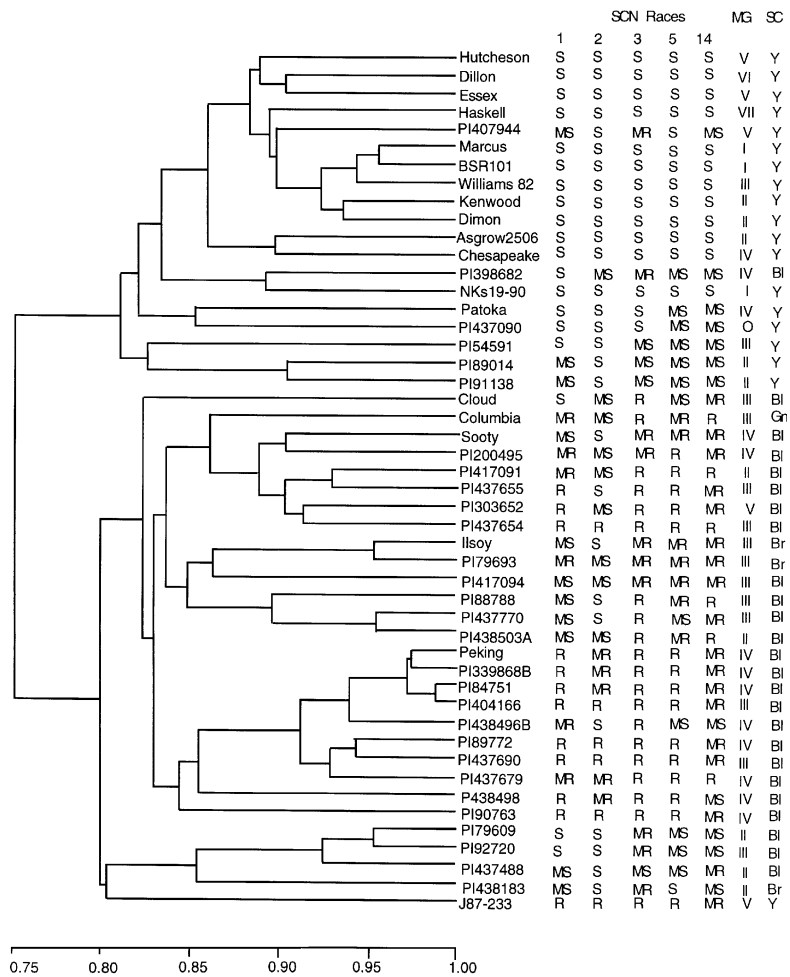

Fig. 4 UPGMA dendrogram of the 48 soybean PIs and cultivars analyzed with eight ASSR primers. Included are: the resistance responses to the five soybean cyst nematode races (*MS* moderately susceptible, *MR* moderately resistant, *R* resistant, *S* susceptible), *MG* Maturity group, *SC* seed coat color (*Bl* black, *Br* brown, *Gn* greenish, *Y* yellow)

markers were mapped to LG A of the soybean molecular linkage map that harbors the *i* locus for seed coat color (Shoemaker and Specht 1995). Five ASSR markers were mapped to linkage group C of the public molecular soybean map, which harbors the *t* locus for pubescence color (Fig. 5). Four ASSR loci were assigned to LG G.

Seventeen ASSR markers (29% of the 58 linked markers) were mapped within distances of less than 5 cM. Fifteen of them (88%) originated from (C-A) ASSRs, $(GT)_6$, $(TG)_6$ and $(CA)_6$, and only 2 markers (12%) were from (G-A) ASSRs, $(GA)_6$, $(AG)_6$, and $(CT)_6$ (Table 4). Biased distribution of markers produced by (C-A) ASSR primers was found on linkage group C and groups 4 and 7 (Fig. 5). All the ASSR markers in the three major clusters on these linkage groups were in coupling phase. Markers generated from (G-A) ASSR primers were distributed among 14 different linkage groups with no obvious clustering (Fig. 5). Twenty-seven markers (32% of the total 85) remained unlinked at LOD 3.0, of which 19 markers were produced by (G-A) ASSR primers and 8 by (C-A) ASSR primers (Table 4).

Discussion

The usefulness of molecular markers relies on their simplicity, cost-effectiveness and the genomic information they provide. Our results demonstrate that (C-A) and (G-A) ASSR primers can be successfully used in the detection of short repetitive sequences in the soybean genome. However, (A-T) ASSR primers failed to give any amplification products. Wu et al. (1994) and Gupta et al. (1994), using ASSR primers, were also unsuccessful in the detection of $(AT)_n$ in several plant species. This can be attributed to self-complimentarity of (A-T) repeats (and probably (G-C) repeats as well), which may self-anneal during amplification.

Short sequence repeats are themselves a useful resource for designing ASSR primers. As shown in these experiments, new banding patterns could also be obtained by designing a 5′ ASSR primer with more than two-base change in anchor sequence. A single-base change or extension of an anchor by one nt at the 5′ end was insufficient to alter banding patterns. The introduction of degenerate anchor sequences

Table 3 Segregation of mapped ASSR markers in the F_2 population of the 'Essex' × PI 437654 soybean cross

Marker[a]	Pattern[b]	χ^2 3:1 or 1:2:1	Probability	Marker[a]	Pattern[b]	χ^2 3:1 or 1:2:1	Probability
Kl_{590}	P	0.553	0.3–0.5	$K16_{540}$	E	2.083	0.1–0.2
Kl_{280}	E	0.001	0.95–0.98	$K16_{360}$	P	0.252	0.5–0.7
$K5_{400}$	E	0.287	0.7–0.8	$K16_{240}$	P	0.101	0.7–0.8
$K5_{260}$	P	1.127	0.5–0.7	$K16_{170}$	E	0.333	0.5–0.7
$K5/MACC_{680}$	E	0.828	0.3–0.5	$K17_{580}$	P	1.000	0.3–0.5
$K5/MACC_{270}$	E	1.104	0.2–0.3	$K17_{350}$	P	0.240	0.5–0.7
$K5/MACC_{200}$	E	1.262	0.2–0.3	$K17_{260}$	E	0.666	0.3–0.5
$K6_{540}$	E	2.083	0.1–0.2	$K17/PLP2_{638/640}$	C	0.089	0.95–0.98
$K6_{480}$	E	1.313	0.5–0.7	$K17/PLP2_{400}$	E	0.355	0.5–0.7
$K6_{350}$	P	0.095	0.8–0.9	$K17/PLP2_{210/220}$	C	0.399	0.8–0.9
$K6_{200}$	P	0.095	0.8–0.9	$K17/IDHF_{800}$	P	0.411	0.5–0.7
$K7_{646/650}$	C	0.001	0.95–0.98	$K17/IDHF_{390}$	P	0.226	0.5–0.7
$K8/PLP2_{480}$	E	0.111	0.7–0.8	$K17/IDHR_{550}$	P	0.167	0.5–0.7
$K8/PLP2_{360}$	P	15.33	< 0.001	$K18_{520}$	E	0.381	0.5–0.7
$K8/PLP2_{280}$	E	0.444	0.5–0.7	$K18_{400}$	P	0.694	0.3–0.5
$K15_{360}$	P	0.252	0.5–0.7	$K18_{390}$	P	0.694	0.3–0.5
$K15_{260/270}$	C	1.064	0.5–0.7	$K18_{200}$	E	0.045	0.8–0.9
$K15_{240}$	P	0.083	0.7–0.8	$K18/PLP2_{540/550}$	C	0.491	0.7–0.8
$K15_{160}$	E	0.139	0.7–0.8	$K18/PLP2_{520/530}$	C	0.964	0.5–0.7
$K15/PLP2_{640}$	P	5.523	0.02–0.05	$K18/PLP2_{400}$	P	0.000	1.00
$K15/PLP2_{270}$	E	1.899	0.1–0.2	$K18/PLP2_{300}$	E	0.051	0.8–0.9
$K15/PLP2_{250}$	P	0.667	0.3–0.5	$K18/PLP2_{240}$	P	0.633	0.1–0.2
$K15/P34F_{400}$	P	3.227	0.05–0.1	$K19_{390}$	P	0.207	0.5–0.7
$K15/P34F_{350}$	P	0.28	0.8–0.9	$K21_{470}$	E	0.001	0.95–0.98
$K15/P34F_{250}$	P	0.470	0.3–0.5	$K21_{430}$	E	2.270	0.1–0.2
$K15/IDHF_{570}$	E	0.517	0.3–0.5	$K21_{320}$	P	0.000	1.00
$K15/IDHF_{410}$	P	0.240	0.5–0.7	$K21_{220}$	E	0.157	0.5–0.7
$K15/IDHF_{280}$	P	0.185	0.5–0.7	$K22/PLP2_{700}$	P	0.144	0.7–0.8
$K15/IDHF_{200}$	P	0.028	0.8–0.9	$K22/PLP2_{330}$	E	0.394	0.5–0.7

[a] Markers are identified by an ASSR primer/arbitrary primer. The molecular size (bp) of the fragments are in subscript
[b] P, Dominance of the PI 437654 pattern; E, dominance of the 'Essex' pattern; C, codominant segregation

enhanced the number of amplified fragments and gave new polymorphisms. Also, the addition of an anchor to the 3′ end of the SSR generated different banding patterns. This observation was in agreement with the results of Zietkiewicz et al. (1994), though the 3′ ASSR primers revealed fewer amplification products and polymorphisms than the 5′ ASSR primers.

In *Arabidopsis*, combinations of ASSR primer and decamers produced different polymorphisms compared with banding patterns produced solely by ASSR primer (Wu et al. 1994). In soybean, under the same PCR conditions as for *Arabidopsis*, the decamers (even 14 mer) generally failed to alter the banding patterns. This observation can be associated with differences in genome complexity of these two species, and the fewer decamers tested in soybean than in *Arabidopsis*. Under the amplification conditions used in our experiments (annealing temperature of 50°C), decamer primers were probably non-functional, despite the possibility of more potential binding sites for decamers than for longer primers. The successful amplification of a product with 2 primers depends on their length, melting temperatures (Tms) and the annealing temperature of the PCR. In some studies, a low annealing temperature (35°C or 36°C) was used for ASSR primer and decamer combinations (Becker and Heun 1995; Sánchez de la Hoz et al. 1996). However, a low annealing temperature may increase non-specific amplification, which means that amplified fragments may not contain ASSR. No sequence information was given by Becker and Heun (1995) or Sanchez de la Hoz et al. (1996). Also, at low annealing temperature, a change in anchor sequences did not alter banding patterns in barley (Sánchez de la Hoz et al. 1996).

The use of 18- to 20-nucleotide primers in combination with ASSR primers changed the amplification patterns. DNA sequencing of two fragments, which were produced by the ASSR K18 and PLP2 primers, revealed the presence of the primers' sequences. This confirmed that using an end-labeled ASSR primer and a 50°C annealing temperature for the first 5 cycles of PCR enhances the binding specificity of ASSR primer and promotes amplification of soybean genomic regions adjacent to the ASSR primer. Using ASSR primer in combination with specific primers for protein motifs may be useful for identifying chromosomal regions which might contain homologous sequences. This has been proven in our case, when ASSR primer K18 was used with primer PLP2 designed for the kinase domain present in several disease resistance

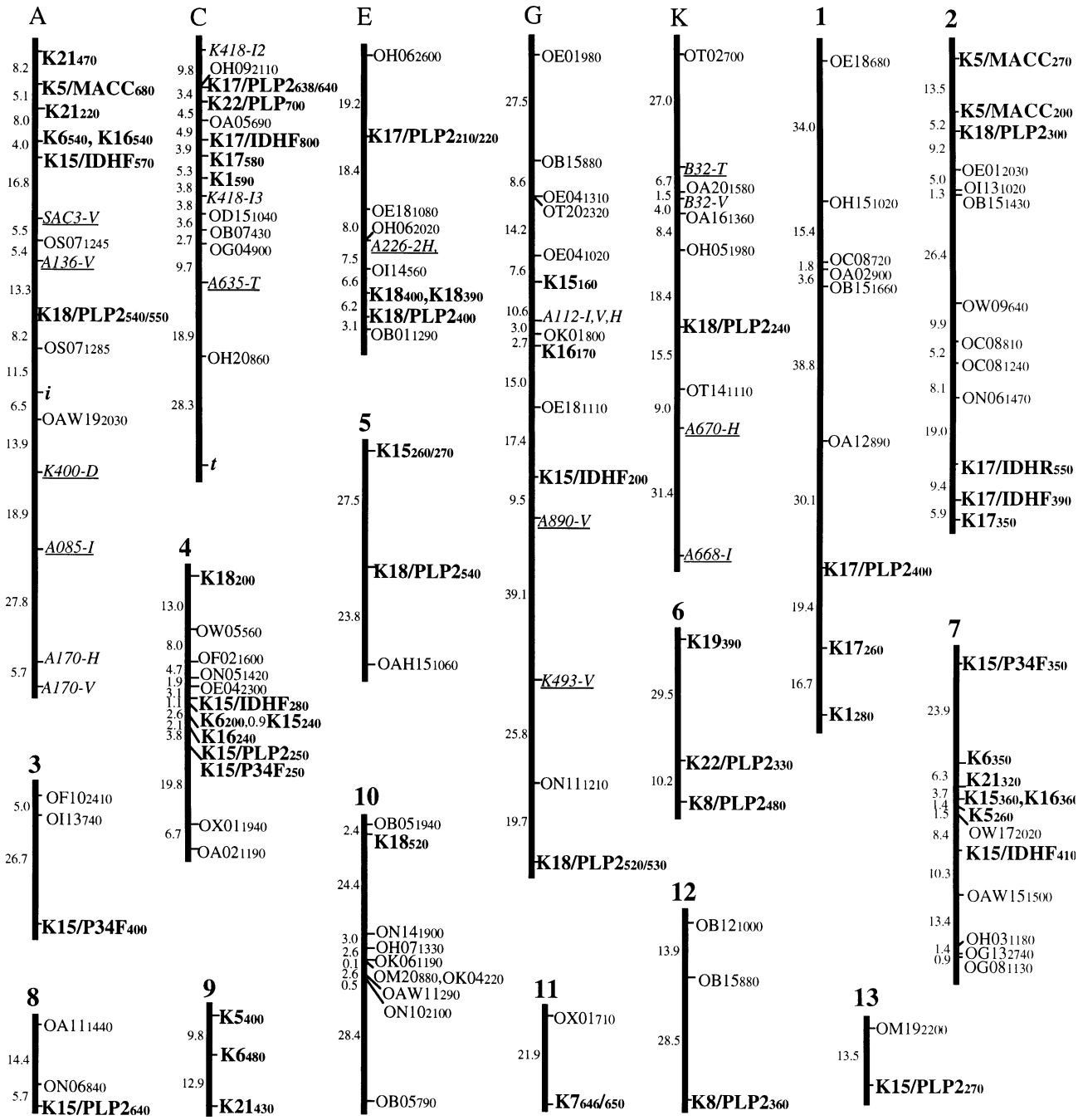

Fig. 5 Linkage assignment of ASSR markers (in bold). Designations are as follows: the ASSR primer name for markers obtained solely with ASSR primers; the ASSR primer/arbitrary primer for markers produced by the ASSR and arbitrary primers. Molecular size of the markers are in *subscript*. The RFLP anchor loci of the public soybean molecular map (Shoemaker and Specht 1995) are in *italics* and *underlined*. The *i* and *t* are two morphological loci for soybean seed-coat color and pubescence color, respectively. The remaining markers are RAPD loci. Genetic distances at LOD 3.0 are in centiMorgans (cM)

genes (Grant et al. 1995). The amino acid sequence from an amplified soybean fragment showed 50% homology to the receptor-like protein kinase precursor from *Arabidopsis* (Walker 1993).

ASSR analysis was useful in characterizing the molecular diversity of soybeans. On average, 7 polymorphisms per ASSR primer were produced. This level of polymorphism is lower than for microsatellites (5–21 alleles per microsatellite, Maughan et al. 1995) and higher than for RFLPs (Keim et al. 1990; Skorupska et al. 1993) and RAPDs (Skorupska et al. 1994; Mahalingam and Skorupska 1995). For linkage analysis

Table 4 Summary of the (G-A) and (C-A) ASSR markers in F2 population of the 'Essex' and PI 437654 cross

ASSR[a] primers	Number of polymorphisms per primer		Length of the amplified fragments (bp)	Polymorphic markers							
				Total markers Number/Percentage		Mapped markers Number/Percentage		Unlinked markers Number/Percentage		Markers linked at less than 5 cM Number/Percentage	
	Germ-plasm	Inter-cross									
(G-A)	6.0	3.9	300–900	39	45.9	20	51.3	19	48.7	2	12
(C-A)	8.0	3.8	100–700	46	54.1	38	82.0	8	17.4	15	88
Total	7.0	3.9	100–900	85	100	58	68.2	27	31.7	17	29.3

[a] (G-A): $(GA)_6$, $(AG)_6$, $(CT)_6$
[b] (C-A): $(GT)_6$, $(TG)_6$, $(CA)_6$

in the 'Essex' and PI 437654 F_2 population, we obtained 4 segregating markers per ASSR primer. The number of ASSR polymorphisms is higher than for genomic clones (0.3 polymorphism per RFLP clone) and RAPDs (1.5 polymorphisms per primer) in the same population but lower than the reported 9 polymorphisms per AFLP primer in an intraspecific cross (Keim et al. 1997). The consistency of ASSRs' classification of genotypes with that of RFLP markers (Diers et al. 1997) illustrates that ASSR-juxtaposed data are useful for describing valuable genome diversity in soybeans. Furthermore, it demonstrates the effectiveness and speed of this analysis compared with genotype classification by RFLPs. All 48 soybean genotypes tested could be distinguished with 8 ASSR primers. Sanchez de la Hoz et al. (1996) suggest that ASSRs, because of their repeated nature, are more reliable than RAPD markers for genetic relationship studies.

Morgante et al. (1994) mapped 7 SSRs to 7 soybean linkage groups; Akkaya et al. (1995) mapped 34 SSR polymorphisms, predominantly $(AT)_n$ and $(ATT)_n$, to 18 soybean linkage groups. With an exception of one cluster, those microsatellite loci appeared to be dispersed randomly in the soybean genome. We assigned 58 ASSR markers to 18 linkage groups. The map information suggests that $(GT)_n$, $(TG)_n$ and $(CA)_n$ have biased distribution, whereas $(CT)_n$, $(GA)_n$ and $(AG)_n$ are dispersed in the soybean genome. Random distribution of (G-A) may be supported by twice the frequency of unlinked (G-A) ASSR markers compared with (C-A) ASSR markers.

In rice, a total of 56 SSR loci, including $(AT)_n$, $(GA)_n$ and $(CA)_n$, were found to be distributed randomly (Akagi et al. 1996). The linkage map of barley implied a clustering of $(GA)_n$ and $(CA)_n$ (Liu et al. 1996). The clustering of $(AG)_n$ was also observed amongst mapped microsatellite markers in *Arabidopsis* (Bell and Ecker 1994). It is clear that SSRs and their genomic distribution differ across plant species. In sugar beet, in situ hybridization with $(GA)_n$ revealed dispersed signals on all chromosomes, while $(CA)_n$ were only present at centromeric regions (Schmidt and Heslop-Harrison 1996). Whether, in soybean, the (C-A) repeats are located at centromeric regions and/or indicate other regions of reduced recombination remains to be investigated.

Acknowledgements The authors thank Drs. V. W. Baird, and E. R. Shipe at Clemson University for helpful comments, and A. P. Rao-Arelli (University of Missouri, Columbia, Mo.) for providing data on response of the soybean genotypes to soybean cyst nematode. The project was supported by the NSF-EPSCoR and the South Carolina Soybean Board. This paper is Technical Contribution No. 4284 of the South Carolina Agricultural Experiment Station, Clemson University, Clemson, S.C., USA.

References

Akagi H, Yokozeki Y, Inagaki A, Fujimura TT (1996) Microsatellite DNA markers for rice chromosomes. Theor Appl Genet 93:1071–1077

Akkaya MS, Bhagwat AA, Cregan PB (1992) Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131–1139

Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA, Cregan PB (1995) Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci 35:1439–1445

Becker J, Heun M (1995) Mapping of digested and undigested random amplified microsatellite polymorphisms in barley. Genome 38:991–998

Bell CJ, Ecker JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics 19:137–144

Cregan PB (1997) Soybean SSR marker development-applications in legumes. In: Heller SR (ed) Plant and animal genome V (Int Conf Status Plant Anim Res). San Diego, Calif., p 45

Cregan PB, Bhagwat AA, Akkaya M, Rongwen J (1994) Microsatellite fingerprinting and mapping of soybean. Methods Mol Cell Biol 5:49–61

Diers BW, Skorupska HT, Rao-Arelli AP, Cianzio SR (1997) The genetic relationship among plant introductions with resistance to soybean cyst nematodes. Crop Sci 37:1966–1972

Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes RW, Dang JL (1995) Structure of the *Arabidopsis* RPM1 gene enabling dual specificity disease resistance. Science 269:843–846

Gupta M, Chyi YS, Romero-Severson J, Owen JL (1994) Amplification of DNA markers from revolutionarily diverse genomes using single primers of simple-sequence repeats. Theor Appl Genet 89:998–1006

Keim P, Shoemaker RC, Olson TC (1988) A rapid protocol for isolating soybean DNA. Soybean Genet Newsl Vol 18. Department of Agronomy, Iowa Stata University and the United States Department of Agronomy, Agricultural Research Service, Ames, Iowa

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

Keim P, Schupp JM, Travis SE, Clayton K, Webb DM (1997) A high density soybean genetic map based upon AFLP markers. Crop Sci 37:535–543

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Lagercrantz U, Ellegren H, Anderson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115

Lander ES, Green P, Abrahamson J, Baklow 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

Liu ZW, Biyashev RM, Saghai Maroof MA (1996) Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theor App Genet 93:869–876

Mahalingam R, Skorupska HT (1995) DNA markers for resistance to *Heterodera glycines*. I. Race 3 in soybean cultivar Peking. Breed Sci 45:435–443

Maughan PJ, Saghai Maroof MA, Buss GR (1995) Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. Genome 38:715–723

Morgante M, Rafalski AJ, Biddle P, Tingey SV, Olivieri AM (1994) Genetic mapping and variability of seven simple sequence repeat loci in soybean. Genome 37:763–769

Powell W, Morgante M, Doyle JJ, McNicol JW, Tingey SV, Rafalski AJ (1996) Genepool variation in genus *Glycine* subgenus *soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics 144:793–803

Rohlf FJ (1994) NTSYS-PC, Numerical taxonomy and multivariate analysis system. Version 1.80. Exeter Software Setauket, N.Y.

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90:43–48

Sánchez de la Hoz MP, Dávila JA, Loarce Y, Ferrer E (1996) Simple sequence repeat primers used in polymerase chain reaction amplifications to study genetic diversity in barley. Genome 39:112–117

Schmidt T, Heslop-Harrison JS (1996) The physical and genomic organization of microsatellites in sugar beet. Proc Natl Acad Sci USA 93:8761–8765

Shoemaker RC, Specht JE (1995) Integration of the soybean molecular and classical genetic linkage groups. Crop Sci 35:436–446

Skorupska HT, Shoemaker RC, Warner A, Shipe ER, Bridges WC (1993) Restriction fragment length polymorphism in soybean germplasm of the Southern U.S.A. Crop Sci 33:1169–1176

Skorupska HT, Choi IS, Rao-Arelli AP, Bridges WC (1994) Resistance to soybean cyst nematode and molecular polymorphism in various sources of Peking soybean. Euphytica 75:63–70

Sneath PHA, Sokal RR (1973) Numerical taxonomy. The principles and practice of numerical classification. WH Freeman, San Francisco

Tautz D (1989) Hypervariability of simple sequences as a source for polymorphic DNA marker. Nucleic Acids Res 17:6463–6471

Udvardi MK, McDermott TR, Kahn ML (1993) Isolation and characterization of a cDNA encoding NADP+ – specific isocitrate dehydrogenase from soybean (*Glycine max*). Plant Mol Bio 21:739–752

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijter A, Pot J, Peleman J, Kulper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Walker JC (1993) Receptor-like protein kinase genes of *Arabidopsis thaliana*. Plant J 3:451–456

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n$, $(dG\text{-}dT)_n$ polymorphism. Genomics 7:524–530

Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using polymerase chain reaction. Am J Hum Genet 44:388–397

Wu KS, Jones R, Danneberger L, Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. Nucleic Acids Res 22:3257–3258

Yu YG, Saghai Maroof MA, Buss GR (1996) Divergence and allelomorphic relationship of a soybean virus resistance gene based on tightly linked DNA microsatellite and RFLP markers. Theor Appl Genet 92:64–69

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183