METHODS IN MOLECULAR AND CELLULAR BIOLOGY 5:49–61 (1994)

Microsatellite Fingerprinting and Mapping of Soybean

PERRY B. CREGAN,[1,*] ARVIND A. BHAGWAT,[1,2] MAHINUR S. AKKAYA,[3] and JIANG RONGWEN[1]
[1]USDA-ARS, Soybean and Alfalfa Research Laboratory, Range 1, Beltsville, Maryland 20705; [2]Agronomy Department, University of Maryland, College Park, Maryland 20742; [3]Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, Maryland 21702, USA

Morphology, pigmentation, disease resistance, and other traits have long been used to identify plant genotypes and to develop genetic linkage maps. These maps have been limited by the number of markers available. DNA-based markers have allowed an almost limitless supply of genetic loci for identification and mapping. In particular, restriction fragment length polymorphisms (RFLP) have been useful in numerous plant species. In soybean, however, the level of genetic polymorphism associated with RFLP is limited. The discovery of microsatellite or simple sequence repeat (SSR) loci in humans and their highly polymorphic nature was recently followed by the elucidation of similar loci in soybean and other plant species. A set of simple protocols is described to allow the development of microsatellite markers in a target species. The steps include the construction of a genomic DNA library, screening to identify clones containing SSR, sequencing of positive clones, and selection of PCR primers flanking the microsatellite. A standard procedure to assess each set of PCR primers, followed by a general approach to the use of SSR markers for fingerprinting and genetic map development, is described. The detection of microsatellite DNA markers using agarose or DNA-sequencing gels is commonly used. However, a number of increasingly sophisticated techniques for the detection of SSR-containing PCR products are now available or will be soon. These include the multiplexing of 10–15 markers per lane on sequencing gels, the detection of fluorescent labeled products, and multi-channel capillary electrophoresis. These and other developing technologies will allow the rapid and efficient molecular genotyping of large numbers of plant genotypes at multiple loci.

INTRODUCTION

DNA markers based upon restriction fragment length polymorphism (RFLP) have been the major type of marker used in the soybean (*Glycine max* (L.) Merr.) for molecular map development and DNA fingerprinting. However, the level of polymorphism detected by RFLP in soybean is relatively low. Because insufficient molecular genetic variation was thought to exist in crosses of cultivated soybean, the USDA/Iowa State University soybean RFLP map (Shoemaker and Olson, 1993) was developed using a mapping population derived from crossing cultivated soybean with wild soybean (*Glycine soja* Sieb.

Received May 10, accepted May 20.

*To whom reprint requests/correspondence should be addressed.

and Zucc.). Other studies have documented the inability of RFLP markers to create unique DNA profiles of soybean genotypes in an efficient and economical fashion (Keim et al., 1989, 1992). Molecular markers that detected much higher levels of genetic variability than is detectable with RFLP would be very useful for DNA fingerprinting and genetic mapping in a species such as soybean.

A DNA marker based upon variation in the number of tandemly repeated 2- to 5-nucleotide DNA core sequences was suggested simultaneously by three human genetics research groups in 1989 (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989). In humans, microsatellite or simple sequence repeat (SSR) markers are abundant and highly polymorphic. Based upon a search of GenBank, Cregan (1992) indicated the presence of microsatellite DNA in a number of plant species and suggested the potential for the use of SSR markers in plants. Shortly thereafter, Akkaya et al, (1992) reported microsatellite loci in soybean with as many as eight alleles. Since that time, SSR markers have been shown to be present and highly polymorphic in a number of plant species (Lagercrantz et al., 1993; Morgante and Olivieri, 1993; Senior and Heun, 1993; Wu and Tanksley, 1993; Zhao and Kochert, 1993). In recent work describing newly identified SSR loci in soybean, Jiang et al. (1994) reported from 11 to 26 alleles per locus with gene diversity [identical with polymorphism information content (Botstein et al., 1980)] values ranging from 0.71 to 0.95. These values contrasted to an average gene diversity of 0.30 for 132 RFLP probes reported by Keim et al. (1992). Because of their greater informativeness and ease of use, microsatellites have now become the marker of choice in human genetic linkage analysis (NIH/CEPH Collaborative Mapping Group, 1992; Weissenbach et al., 1992). Similarly, in soybean and other species SSR markers appear to be an excellent alternative to RFLP.

METHODOLOGY

Our research group is currently constructing a microsatellite map of soybean that will contain between 250 and 300 loci. We are identifying the most informative of these loci with the intention of selecting a group of SSR markers that will provide the basis of a soybean DNA fingerprint database. In identifying and characterizing microsatellite markers we have attempted to use the least expensive and simplest techniques available. It is our belief that this technology will see its widest adoption by the plant breeding and genetics community if the difficulty and expense of developing and using SSR markers are minimized.

The development of microsatellite markers consists of a series of steps including the construction of a genomic DNA library, screening to identify clones containing SSR, sequencing of positive clones, and selection of PCR primers flanking the microsatellite.

Genomic DNA Library Construction

To isolate a DNA clone containing a short sequence that occurs at a low frequency, appropriate strategies might include the creation of a library from DNA fragments that had been enriched for the sequence, or the screening of a library carrying large inserts followed by subcloning. However, a number of reports (Condit and Hubbell, 1991; Cregan, 1992; Lagercrantz et al., 1993; Morgante and Olivieri, 1993; Zhao and Kochert, 1993) estimate that certain microsatellites such as $(AT)_n$/$(TA)_n$, $(CT)_n$/$(GA)_n$, and $(CA)_n$/$(GT)_n$ occur as frequently as every 150 to 200 kbp in soybean and/or other plant genomes. Because of this high frequency, it is reasonable and technically less demanding to create libraries of unselected genomic DNA fragments that can be sequenced without subcloning. In this way one is not required to use protocols to

enrich for the presence of fragments carrying specific microsatellites (Karagyozov et al., 1993) or procedures designed to directly amplify or sequence SSR-flanking regions from large clones (Baron et al., 1992; Browne and Litt, 1992; Edwards et al., 1991; Herman, 1993).

High-quality genomic DNA of the soybean cultivar Williams was either sheared using sonication or cut with a restriction endonuclease and size selected in a 0.7% agarose gel. The sonicated DNA gave a random sampling of genomic DNA rather than a subset of fragments flanked by specific endonuclease restriction sites. DNA fragments in the 450 to 550 bp range were isolated from the gel using Geneclean II (Bio101 Inc., La Jolla, CA) or a similar product. In the case of the sheared DNA fragments, S1 nuclease (Roberts et al., 1979) or mungbean exonuclease treatment (Kroeker et al., 1976) was used to assure blunt ends. Fragments were ligated into pBluescript (Stratagene, La Jolla, CA), pGEM (Promega Corp., Madison, WI), pUC18 (Sambrook et al., 1989), or another plasmid vector that was used to transform freshly prepared competent cells of *Escherichia coli* such as DH5α or XL1-Blue. Transformed cells were plated on selective medium containing the appropriate antibiotics and X-gal for blue/white color selection to assure transformed cells carrying plasmids with genomic inserts. Between 400 and 500 transformants on an 88-mm petri plate is a reasonable number, and allows the screening of 200–250 kbp of genomic DNA.

Screening Clones for the Presence of SSR

Two colony lifts were made from each petri plate by placing an 82-mm diameter supported nitrocellulose membrane (0.45 μm pore size) (Schleicher & Schuell, Keene, NH) on the surface of the petri plate containing a 12-h growth of transformed cells. After lifting, each membrane was placed (colony side up) on LB agar containing appropriate antibiotics in large (23 cm × 33 cm) baking pans, and a second membrane was placed on top of each. Colonies were allowed to grow for 8 h at 37°C. The membranes, still face to face, were placed on LB agar containing chloramphenicol (175 μg/mL) for 12–15 h at 37°C to further increase plasmid copy number after which they were placed on Whatmann 3MM paper saturated with 10% SDS at room temperature for 5 min. Membranes were transferred to Whatmann 3MM paper saturated with denaturing solution (0.5 N NaOH, 1.5 M NaCl) for 10 min followed by transfer to Whatmann 3MM paper saturated with neutralization solution (1.5 M NaCl, 0.5 M Tris-HCl, pH 8.0). Subsequently, membranes were placed on Whatmann 3MM paper saturated with 2× SSC for 10 min; the membranes were then separated and allowed to dry for 1/2 h at room temperature followed by baking in a vacuum oven for 1–2 h at 80°C. This protocol is very similar to that described by Sambrook et al. (1989).

Each of the four membranes from a petri plate was screened by colony hybridization using a different radiolabeled oligonucleotide probe. For example, membranes were initially screened for the presence of $(CT)_n/(GA)_n$, $(CA)_n/(GT)_n$, $(AT)_n/(TA)_n$, and $(ATT)_n/(TAA)_n$ SSR. To increase the specific activity and life of the probe it was preferable to label the internal bases of the probe rather than just the 5′

end. Thus, an $GG(AG)_{15}$ oligonucleotide probe was made using a reaction mix containing equal molar amounts of a $(CT)_{15}$ CC oligonucleotide template and a $GG(AG)_3$ primer, KGB buffer (Sambrook et al., 1989), Klenow polymerase, dGTP, α-^{32}PdATP. In the case of the $GGG(AAT)_{10}$ probe, the reactants were similar except a $(ATT)_{10}CCC$ template and a $GGG(AAT)_2$ primer were used. After labeling, the reactants were filtered through a BioSpin 6 chromatography column (Bio-Rad, Richmond, CA) for removal of unincorporated nucleotides. Scintillation counts before and after removal of unincorporated nucleotides usually indicated 75% or greater incorporation of α-^{32}PdATP.

Simultaneous screening with a number of oligo probes. Another useful method to screen for tri- or tetranucleotide microsatellites that occur at a relatively low frequency is to screen for more than one core motif at a time. The probes used in multiple screening had similar or identical Tm values. The oligonucleotide probes $GG(CAA)_{10}$, $GG(CAT)_{10}$, $GG(CTA)_{10}$, and $GG(CTT)_{10}$, all with a core sequence of 33% GC were labeled using α-^{32}PdCTP, the appropriate template, [e.g., $(TTG)_{10}CC$], 8-mer primer [e.g., $GG(CAA)_2$], buffer, dATP and/or dTTP, and Klenow polymerase as described above. The four probes were then pooled and used simultaneously. Likewise, $GG(ACC)_{10}$, $GG(ACG)_{10}$, $GG(AGC)_{10}$, and $GG(AGG)_{10}$ probes with core sequences of 67% GC were labeled using α-^{32}PdATP, pooled, and used simultaneously in hybridization.

Pre-washes and pre-hybridization. Before pre-hybridization, the membranes were washed for 5 min in 6 × SSC at room temperature followed by two 1-h pre-washes [0.05 M Tris-HCl (pH 8.0), 1.0 M NaCl, 0.001 M EDTA, 0.1% SDS] at 42°C and finally a 1-h wash (0.1 SSC, 0.5% SDS) at 65°C. The membranes were then transferred to pre-hybridization solution (6 × SSPE, 5 × Denhardt's solution, and 1% SDS) for 2 h at 37°C. Membranes were removed from the pre-hybridization solution and placed in hybridization tubes (30 mm diameter × 30 cm length) to which 15 ml of hybridization solution (6 × SSPE and 1% SDS) was added. Tubes were cooled at 4°C and the denatured probe was added. The addition of the denatured oligo probe to the hybridization tubes cooled at 4°C seemed to be helpful in obtaining better hybridization in the case of the $(AT)_n$ probe that was particularly prone to renature. Hybridization was done overnight at 38°C, then the hybridization solution was removed and reused in subsequent screenings.

Membrane washing. The goal of this screening protocol was to obtain clones containing microsatellite sequences with more than 10 dinucleotide or 7 tri-nucleotide repeats. According to Weber (1990), these are the numbers of repeat units required to obtain a reasonable expectation of simple sequence length polymorphism (SSLP) (Dietrich et al., 1992) at a microsatellite locus. Thus, careful determination of membrane washing temperature to accommodate the Tm of the probe can assure that most clones selected for sequencing will contain SSR meeting these minimum length criteria. Following hybridization, membranes were washed 2 times for 1/2 h in 1× SSC with 0.1% SDS. The $GG(AT)_{15}$ probe was washed at 38°C, the $GGG(AAT)_{10}$ probe at 45°C and the $GG(AG)_{15}$ and $GG(CA)_{15}$ probes at 55°C. Even higher temperatures for the $GG(AG)_{15}$ and $GG(CA)_{15}$ may be desirable. After washing, the membranes were blotted dry, taped to Whatmann paper, and exposed to X-ray film for an appropriate period of time to give small but

clearly defined hybridization signals that were used to identify the colony or small group of colonies on the petri plate from which the signal was obtained.

Colony selection and rescreening. After aligning the X-ray film to identify the colony or colonies from which the hybridization signal was obtained, a standard number of colony picks (6 or 12) were made from individual colonies (when possible) on the petri plate. More than one colony was picked because it is often difficult to identify the exact colony that produced the hybridization signal. Each pick was made into a separate microtiter plate well containing 100 µl of LB medium with appropriate antibiotics to eliminate plasmidless *E. coli* and contaminants. After overnight growth at 37°C, the contents of each microtiter plate were stamped onto nitrocellulose membrane using a 96-prong stamper, and grown on LB agar in baking trays as described above. The growth on LB, lysis, denaturation, neutralization, hybridization, washing, and exposure to X-ray film were as described above. The only difference in this screening was in the denaturation of the probe. After boiling, the hybridization solution containing the probe was quickly frozen in –70°C ethanol and then allowed to melt at 4°C over a number of hours.

From each set of six or 12 microtiter plate wells (each set traceable to one hybridization signal from the original screening of the colony lifts), one colony with the strongest hybridization signal was plated on LB selective medium containing the appropriate antibiotics and X-gal for blue/white color selection. Six colonies were picked from each plate into separate microtiter plate wells containing 100 µl of LB medium with appropriate antibiotic(s), as described above. After overnight growth the colonies were stamped onto nitrocellulose membrane and again hybridized to the oligo probe as described above. About 33 µl of 50% glycerol was added to each well of these microtiter plates and the plates were then stored at -70°C for long-term preservation.

This procedure should yield clones with 500-bp genomic inserts with a high probability of containing an SSR of acceptable length to produce SSLP. Such plasmid clones can be easily sequenced using most of the commonly available sequencing protocols.

Sequencing of SSR-Containing Clones

Numerous protocols are available to isolate sequencing-quality plasmid DNA. One commonly used protocol that can cause difficulties is the PCR amplification of the insert. While this strategy is appealing, it is one that may not always be appropriate with SSR-containing sequences. The phenomenon of "polymerase slippage" or "stuttering" (Edwards et al., 1991) during the PCR amplification of a microsatellite with *Taq* polymerase may create one dominant PCR product with one or more products containing fewer or greater repeat units than the dominant product. As a result, the sequence data on the 3′ side of the SSR may be unreadable. Polymerase slippage is less likely to be a problem with tri- and tetra-nucleotide repeats. Furthermore, colonies can be sequenced from both ends to obtain clean sequences in both SSR flanking regions.

PCR Primer Selection

There are many computer programs available to select PCR primers. One that worked well and is available in both PC and MacIntosh versions is called PRIMER; it can be obtained at no cost through the Lander Lab (Whitehead Institute, Cambridge, MA 02142). Using PRIMER, we selected primers to soybean microsatellites to have a Tm of 50.5 ± 0.5°C with an expected PCR fragment size between 100 and 200 bp. The homogeneous Tm values allowed PCR amplifications of all SSR loci to be carried out at a standard annealing temperature. We had originally intended to use a higher Tm, but owing to the low GC content of many soybean clones 50°C was a necessity if primer length was to be approximately 20 to 23 bases.

Evaluation of PCR Primers

The initial evaluation of each primer set was conducted using two DNA templates. One was the plasmid containing the insert from which the primer sequences were derived, and the second was genomic DNA of Williams soybean, the source of the plasmid library. Reaction mixes contained either 0.5 ng of plasmid template or 30 ng of soybean genomic DNA, 1 mM Mg^{2+}, 0.15 μM of 3′ and 5′ end primers, 100 μM of each nucleotide, 0.1 μL of 3,000 Ci/mmol α-^{32}PdATP, 1× PCR buffer containing 50 mM KCl, 10 mM Tris-HCl pH 8.3, 1 unit *Taq* DNA polymerase (Promega Corp., Madison, WI) in a total volume of 10 μL. (^{32}P end-labeling of one or both of the PCR primers is also commonly done.) Cycling consisted of 25 s denaturation at 94°C, 25 s annealing at 47°C, and 25 s extension at 68°C for 32 cycles on a MJ Research model PTC-100 thermocycler (MJ Research, Inc., Watertown, MA). PCR products (3 μL/lane) were separated on a standard DNA sequencing gel containing 6% polyacrylamide, 8 M urea, and 1× TBE, at 60 Watts constant power for 2–3 h. To estimate the exact length of the denatured PCR product and to allow a comparison with predicted length, sequencing reactions of M13 ssDNA were used as molecular weight standards. The M13 sequencing reaction containing dideoxy-A was run in one lane and those containing dideoxy G, C, and T were combined and run in a second lane. These were run adjacent to the PCR products. Following drying, gels were exposed to X-ray film.

Comparison of the PCR products with the sequencing ladder allowed a determination of actual versus predicted size. Soybean is an inbreeding species, thus all loci are expected to be homozygous. PCR primers specific to a single locus would therefore be anticipated to produce one amplification product. Based on the initial primer evaluation, we discarded those primer sets that did not give single amplification products of the predicted length from both the plasmid and Williams soybean templates.

An estimate of the size of the allele present at the SATT9 microsatellite locus in Williams soybean can be obtained from Figure 1 (Lane 1). The predicted length of the PCR product from Williams was 152 bp. Using the M13 sequencing ladder, the lower band from Williams soybean was estimated to be 153 bases in length. This is one greater than the predicted length and is likely the result of the addition of a non-

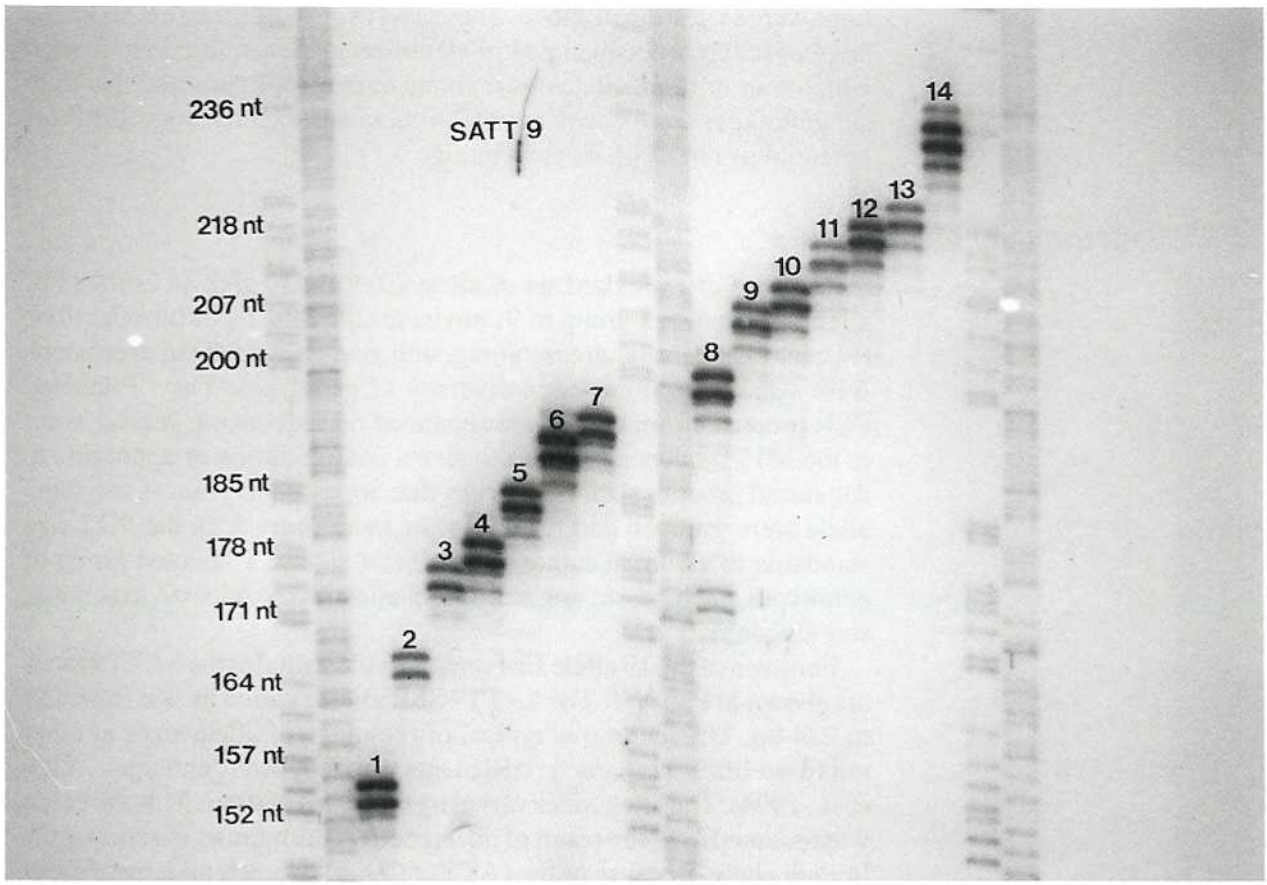

Fig. 1. The SATT9 microsatellite locus showing 14 of the 19 alleles detected among 96 soybean genotypes. The denatured DNA fragments in **lanes 1–14** are derived from PCR products (SSR alleles) from the soybean genotypes indicated with the estimated allele size (in base pairs) in parentheses: 1: Williams (153); 2: PI 101404B (165); 3: Lee (174); 4: PI 180501 (177); 5: PI 86063 (183); 6: PI 326582A (189), 7: PI 200492 (192); 8: PI 240664 (198 and 174); 9: Aoda (207); 10: S-100 (210); 11: Wuchang (216); 12: Centennial (219); 13: Haberlandt (222); and 14: Tracy (234).

templated nucleotide on the 3′ end of the PCR product, as described by Clark (1988).

The denatured PCR product amplified from Williams and genomic DNA of other soybean genotypes consisted of two major bands (Fig. 1). Because the microsatellite present at the SATT9 locus is $(ATT)_n/(TAA)_n$, one strand of the PCR product was expected to be purine rich and of higher molecular weight in relation to the other. In the case of Williams soybean, the two major bands differed by the equivalent of about 2 bases in length (as estimated from the M13 ladder), i.e., 153 versus 155 bases, or about 1.3%. This difference was very close to the molecular weight difference that would be anticipated between opposite strands of the PCR product amplified from Williams soybean at the SATT9 locus.

Simple Sequence Length Polymorphism (SSLP)

Those primer sets found to be acceptable in the initial evaluation were used in assays of six soybean cultivars with diverse origins. The PCR reaction mixtures, cycling conditions, and sequencing gel separa-

tions were as described above. These assays gave information about the level of polymorphism at newly identified microsatellite loci. Loci at which four or more alleles were found among Williams and the other six genotypes were considered to possess sufficiently high SSLP for potential use in DNA fingerprinting.

DNA Fingerprinting of Soybean

To develop a standard set of allele sizes and to give an estimate of allelic frequency, a group of 91 soybean cultivars, Plant Introductions (PI), and Chinese landraces along with five wild soybean accessions were selected to represent a diversity of genotypes. The ^{32}P-labeled PCR products from these were separated on sequencing gels adjacent to the M13 sequencing ladder to give a determination of approximate denatured product size. Genotypes that appeared to possess the same allele were grouped and re-run one or more times with the M13 size standards to assure accurate sizing. After sizing, a selected group of genotypes, each possessing a different allele, were selected to serve as size standards.

Fourteen of the 19 allele size standards identified at the SATT9 locus are shown in Figure 1. The SATT9 SSR alleles varied in size from 153 to 234 bp. The range was typical of the range of allele sizes at other microsatellite loci characterized in this set of soybean genotypes (Jiang et al., 1994). The allele sizes varied by multiples of three. This variation was assumed to be the result of differences in the number of repeat units in each allele. Because of the $(ATT)_n/(TAA)_n$ trinucleotide core motif of the SATT9 locus, it was anticipated that alleles would differ by multiples of three base pairs.

To determine the allele present in an unknown genotype at a particular locus, an approximate determination of allele size is made by comparison with the M13 sequencing ladder followed by a comparison with the appropriate allele size standard.

Tri-nucleotide vs. Di-nucleotide Microsatellites for DNA Fingerprinting

It can be noted that there are a number of "shadow" bands present in the SSR alleles at the SAT43 and SAT1 loci (Fig. 2, 3). These loci have an $(AT)_n/(TA)_n$ core motif. At these two loci the shadow bands were the result of PCR products that differ in length from the most prominent band by multiples of two base pairs. Such bands were also present in PCR products from SSR loci with trinucleotide core motifs such as the SATT9 locus (Fig. 1), but to a lesser extent. As noted earlier, the shadow bands were particularly prevalent with dinucleotide SSR and have been shown to be less prevalent with tri- and tetra-nucleotide SSR (Edwards et al., 1991). It required many comparisons of genotypes, particularly at the SAT1 locus, to be certain of the exact allele length of each genotype. In contrast, it was much less difficult to unambiguously determine allele lengths at loci with tri-nucleotide core motifs such as SATT9 and other similar loci (Cregan et al., 1994; Jiang et al., 1994). Thus, for the purpose of fingerprinting, it is our conclusion that SSR loci with tri-nucleotide core motifs were clearly preferable to those with di-nucleotide cores because ambiguous allele sizing was significantly less difficult.

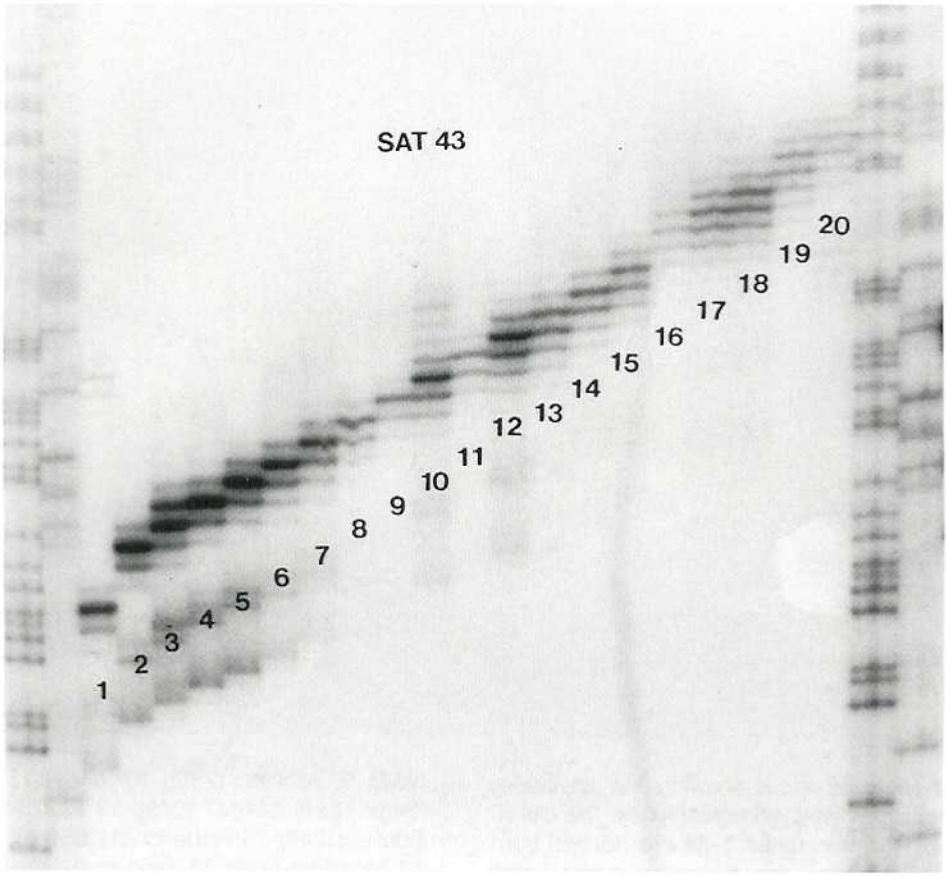

Fig. 2. The SAT43 SSR locus showing 20 of the 21 alleles detected among 96 soybean genotypes. The denatured DNA fragments in **lanes 1–20** are derived from PCR products (SSR alleles) from the soybean genotypes indicated with the estimated allele size (in base pairs) in parenthesis: 1: PI 101404B (157); 2: PI 80471-1 (163); 3: Taixin Black (165); 4: Forrest (167); 5: Peking (169); 6: Calland (171); 7: Illini (173); 8: Laredo (175); 9: Haberlandt (177); 10: Flambeau (179); 11: Kanro (181); 12: PI 86063 (183); 13: Goldsoy (185); 14: Capital (187); 15: PI 342619A (189); 16: Bragg (193); 17: Lee (195); 18: PI 273483B (197); 19: Biloxi (201); and 20: PI 240664 (203). From Jiang et al. (1994), by permission of the publisher.

Genetic Mapping of Microsatellite Loci

The ability to determine the exact size of SSR alleles, so important in DNA fingerprinting is not as critical in most aspects of genetic mapping. Usually, the alleles present in the mapping population at a given locus are known. Thus, it is simply a matter of determining which allele or alleles are present in an individual without the necessity of knowing exact allele size. We have mapped SSR loci in the population derived from a single cross between two soybean cultivars (Specht et al., 1994). For those loci at which the segregating alleles differed by less than 6–8 base pairs in length, ^{32}P-labeled PCR products were visualized as described above using a sequencing gel followed by autoradiography. Homozygotes that carry alleles varying by 2 or 3 base pairs in length as well as heterozygotes that carry both such alleles were distinguishable in this manner (Fig. 4). If segregating alleles differed by more than 6–8 base pairs in length, a 1.25% agarose + 1.25% Synergel (Diversified Biotech, West Roxbury, MA) horizontal gel with ethidium bromide incorporated in the gel was used. MetaPhor agarose (2.5%)

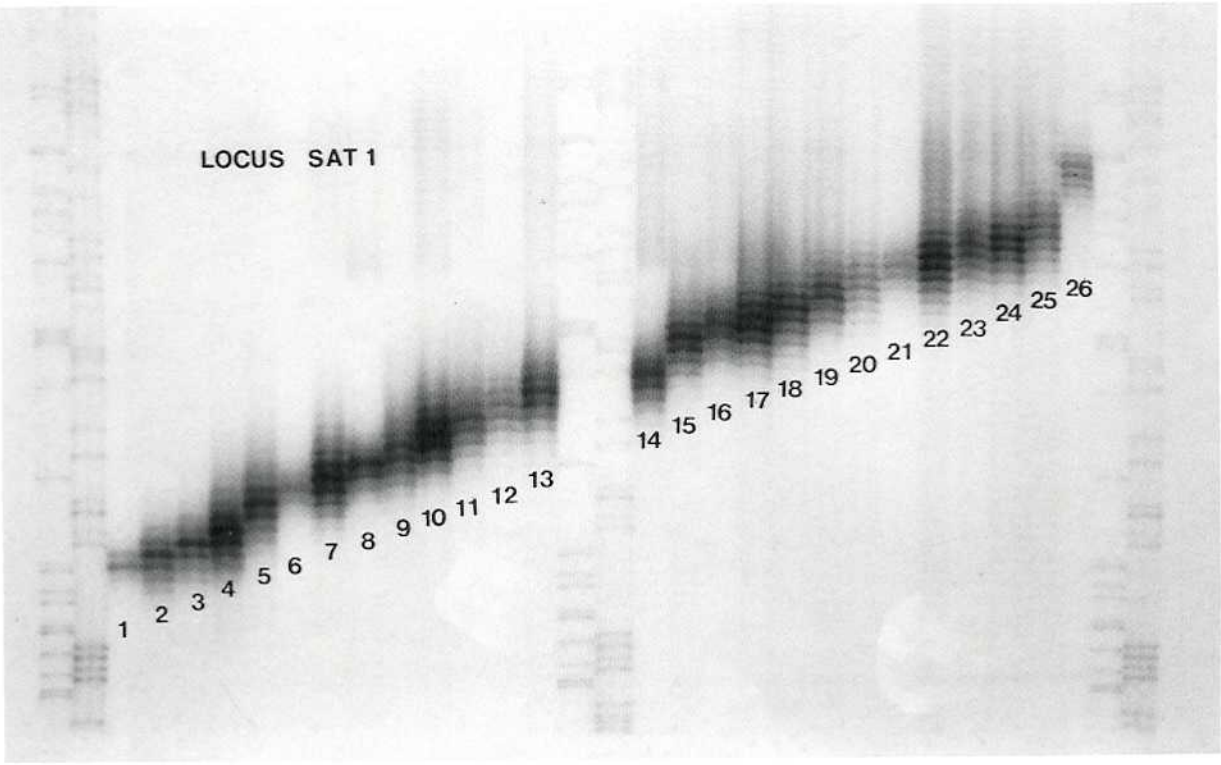

Fig. 3. The SAT1 SSR locus showing the 26 alleles detected among 96 soybean genotypes. The denatured DNA fragments in **lanes 1–26** are derived from PCR products (SSR alleles) from the soybean genotypes indicated with the estimated allele size (in base pairs) in parentheses: 1: Huangmao T.J. (221); 2: Ralsoy (223); 3: PI 342619A (225); 4: Xiangdu #3 (227); 5: Fiskeby V (231); 6: PI 65338 (233); 7: Palmetto (235); 8: PI 180501 (237); 9: Jackson (239); 10: 54610 (241); 11: Kanro (243); 12: PI 349647 (245); 13 PI 273483B (247); 14: Dunfield (249); 15: Haberlandt (255); 16: PI 85356 (257); 17: Mandarin (259); 18: Mandarin (Ottowa) (261); 19: S-100 (265); 20: Illini (267); 21: Centennial (269); 22: Mammoth Yellow (271); 23: Calland (273); 24: Suxie (275); 25: PI 84987A (277); and 26: Ogden (289). From Jiang et al. (1994), by permission of the publisher.

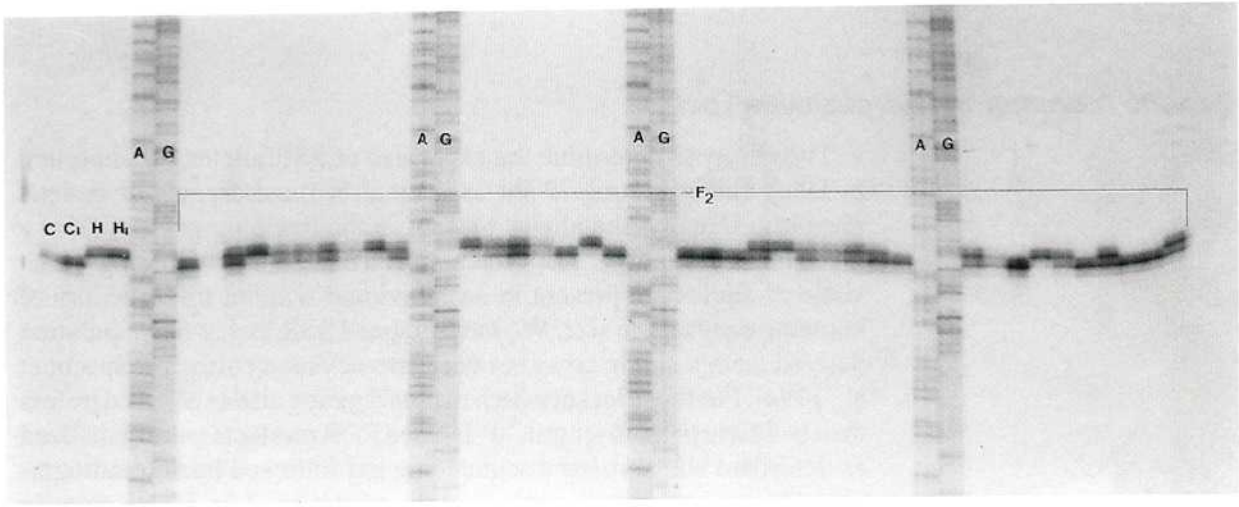

Fig. 4. The determination of the SSR allele present at the SoyPRP1 SSRlocus (Akkaya et al., 1992) in genotypes Clark **(C)**, an isoline of Clark **(C_1)**, Harosoy **(H)**, an isoline of Harosoy **(H_1)**, and a series of F_2 plant progeny from the cross of $C_1 \times H_1$. The segregating alleles differ by three base pairs in length. Homozygous parental types and heterozygotes can usually be easily distinguished. **Lanes A** and **G** were obtained from sequencing reactions using M13 ssDNA.

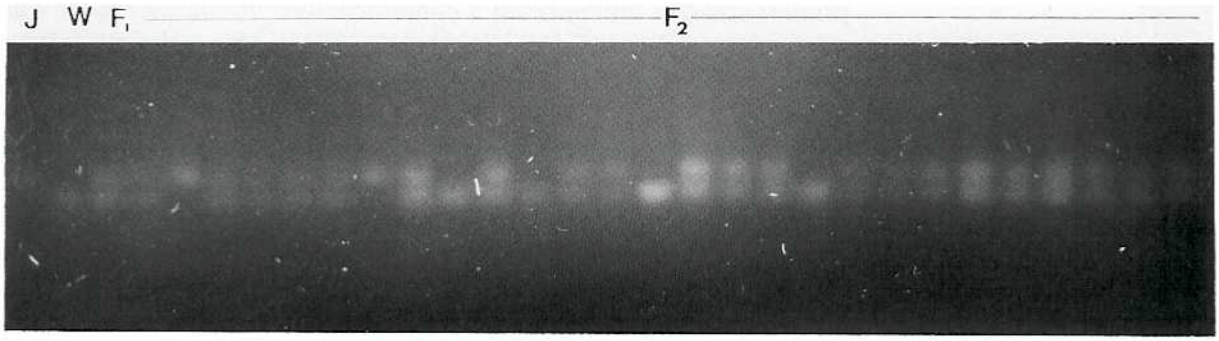

Fig. 5. The analysis of segregation at the SATT5 soybean SSR locus of a series of F2 progeny from a cross of Jackson × Williams soybean. The three lanes at the left contain Jackson **(J)**, Williams **(W)**, and their F_1 hybrid. The SSR allele carried by Williams is 142 bp in length and that of Jackson 151 bp. A total of 25μL from a 30 μL PCR amplification were loaded on a 1.25% agarose + 1.25% Synergel (Diversified Biotech, West Roxbury, MA) horizontal gel with ethidium bromide incorporated in the gel. The gel was run for 15 h at 50 volts.

(FMC Bioproducts, Rockland, ME) functions equally well to separate alleles that differ by only 6–8 base pairs. Under these conditions codominant segregation was readily determined (Fig. 5). When ethidium bromide visualization was used, the PCR volume was increased to 30 μl to produce a greater amount of product. The use of agarose for the separation of microsatellite alleles can be adopted easily by researchers who currently are equipped to use agarose in RFLP or RAPD (random amplified polymorphic DNA) (Williams et al., 1990) mapping.

Multiplex PCR. Genetic mapping can be expedited by using multiplex PCR (see Parrish and Nelson, this volume). Primer sets should be selected for compatibility, i.e., no cross homologies. Two or more primer sets can then be used in the same PCR reaction and separated simultaneously in a single lane of a gel. The allele sizes from the co-amplified loci must be dissimilar.

Other Technologies for the Detection of SSR Alleles

Numerous methods are available for labeling and detection of microsatellite alleles. A brief listing of some of these is provided below.

Ethidium bromide or silver staining of polyacrylamide gels. PCR products containing microsatellites can readily be separated on 8–10% non-denaturing polyacrylamide gels. These can be stained with ethidium bromide or silver stained as suggested by Klinkicht and Tautz (1992) and others (Bassam et al., 1991). We have often seen additional bands of higher than expected molecular weight when using non-denaturing polyacrylamide gels. These additional bands can be confusing when working with heterozygous and/or autopolyploid genotypes.

Multiplex PCR, sequencing gel separation, gel blotting, and non-radioactive detection. Hazan et al., (1992) developed a protocol that was refined and described in detail by Vignal et al. (1993) in which PCR products from as many as 16 SSR loci from the same individual were combined and separated in a single lane on a sequencing gel. The gel was blotted onto nylon membrane and sequentially probed with one of the primers of each pair used in the PCR amplifications. While the

primers could be radiolabeled, a nonradioactive probing technique was suggested and described by Vignal et al. (1993). This technique is now the basis of the large-scale human mapping project at Genethon in Evry, France (Weissenbach et al., 1992).

Fluorescent tagging and automated detection of microsatellite-containing PCR fragments. Chemistry to produce, and equipment to detect, single-stranded fluorescently labeled DNA fragments have been available (Applied Biosystems International, Forest City, CA) for DNA sequencing for some time. Products were labeled with four different fluorescent dyes corresponding to the A, G, C, and T dideoxy termination reaction used in sequencing and detected with an argon ion laser beam that scans across the lanes of a sequencing gel. This technology has been adapted for use in the size determination of SSR alleles (Ziegle et al., 1992). SSR-containing PCR products were produced in reactions in which one of the primers carried a fluorescent label. Three different fluorescent labels were available. Denatured products from three PCR reactions, each performed using a primer labeled with a different fluorescent label, were run in a single lane along with ssDNA size standards labeled with a fourth fluorescent label. The size of the microsatellite containing PCR products was determined in reference to this internal size standard.

Capillary electrophoresis. Capillary electrophoresis is a promising method for performing high-speed separation and sizing of DNA fragments. A number of different separation matrices including acrylamide, liquefied agarose, and polyvinyl alcohol have been used. Marino et al. (1994) separated SSR alleles from the soybean genotypes Williams and Jackson that differed by only 9 base pairs. The two alleles in the F_1 of Jackson × Williams were easily distinguished. Capillary electrophoresis used a 100 μM, 3% polyacrylamide gel capillary with an effective length of 40 cm. Detection was by absorbance at 260 nm. More recently the possibility of high-speed parallel capillary electrophoresis separation and sizing of DNA fragments was suggested by Clark and Mathies (1993). Such a system would allow simultaneous electrophoretic separation and sizing of SSR alleles in a number of capillaries using a sieving medium such as hydroxyethyl-cellulose. Combining this type of parallel separation technology with highly sensitive fluorescence detection promises to dramatically increase the speed and sensitivity of SSR allele size determination. These and related technologies that have and will result from the major investment in the Human Genome Project will have a major impact on plant genome research. In the near future rapid, accurate, and efficient molecular genotyping of large numbers of plant genotypes at multiple loci will be easily accomplished.

ACKNOWLEDGMENTS

The work reported here was supported in part by funds from the USDA Cooperative State Research Service, National Research Initiative, Competitive Grants Program grant 9300874.

LITERATURE CITED

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

Baron B, Poirier C, Simon-Chazottes D, Barnier C, Guenet JL (1992): A new strategy useful for rapid identification of microsatellites from DNA libraries with large size inserts. Nucleic Acids Res 20:3665–3669.

Bassam BJ, Caetano-Anollés G, Gresshoff PM (1991): A fast, single and sensitive silver-staining for DNA in polyacrylamide gel. Anal Biochem 196:80–83.

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

Browne DL, Litt M (1992): Characterization of (CA)n microsatellites with degenerate sequencing primers. Nucleic Acids Res 20:141.

Clark JM (1988): Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. Nucleic Acids Res 16:9677–9685.

Clark SM, Mathies RA (1993): High-speed parallel separation of DNA restriction fragments using capillary array electrophoresis. Anal Biochem 215:163–170.

Condit R, Hubbell SP (1991): Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34:66–71.

Cregan PB (1992): Simple sequence repeat DNA length polymorphisms. Probe Spring 1992:18–22.

Cregan PB, Akkaya MS, Bhagwat AA, Lavi U, Jiang Rongwen (1994): Length polymorphisms of simply sequence repeat (SSR) DNA as molecular markers in plants. In Gresshoff PM (ed): Plant Genome Analysis, Boca Raton, FL: CRC Press, pp 43–49.

Dietrich W, Katz H, Lincoln SE, Shin H-S, Friedman J, Dracapoli NC, Lander ES (1992): A genetic map of the mouse suitable for typing intraspecific crosses. Genetics 131:423–447.

Edwards A, Civitello A, Hammond HA, Caskey CT (1991): DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756.

Hazan J, Dubay C, Pankowiak M-P, Becuwe N, Weissenbach J (1992): A genetic linkage map of human chromosome 20 composed entirely of microsatellite markers. Genomics 12:183–189.

Herman GE (1993): A method for the isolation of microsatellite repeats from YACs using IRS-PCR and the CloneAmp system. Focus 15:51–53.

Jiang Rongwen, Cregan PB, Akkaya MS, Bhagwat AA, Lavi U (1994): The use of simple sequence repeat DNA markers for soybean genotype identification. Theor Appl Genet (in press).

Karagyozov L, Kalcheva ID, Chapman VM (1993): Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Res 16:3911–3912.

Keim P, Shoemaker RC, Palmer RG (1989): Restriction fragment length polymorphism diversity in soybean. Theor Appl Genet 77:786–792.

Keim P, Beavis W, Schupp J, Freestone R (1992): Evaluation of soybean RFLP marker diversity in adapted germ plasm. Theor Appl Genet 85:205–212.

Klinkicht M, Tautz D (1992): Detection of simple sequence length polymorphisms by silver staining. Mol Ecology 1:133–134.

Kroeker WD, Kowalski D, Laskowski M (1976): Mung bean nuclease I. Terminally directed hydrolysis of native DNA. Biochemistry 15:4463–4467.

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

Litt M, Luty JA (1989): A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401.

Marino MA, Turni LA, DelRio SA, Williams PE, Cregan PB (1994): The analysis of simple sequence repeat DNA in soybean by capillary gel electrophoresis. Appl Theo Electrophor (in press).

Morgante M, Olivieri AM (1993): PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182.

NIH/CEPH Collaborative Mapping Group (1992): A comprehensive genetic linkage map of the human genome. Science 258:67–86.

Roberts TM, Kacich R, Ptashne M (1979): A general method for maximizing the expression of a cloned gene. Proc Natl Acad Sci USA 76:760–764.

Sambrook J, Fritsch EF, Maniatis T (1989): Molecular Cloning: A Laboratory Manual, 2nd Edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Senior ML, Heun M (1993): Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. Genome 36:884–889.

Shoemaker RC, Olson TC (1993): Molecular linkage map of soybean (*Glycine max* L. Merr.) In O'Brien SJ (ed): Genetic Maps: Locus Maps of Complex Genomes. Cold Spring Harbor, Cold Spring Harbor Press, pp 6.131–6.138.

Specht JE, Shoemaker RL, Cregan PB (1994): Registration of the soybean mapping population NE-SMP-1. Crop Sci (in press).

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

Vignal A, Gyapay G, Hazan J, Nguyen S, Dupraz C, Cheron N, Becuwe N, Tranchant M, Weissenbach J (1993): Nonradioactive multiplex procedure for genotyping of microsatellite markers. In Adolph KW (ed): Methods in Molecular Genetics—Gene and Chromosome Analysis, Part A, New York: Academic Press, pp 211–221.

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

Weber JL, May PE (1989): Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396.

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992): A second-generation linkage map of the human genome. Nature 359:794–781.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990): DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535.

Wu K-S, Tanksley SD (1993): Abundance, polymorphism and genetic mapping of microsatellites in rice. Mol Gen Genet 241:225–235.

Zhao X, Kochert G (1993): Phylogenetic distribution and genetic mapping of a $(GGC)_n$ microsatellite from rice (*Oryza sativa* L.). Plant Mol Biol 21:67–614.

Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Kiehl SR (1992): Application of automated DNA sizing technology for genotyping microsatellite loci. Genomics 14:1026–1031.