Journal of Applied Biosciences 21: 1237 - 1245
ISSN 1997–5902

Microsatellite DNA marker isolation from *Picrorhiza kurrooa* Royle ex Benth by magnetic capture

*** Hussain, M.A.[1], Aggarwal, R.K.[2], Bedi, Y.S.,[1] and Abdin, M.Z.[3]**
[1] Plant Biotechnology Division, Indian Institute of Integrative Medicine, Canal Road, 180001, Jammu Tawi, India.
[2]Centre for Cellular and Molecular Biology, Hyderabad, 500007, India.
[3]Centre for Transgenic Plant Development, Department of Biotechnology, Hamdard University, New Delhi, 110062, India.

*Corresponding author email: abid_rrl@yahoo.com ; Cell: +919419110622

Published at www.biosciences.elewa.org on September 7, 2009

ABSTRACT
Objective: To establish enriched libraries by magnetic capture of microsatellites.
Methodology and results: Microsatellite containing sequences were isolated from enriched genomic libraries of *Picrorhiza kurrooa* Royle ex Benth. The enrichment was based on magnetic capture of microsatellite motifs by streptavidin coated magnetic beads and mixed biotin labeled probes hybridization capture strategy. Out of 96 colonies randomly picked for sequencing, 152 loci containing unique SSR motifs were identified. The majority of these (70%) were dinucleotide and trinucleotide repeats. The GA/GT repeat was the most common accounting for 58% of all repeat types. The average clone size of the positive clones was 500 to 1000 bp. The enrichment efficiency obtained by magnetic capture was 75% as compared 0.4 to 0.5% realized with conventional colony hybridization protocols.
Conclusion and application: Microsatellite sequences (EU883601 to EU883622 and FJ617210 to FJ617223) were successfully, and for the first time, isolated from the endangered medicinal herb *Picrorhiza kurrooa* Royle ex Benth (Royle, 1835). These markers can be applied in the assessment of genetic diversity or similarity analysis in order to assign specimens to heterotic groups on the basis of repeat length polymorphism.

Key words: magnetic capture, microsatellites, molecular marker, *Picrorhiza kurrooa*, SSR

INTRODUCTION
Microsatellites (Litt and Luty, 1989) or Simple Sequence Repeats (SSR's) (Tautz, 1989), are co-dominant markers used to identify both the alleles of a gene in a heterozygous individual. Microsatellite markers isolated de novo from a species render them species specific and are used for single-locus genotyping of a species. Multi-locus approaches may be convenient but have some technical and/or analytical drawbacks, such as dominance (i.e. only one allele identified where there is no possibility to discriminate between homozygous and heterozygous individuals). As a consequence of simultaneous visualization of many marker alleles, multi-locus data are typically analyzed by pair wise comparison of complex banding patterns that only have meaning relative to others in the same study, thus results are to a limited extent comparable among studies.

In contrast, single-locus markers are usually characterized by co-dominance (i.e. both alleles identified in heterozygous individuals) and thus are more flexible and supply more robust and comparable data (Karp, 2002). Since techniques detecting heterozygotes (i.e. co-dominant markers) and providing data on allelic differences are desirable,

microsatellites (SSR's) are considered to be the most suitable markers. The abundance of microsatellites depends upon the density of these motifs in the genome, which varies from organism to organism. Microsatellite isolation by conventional non–enriched protocols (Condit *et al.,* 1991; Akkaya *et al.,* 1992; Morgante *et al.,* 1993) based on colony hybridization for screening of a large number of colonies for microsatellite motifs is labour intensive and less effective due to the occurrence of artifacts or hybridization signals with false positives.

As a result microsatellite marker development in Pigeonpea (*Cajanus cajan*) by non-enriched genomic library approach required 753 colonies to be screened, 208 were identified as positive while only 48 (23%) actually contained a microsatellite motif when sequenced (Damaris, 2006). Traditional microsatellite isolation studies have therefore reported an average of only 2.3% final positive clones after sequencing but only after getting rid of 77% false positives (Squirrel *et al.*, 2003). To overcome this problem various strategies to increase the efficacy of mining microsatellites have been successfully applied (Zane *et al*., 2002). These strategies primarily employ magnetic capture of microsatellite motifs by streptavidin coated magnetic beads and mixed biotin labeled probes hybridization capture strategy. The establishment of enriched libraries by magnetic capture of microsatellites is a fast and low cost method for microsatellite DNA marker development.

The libraries so generated are rich in the proportion of clones containing the microsatellite repeat of interest. By this technique, microsatellite sequences (EU883601 to EU883622 and FJ617210 to FJ617223) were isolated from an endangered medicinal herb *Picrorhiza kurrooa* Royle ex Benth (Royle, 1835). The species specific microsatellite DNA markers isolated from *Picrorhiza kurrooa* can be highly useful in the assessment of genetic similarity or diversity in order to assign specimens to heterotic groups on the basis of repeat length polymorphism. *Picrorhiza kurrooa* specific microsatellite DNA markers can even be used in forensic studies for the identification of unknown market samples and for conservation studies. Our study is the first report of the isolation of species-specific microsatellite DNA markers from *Picrorhiza kurrooa.*

MATERIALS AND METHODS

Microsatellite DNA markers were isolated from *Picrorhiza kurrooa* Royle ex Benth by enriched genomic library development strategy. The protocols primarily imply hybridization of biotinylated repeat oligomers with restriction digested, adaptor ligated genomic DNA and further capturing by streptavidin coated magnetic beads, amplification using adaptor primer, and cloning in a competent vector followed by sequencing of the insert (Kijas *et al*., 1994; Fischer and Bachmann, 1998; Hamilton *et al*., 1999). The enriched genomic library that is based on selective hybridization of DNA fragments produces higher yield of microsatellite loci per hundred clones sequenced than the conventional non-enrichment protocols (Zane *et al.,* 2002).

DNA isolation: *P. kurrooa* was collected from its high altitude natural habitat in the ranges of western Himalayas of Kashmir. Total genomic DNA was isolated from fresh leaves following the protocol of Dellaporta (1983).

Size fractionation of genomic DNA: The genomic DNA (10 µg) was digested by restriction enzyme *Dra* I, *Rsa* I, *Hae* III (10 U) seperately in three reaction volumes of 200 µl each at 37°C for 8 to 12 hours in a supplier recommended buffer (Fermentas, USA). The digestion resulted in a diverse population of blunt ended restriction fragments with an average size of 200 to 1500 bp (Fig 1). After digestion the restriction enzyme was inactivated at 80°C for 10 min. The digested DNA was purified using phenol chloroform (1:1 v/v) treatment, precipitated with twice the volume of absolute alcohol and dissolved in a final volume of 20 µl Tris-EDTA (pH 8).

Preparation of double stranded adaptors: Mlu 1 adaptors were used for ligation. Mlu-Forward oligo: 5/-CTCTTGCTTACGCGTGGACTA-3/ and Mlu-Reverse oligo: 5/-TAGTCCACGCGTAAGCAAGAGCAC-3/ as proposed by Edwards *et al*. (1996) were synthesized from a commercial supplier (Sigma Aldrich). The oligos were designed to have *Mlu* 1 restriction enzyme site in it after making it double stranded. The synthetic oligos did not contain 5/ phosphate group and thus could not be used as such for ligation, thus phosphorylation was the first step in the process of preparing adaptors followed by making double stranded oligo. 600 pM of Mlu-Reverse oligo was phosphorylated with T4 polynucleotide kinase (20 units) in a 20 µl reaction volume at 37°C for 1 hour. The enzyme was denatured at 70°C for 10 min. To prepare double stranded adaptors, phosphorylated Mlu-Reverse oligo: 5/p-TAGTCCACGCGTAAGCAAGAGCAC-3/ (p = 5/ phosphate) was mixed with an equal volume of Mlu-F oligo (600 pM

each). The oligos were incubated at 25°C at room temperature for 3 hours to produce double stranded adaptor molecules and then cooled on ice.

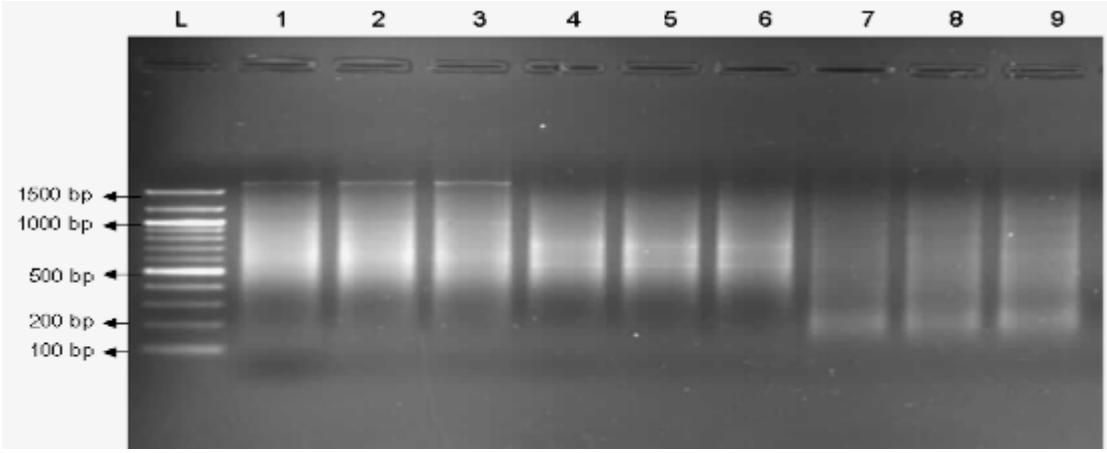

Figure 1: Restriction pattern of genomic DNA generated by three restriction enzymes (Size range 200 to 1500 bp). Lane 1 to 3 = *Dra* I digest, Lane 4 to 6 = *Rsa* I digest, Lane 7 to 9 = *Hae* III digest. L is 100 bp plus DNA ladder (Fermentas).

Ligation of adaptors to digested genomic DNA: Digested genomic DNA fragments from each of the restriction digestion above (1 µg), were resuspended in a 30 µl ligation reaction containing 5 pM of double stranded adaptor molecules and 30 Weiss units of T4 DNA ligase. Ligation was carried out at 16°C for 8-12 hours.

Setting up of an Mlu amplification PCR: PCR with Mlu-F oligo as a primer is essential as it selects the fragments, which have adaptors at the ends. During the hybridization step, this selection reduces the competition between the ligated and the unligated fragments. Preamplification of adaptor ligated products was performed in a 15 µl reaction volume using 2 µl of the ligation reaction as a template, 2 µl of 10x PCR buffer, 1 µl of 1.5 mM $MgSO_4$, 2 µl of 2 mM dNTP, 1 µl of 5 pM Mlu-F oligo and 1 µl of 1U Taq DNA polymerase. PCR was set up with an initial denaturation at 94°C for 5 min, 30 cycles for 94°C at 45 sec, 1 min annealing at 60°C for 45 sec, 72°C for 2 min and a final extension step at 72°C for 5 min. Setting gel electrophoresis of the PCR products showed a primary smear of 250-1500 bp as expected when adaptors ligate to the genomic DNA fragments at high frequency (Fig 2).

Enrichment strategy: The enrichment was based on mixed, biotin labeled probes and streptavidin coated magnetic beads hybridization capture strategy. Hybridization was performed in a total volume of 200 µl containing 100 ng of the preamplified adaptor ligated product, 5 pM each of selected biotinylated oligos $b(CA)_{15}$, $b(GA)_{15}$, $b(GAA)_{10}$, $b(CAA)_{10}$, (b = 5′ biotinylation), in a hybridization solution containing 6x SSC (0.9 M NaCl, 90 mM sodium citrate, pH 7) and 0.1% SDS (sodium dodecyl sulfate). The mixture was denatured for 15 min at 95°C and then incubated at 65°C overnight for hybridization between preamplified adaptor ligated product and selected biotinylated oligos. The hybridization reaction was then mixed with 100 mg of streptavidin coated magnetic beads (Dynabeads M-280 streptavidin, Dynal) previously equilibrated with 6x SSC and then incubated at 65°C (before use, the Dynabeads were washed and stored in a supplier recommended buffer) The mixture was incubated at 65°C for 30 min followed by incubation at 70°C for 2 hours. The beads were separated on a magnetic separator provided with the kit components (Invitrogen), the supernatant was removed by pipetting, the beads were washed four times with 1 ml 2x SSC followed by four washings with 1 ml 1x SSC. After washing with SSC, the beads were rinsed briefly with 100 µl Tris-EDTA. Finally the beads were suspended in 50 µl Tris-EDTA and used directly for PCR.

PCR amplification and cloning in pJET1.2/blunt cloning vector: One microliter 1 µl of the beads were used as a template for 30 cycles of PCR in a 15 µl reaction volume using primer Mlu-F (5 pM) as described in Mlu amplification PCR. PCR products with 3′-dA overhangs generated using Taq DNA polymerase were blunted with DNA blunting enzyme included in the kit (CloneJET™ PCR Cloning Kit) prior to ligation. One microliter 1 µl of the blunting enzyme was used for 2 µl of the PCR product in 10 µl of the reaction buffer in a total volume of 18 µl. The reaction was incubated at 70°C for 5 min and then kept on ice. The

pJET1.2/blunt cloning vector was linearized with EcoRV (Eco321) to generate blunt ends. The enzyme was then denatured at 70°C for 10 min. One microliter 1 µl of the linearized vector was ligated with 18 µl of the insert in the presence of 1 µl T4 DNA Ligase (5 $U\mu l^{-1}$) at 22°C for 30 min.

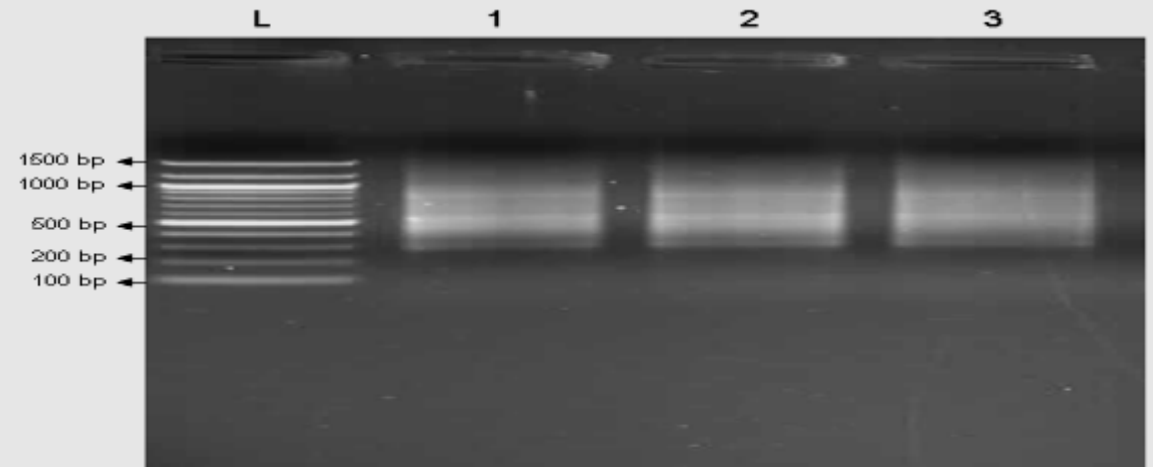

Figure 2: Adaptors successfully ligated to the genomic DNA (Size range 250 to 1500 bp). Lane 1 = Dra I digest, Lane 2 = *Rsa* I digest, Lane 3 = *Hae* III digest. L is 100 bp plus DNA Ladder.

The ligated products were used to transform electro competent *E. coli* strain DH5α which were then plated on LB medium containing 50 $\mu g ml^{-1}$ ampicillin. Recombinant colonies were identified by positive selection through insertional inactivation of the *ccd* B (control of cell death) open reading frame. The vector contains the lethal restriction enzyme gene which is disrupted by ligation of a DNA insert in to the cloning site as a result of which only cells with recombinant plasmids are able to propagate (CloneJET™ PCR Cloning principle, Fermentas, USA). Colonies were transferred to 96 well micro titer plates containing 200 µl LB/glycerol (1:1 ratio) incubated overnight at 37°C and kept for long term storage at -70°C.

Check for insert: Plasmids were isolated from the colonies using plasmid isolation kit (Qiagen, USA). Inserts were amplified using primers flanking the insert. Five picomoles 5 pM each of pJET-F and pJET-R primers was used in a 15 µl reaction volume, containing 250 ng plasmid DNA. PCR was set up with an initial denaturation at 94°C for 5 min, 35 cycles of 94°C for 45 sec, 55°C annealing for 45 sec, 72°C extension for 2 min and the final extension step at 72°C for 5 min. Three microliter 3 µl of the PCR product was resolved on 1.5% agarose gel along with 100 bp DNA ladder to check for the insert (Fig 3).

Sequencing of the insert: pJET-F and pJET-R oligos were used as primers for sequencing of the forward and reverse strands of the insert, respectively. The sequencing was performed in a 5 µl reaction volume containing 1 µl (100 ng) of the PCR product, 0.3 µl forward sequencing primer pJET-F (5 pM) and 2 µl of the Big Dye terminator mix (Big Dye terminator v3.1 cycle sequencing kit, Applied Biosystems). Sequencing PCR was set for 30 cycles at 95°C for 10 sec, 50°C for 5 sec and 60°C for 4 min in a thermocycler (Eppendorf, Germany) and the extension products were processed for purification.

To 5 µl product of the sequencing PCR, 95 µl of sterile Millipore water was added to make up the final volume of 100 µl. Extension product was precipitated by adding 200 µl ethanol and 10 µl of 3 M sodium acetate (pH 5.6). The contents were mixed gently and centrifuged at 12,000 × g for 20 min, the supernatant was discarded and the pellet was washed with 250 µl of 70% ethanol twice at 12,000 × g for 5 min. The supernatant was removed and the pellet air-dried. The pellet was dissolved in 12 µl of 10% formamide and incubated at 95°C for 2 min on a heat block. The tubes were immediately put on ice before loading the samples on the DNA sequencer. The samples were loaded onto the 96-capillary 3730x DNA analyzer (Applied Biosystems) for sequencing using the Big Dye Terminator chemistry (Fig 4). The DNA sequences thus obtained were edited, assembled and analyzed by autoassembler software (Sequencher 4.9, Macintosh, USA).

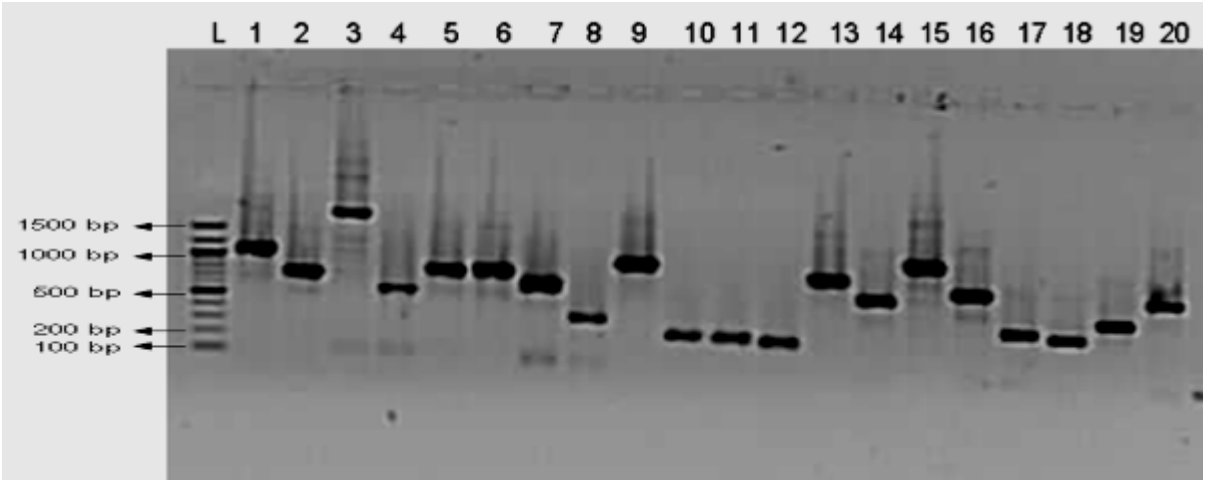

Figure 3: Insert size variation of plasmid library clones resolved in 1.5% agarose gel electrophoresis. L is 100 bp plus DNA ladder.

RESULTS AND DISCUSSION

A total of 96 clones were sequenced. A search for microsatellite repeat motifs was performed with the online software tool, Tandem Repeats Finder (Benson, 1999). Primers were designed in pairs called forward and reverse flanking the repeat region using online bioinformatic tool Primer 3 software. The length of a primer was kept 18-24 nucleotides long as is considered ideal and highly specific. Primers were selected with melting temperatures within 1–4^{o}C between each other and in the 40–45^{o}C range. It is recommended for each forward and reverse primer to have a similar melting temperature, at which one half of the primer-dimer duplex will dissociate to single stranded form, so that they can both work under optimal conditions in a reaction. It is considered that none of the primers should have a 3/ end as T, primers with a 3/ T have greater tolerance of mismatch. Further it is advisable for each primer to have at least one A or T within the 3/ most triplet codon to prevent mismatch tolerance of primers with consecutive G's or C's. A prime consideration is that the primers should be complex enough so that the likelihood of annealing to sequences other than the chosen target is very low. Both the forward and the reverse primer sequences are designed from the coding strand of the DNA molecule. The reverse primer must be the reverse compliment of the coding strand target site (Rychlik, 1993). In terms of the frequency of the distribution of microsatellite repeats, the dinucleotide SSR's were the most abundant (58%), trinucleotide SSR's occurred in about 12% of the cases, 6% mononucleotide repeats and 3% hexanucleotide repeats (Fig 5). Other types of SSR's had only <1% representation. Out of 96 clones, 75 clones were non-redundant sequences. The GA/GT repeat motif had the highest frequencies in the library, followed by TG/CA; CT; AC/TC; AG/AT; AAG/CCT/CTT; GAA; and A/T repeats. The GA/GT repeat is reported to be rich in other plant species (He *et al*., 2003), and was found to be the most frequently dispersed SSR's of *Picrorhiza kurrooa* (Fig 6)

Overall, 36% of the microsatellites sequenced were perfect, 25% were imperfect and 14% were compound, following Weber's (1990) definitions that state that perfect repeats have no interruptions, imperfect repeats are interrupted by non-repeat bases, and compound repeats consist of two or more repeat runs adjacent to each other. The probes used during enrichment were synthesized and 5/ biotinylated by a commercial supplier (Sigma Aldrich). All of the four probes: $(CA)_{15}$, $(GA)_{15}$, $(GAA)_{10}$, $(CAA)_{10}$ were directly related to these sequences, with probe $b(GA)_{15}$ providing the highest frequency of repeats. Since biotin has got a very high affinity towards streptavidin therefore biotin labeled probes attached to the complimentary repeat region in the genomic DNA were readily pulled out using streptavidin coated magnetic beads on a magnetic surface (Fig 7).

Our protocol based on selective hybridization uses microsatellite bound magnetic beads directly for the PCR and is a slight modification of the earlier protocols. The modified protocol skips further washing and purification steps to remove bound microsatelite motifs from the beads. These results reveal the relative abundance of different repeat motifs in *Picrorhiza kurrooa* which are highly mutable and show substantial variation in size polymorphism. With this strategy we were able to achieve a constant success rate of 75% enrichment, which is significantly higher compared to the conventional protocols.

.

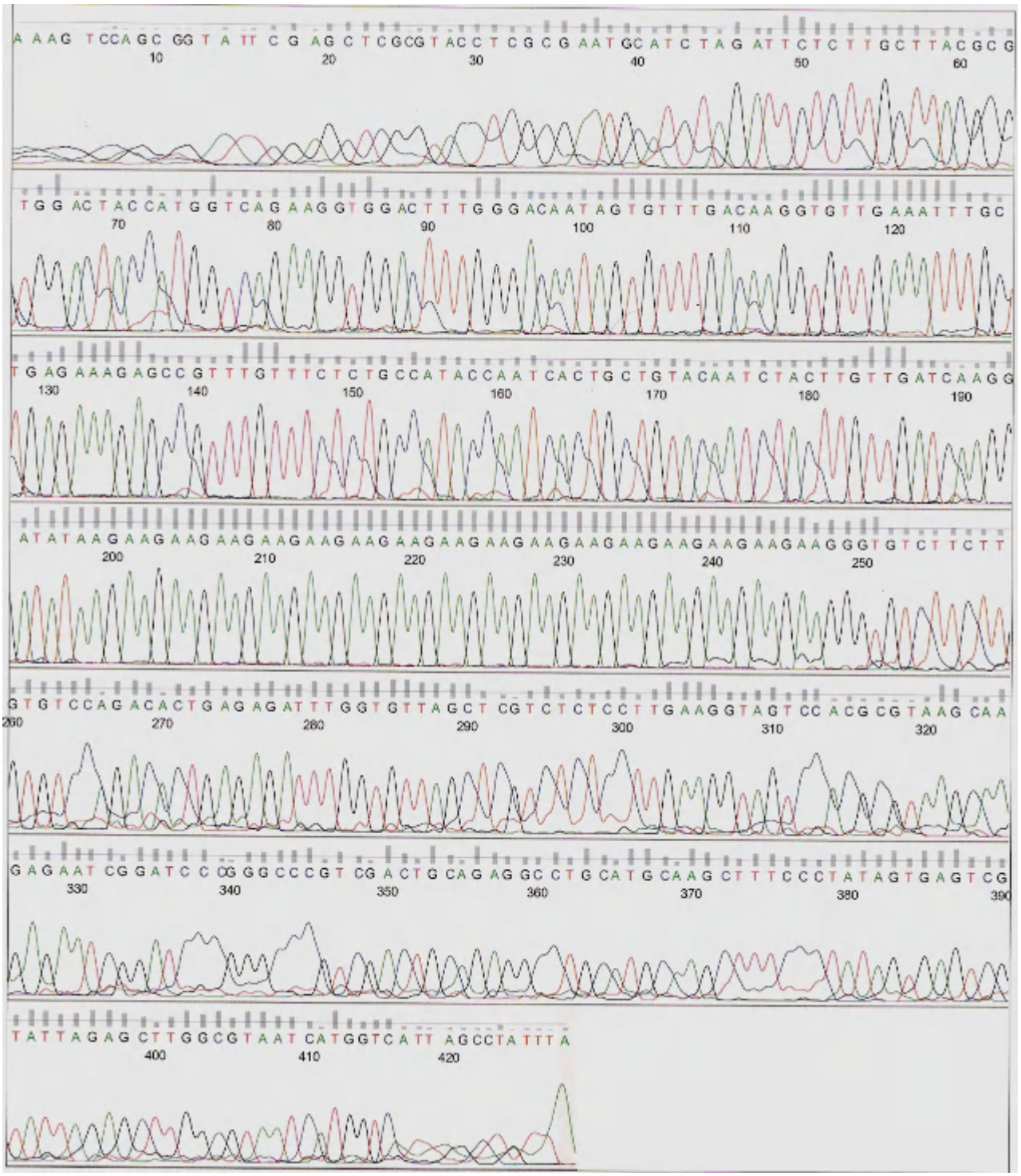

Figure 4: A typical electrophoretogram showing the microsatellite repeat motif (AAG)17 present in the enriched genomic library of *Picrorhiza kurrooa* at locus PKssrD2 (EU883611) and its flanking sequence.

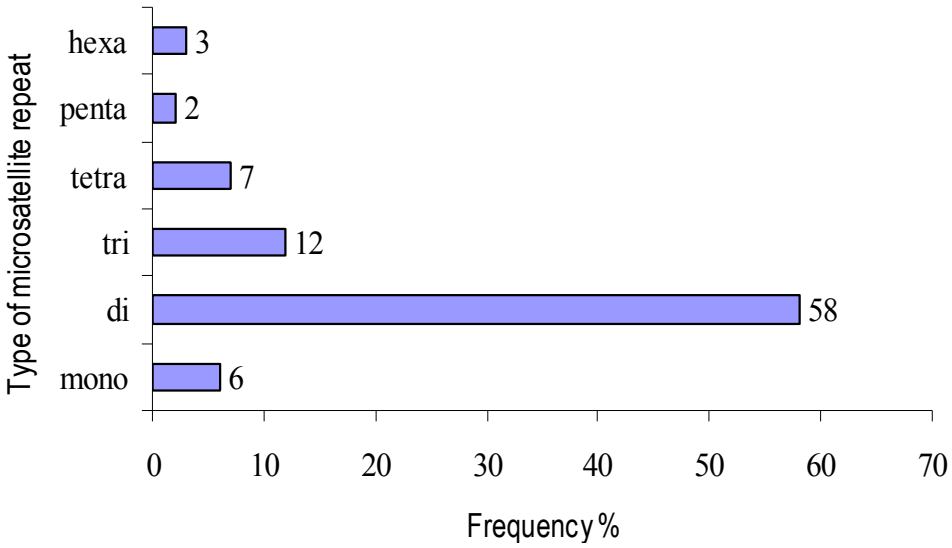

Figure 5: Frequency of different SSR classes in *Picrorhiza kurrooa.*

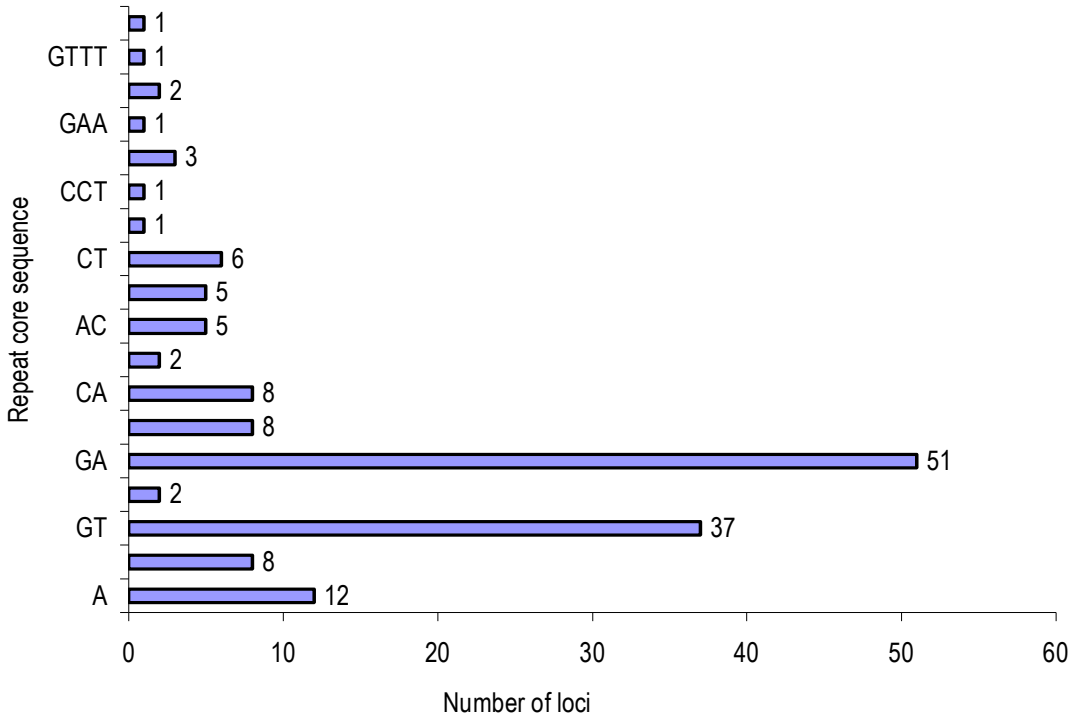

Figure 6: Frequencies of SSR loci in *Picrorhiza kurrooa* with different core sequences.

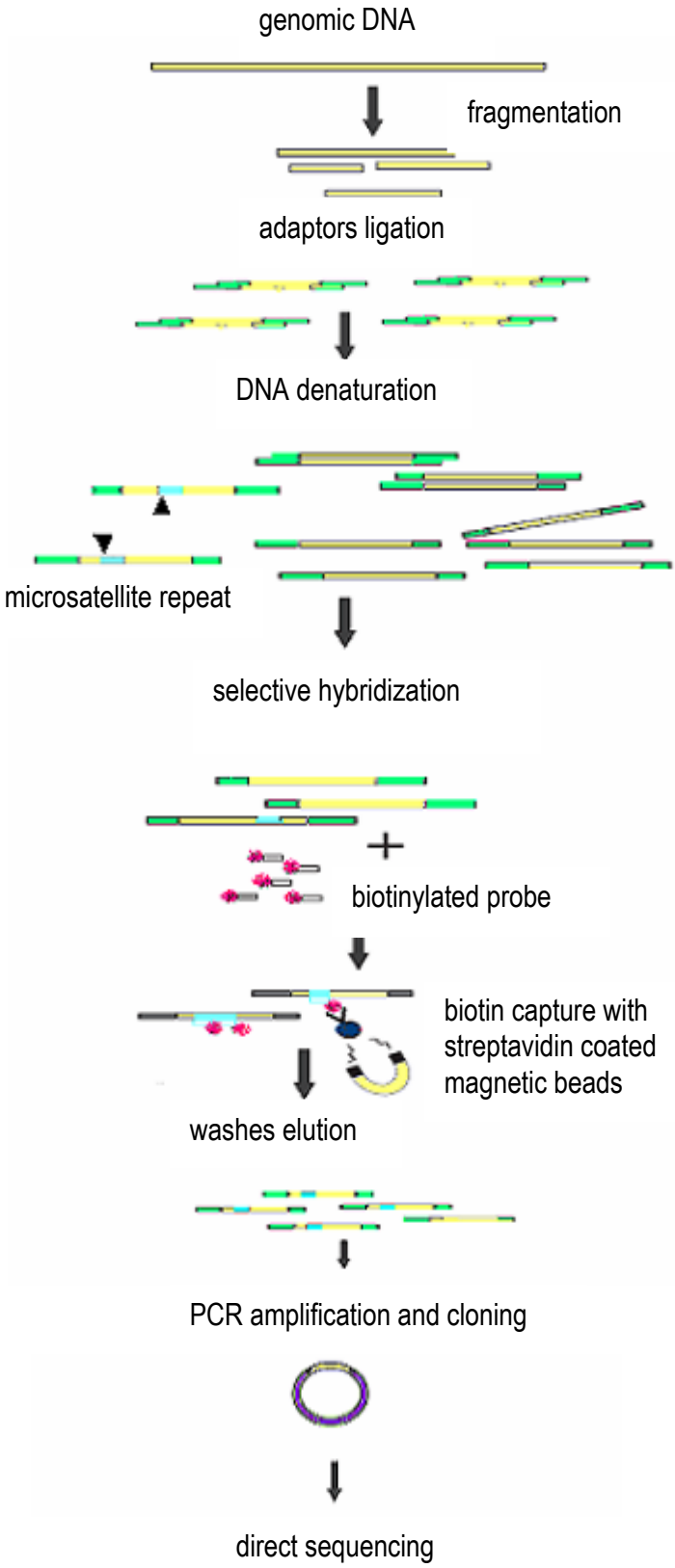

Figure 7: Scheme for microsatellite enrichment by selective hybridization.

ACKNOWLEDGEMENTS: The author gratefully acknowledges the Council of Scientific and Industrial Research, India for providing financial assistance to the project in the form of fellowship.

REFERENCES

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

Benson G, 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research Vol. 27, No. 2, pp. 573 - 580.

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

Dellaporta SL, Wood J, Hicks JB, 1983. A plant DNA minipreparation: version II. Plant Mol Biol Rep 1: 19 - 21.

Damaris Achieng Odeny, 2006. Microsatellite development and application in Pigeonpea (*Cajanus cajan* (L.) Millspaugh). Ecology and development series, No. 45.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A, 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20: 758 - 760.

Fischer D and Bachman K, 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). BioTechniques 24: 796 - 802.

Hamilton MB, Pincus EL, Di Fiore A, and Fleischer RC, 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. BioTechniques 27: 500 - 507.

He Guohao, Meng Ronghua, Newman Melanie, Gao Guoqing, Pittman RN, Prakash CS, 2003. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.). BMC Plant Biology, 3: 3.

Karp A, 2002. The New Genetic Era: Will it help us in Managing Genetic Diversity? In: Managing Plant Genetic Diversity. (Eds: Engels JMM, Ramanatha Rao V., Brown AHD, Jackson MT), IPGRI Rome, pp. 43 - 56.

Kijas JMH, Fowler JCS, Garbett CA, Thomas MR, 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. BioTechniques 16: 657- 662.

Litt M and 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- 40.

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

Rychlik W, 1993. Selection of primers for polymerase chain reaction, in PCR Protocols. Current Methods and Applications (White, B. A., ed.), Humana, Totowa, NJ, pp. 31– 40.

Royle JF, 1835. Illustration of botany of Himalayan Mountain, pp. 291.

Squirrel J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W, 2003. How much effort is required to isolate nuclear microsatellites from plants. Mol Ecol. 12: 1339 - 1348.

Tautz D, 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463 - 6471.

Weber JL, 1990. Informativeness of human (dC-dA)n. (dG-dT)n polymorphism. Genomics 7: 524 - 530.

Zane L, Bargelloni L, Patarnello T, 2002. Strategies for microsatellite isolation: a review. Mol Ecol 11:1 - 16.