Microsatellite allele size profiling to determine varietal identity and genetic diversity among groundnut varieties in Bangladesh

MR Molla[1]*, MN Islam[2], MM Rohman[3] and L Rahman[4]

[1]Plant Genetic Resources Centre, Bangladesh Agricultural Research Institute, Joydebpur, Gazipur, Bangladesh
[2]Department of Biotechnology, Sher-e-Bangla Agricultural University, Dhaka
[3]Molecular Breeding Laboratory, Bangladesh Agricultural Research Institute, Joydebpur, Gazipur, Bangladesh
[4]Department of Genetics and Plant Breeding, Bangladesh Agricultural University, Mymenisingh
rezwanbt@gmail.com, motiar_1@yahoo.com

Abstract: Cultivated peanut or groundnut (*Arachis hypogaea* L.) is an important crop for oil and protein source in Bangladesh. A number of varieties have so far been released or registered for cultivation in the country. These have been distinguished only through morphological traits. Microsatellite markers, also known as SSRs (Simple Sequence Repeats), have proved to be an excellent tool for the identification of the plant varieties and determining genetic relationship between the varieties of a crop species. A set of three SSR markers namely, PM36, PM50 and PM238 were used to identify ten cultivated groundnut varieties (Dhaka-1, Bashanti, Tridana, Zhinga badam, BARI badam 5,6,7; BINA Cheenabadam 1,2,3) available in Bangladesh. All the cultivars were successfully discriminated by these three SSR primers. The primer PM50 alone was able to distinguish four varieties (Dhaka-1, Bashanti, Tridana and Zhinga Badam). Six variety- specific alleles were identified, these are, PM36/227, PM50/110, PM50/116, PM50/118, PM50/137, PM238/200. The three primers produced a total of 13 alleles with size ranging from 109bp to 241bp. The PIC (Polymorphism Information Content) value for the primer PM36, PM50 and PM238 was found 0.81, 0.76 and 0.82 respectively. This approach will be useful for developing a set of limited number of SSR loci for the identification of commercially important groundnut varieties for purpose of obtaining plant variety protection (PVP) in Bangladesh.

[Molla MR, Islam MN, Rohman MM, Rahman L. Microsatellite Allele Size Profiling to Determine Varietal Identity and Genetic Diversity among Groundnut Varieties in Bangladesh. Nature and Science 2010; 8(12):123-129]. (ISSN: 1545-0740). http://www.sciencepub.net.

Keywords: Groundnut, Microsatellite marker, Variety Identification.

1. Introduction

The cultivated peanut (*Arachis hypogaea*) or groundnut is an important legume crop across the Americas, Africa and Asia, where it is grown for local consumption and international trade, as a food and oil product. Cultivated groundnut (*Arachis hypogaea* L.) is an important crop for oil and protein source in Bangladesh. Nearly all *Arachis* species are diploid, but cultivated peanut (*Arachis hypogaea* L.) is an allotetraploid (genome AABB). Although there is high level of morphological diversity among varieties of *A. hypogaea*, this has not been generally reflected in the level of detectable genetic diversity (Halward et al. 1992; Lacks and Stalker, 1993; Lanham et al. 1994). With the introduction of Plant Variety and Farmers Right Protection Act (PVFRPA), there is a need for adequate and well characterization and documented information of the varieties already approved for production in the country. Proper cataloguing of the varieties will decrease the chances of loss of our genetic materials of high potentials either through piracy or through informal system of sharing the benefit. In recent years it has been proved beyond doubt that only identification of crop varieties by quantitative terms is not adequate. The visible distinctness is shown mostly through qualitatively controlled morphological traits. It seems important to start Genetic Fingerprinting of the varieties developed so that the SSR marker can indicate the variety position in addition to its morpho-physiological traits that are now being used as criteria for variety release/registration. Molecular marker analysis, joined to phenotypic evaluation, is a powerful tool for grouping of genotypes based on genetic distance data and for selection of progenitors that might constitute new breeding populations. However, little variation has been detected at the DNA level using techniques such as RAPDs (Random Amplified Polymorphic DNAs), AFLPs (Amplified Fragment Length Polymorphisms) and RFLPs (Restriction Fragment Length Polymorphisms) (Kochert et al., 1996; Hilu et al., 1995; He and Prakash, 1997; Subramanian et al., 2000; Gimenes et al., 2002; Herselman , 2003). Among this marker SSRs are ideal tools for such studies as they are PCR-based markers,

genetically defined, typically co-dominant, multiallelic, and uniformly dispersed throughout plant genomes. In *Arachis*, SSR markers have been recently developed and proved to be useful for accession discrimination and assessment of genetic variation. Moreover, it is also used to protect intellectual property of the plant breeder (plant breeders' rights), which encourages the continuous development of new varieties. Varieties eligible for certification have resulted either from natural selection or through systematic plant breeding techniques. In either case, without a planned method for maintaining genetic purity, there is grave danger of losing varietal identity. Hence the advent of plant variety protection lends added urgency to the search for solutions to the conservation of plant genetic diversity. Molecular markers have been successfully applied in registration activities like cultivar identification (Mailer et al., 1994), or controls of seed purity of hybrid varieties (Marshall et al., 1994). However, with the necessity of varietals identity and protection against adulteration, present study was an attempt to reveal genetic variation and varietal identification of 10 groundnut varieties in Bangladesh using Microsatellite DNA markers.

2. Materials and Methods

2.1 Genomic DNA isolation:

Seeds of 10 groundnut varieties were collected from Bangladesh Agricultural Research Institute and Bangladesh Institute of Nuclear Agriculture (Table 1). Genomic DNA was isolated following protocol described by Saghai-Maroof et al. (1984) with some modifications. Juvenile leaves (unfolded) of 30 days old plants were used as source of Genomic DNA. Leaf tissues were cut into small pieces, homogenized and digested with extraction buffer (pH= 8.0): 50 mM Tris-HCl, 25 mM EDTA (Ethylenediaminetetraacetic acid), 300 mM NaCl and TEN buffer + 5% SDS (Sodium Dodecyl Sulfate) +10% PVP (Poly Vinyl Pyrolideone) +20% CTAB (Cetyl Trimethyl Ammonium Bromide)]. After incubation for 20 minutes at 650C with intermittent swirling, the mixture was emulsified with an equal volume of phenol: chloroform: isoamyl alcohol (25 : 24 : 1, v/v/v). DNA was precipitated using two volume of absolute alcohol in presence of 0.3M sodium acetate and pelleted by centrifugation. The pellets were then washed with 70% ethanol, air dried and resuspended in an appropriate volume of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH=8.0) treated with 2 µl of RNAse for removing of RNA. DNA quality was checked by electrophoresis in a minigel and quantification was accomplished using a spectrophotometer (Spectronic® Genesis™, Spectronic Instruments Inc., USA).

Table 1. List of variety used in this study

Sl. No.	Variety	Year of Release	Source
1.	Dhaka-1	1976	Bangladesh Agricultural Research Institute
2.	Bashanti (DG-2)	1979	Bangladesh Agricultural Research Institute
3.	Tridana (DM 1)	1987	Bangladesh Agricultural Research Institute
4.	Zhinga badam	1988	Bangladesh Agricultural Research Institute
5.	BARI badam 5	1999	Bangladesh Agricultural Research Institute
6.	BARI badam 6	1999	Bangladesh Agricultural Research Institute
7.	BARI badam 7	2004	Bangladesh Agricultural Research Institute
8.	BINA Cheenabadam 1	2000	Bangladesh Institute of Nuclear Agriculture
9.	BINA Cheenabadam 2	2000	Bangladesh Institute of Nuclear Agriculture
10.	BINA Cheenabadam 3	2000	Bangladesh Institute of Nuclear Agriculture

Table 2: Primers uses in this study

Sl.#	Locus	Forward Primer	Reverse Primer	Ann.T.	Ref.
1.	PM36	act-cgc-cat-agc-caa-caa-ac	cat-tcc-cac-aac-tcc-cac-at	48°C	Guohao et al. (2003)
2.	PM50	caa-ttc-atg-ata-gta-ttt-tat-tgg-aca	ctt-tct-cct-ccc-caa-ttt-ga	48°C	Guohao et al. (2003)
3.	PM238	ctc-tcc-tct-gct-ctg-cac-tg	aca-aga-aca-tgg-gga-tga-aga	48°C	Guohao et al. (2003)

2.2 Microsatellite Markers and PCR amplification:

A set of five microsatellite primer pair (PM3, PM36, PM50, PM150 and PM238) have been used to estimate the potential of these marker for variety identification. Finally three primer pairs PM36, PM50 and PM 238 were selected based on their performance for SSR data analysis (Table 2). Polymerase Chain Reactions were done in a volume of 10 µl containing 1µl 10 x PCR

Buffer, 0.25 mM each of the dNTPs, 10 µM of each of primer, 1 unit ampli Taq DNA polymerase, 75 ng template DNA and a suitable amount of sterile deionized water. Amplification were carried out in a oil free thermal cycler (Thermal cycler gradient, Eppendorf) with the following thermal profile: initial denaturation step at 94 ^{o}C for 3 min followed by 40 cycles at 95 ^{o}C for 30 sec., 48oC for 30 sec, and 72oC for 1 min and a final cycle at 72 ^{o}C for 7 min. PCR products were confirmed by electrophoresis on 2% agarose gel.

2.3 Electrophoretic separation and visualization of PCR products:

PCR-products were electrophoresed on a 6% denaturing polyacrylamide gel containing 19:1 acrylamide: bis-acrylamide and 8M urea. Electrophoresis was conducted using the SequiGen GT sequencing gel electrophoresis system (BIO-RAD Laboratories, Hercules, CA). A pre-run of the gel for 30 mins at 120W was followed by a final run at 60W and 50ºC upon loading of denatured PCR products for a specified period of time depending on the size of amplified DNA fragment (usually 1 hour for 100 bp). A molecular weight marker DNA (100 bp DNA ladder, Biolabs, New England) was loaded on either side of the gel. After completion of electrophoresis, the DNA fragments were visualized following the Promega (Madison, WI) silver-staining protocol.

2.4 Analysis of microsatellite data:

The bands representing particular alleles at the microsatellite loci were scored manually and designated the bands as A, B, C etc. from the top to the bottom of the gel. The genotypes of different strains were scored as AA, BB, CC, etc. for homozygous or as AB, AC, BC etc. for heterozygous. A single genotypic data matrix was constructed for all loci. Polymorphism Information Content (PIC) was computed by adding the square values for all the frequencies of different alleles produced by a single marker locus and then deducted from one (PIC = 1- ΣXi^2, where, X_i is the frequency of the i-th allele of a particular locus). PIC provides an estimate of the discriminatory power of a marker by taking into account, not only the number of alleles that are expressed, but also the relative frequencies of those alleles. PIC values range from 0 (monomorphic) to 1 (very high discriminative, with many alleles in equal frequencies (Smith et al., 1997). The software DNA FRAG version 3.03 (Nash, 1991) was used to estimate marker length and allelic length. Expected (He) and observed heterozygosity (H_o) were also calculated using the software, POPGENE (version 1.31) (Yeh et al., 1999). Estimation of Nei's (1972) genetic distance and construction of a dendrogram was done using the computer programme POPGENE (Version 1.31) (Yeh et al., 1999).

3. Results and Discussion

3.1 Microsatellite polymorphism

All three microsatellite markers were found to be polymorphic, revealing a total of 19 alleles with an average number of 6.33 alleles per locus were found in the present study. Number of alleles was highest for locus PM238 (7) and lowest for PM36 and PM50 (6). Allele sizes were ranged between 186-241 bp, 109-137bp and 193-219 bp for the loci PM36, PM50 and PM238, respectively. Eight alleles ranging between 190 and 240bp for PM36, eight alleles ranging between 94 and 224 bp for PM50 and four alleles ranging between 164 and 170 bp for PM238 in groundnut cultivars were reported (Guohao et al., 2003). Observed allelic lengths which were almost similar to the previous study although some variation occurred might be due to mutation of di-nucleotide repeat units. The average number of alleles of the present study was comparison to the previous study in groundnut. The possible reason for this result might due to the less number of diverse groundnut varieties used in this study. Allele frequency ranged from 0.100 to 0.400 in the present study (Table 3).

Polymorphism Information Content (PIC) values for SSRs are presented in Table 3. PIC values in the present study ranged from 0.76 to 0.82 which were lower for than the average PIC value reported by Guohao et al. (2003). High PIC values were observed might be due to use of di-nucleotide repeats and also due to varietal differences.

3.2 Allele size profiling and cultivar identification

Allele sizing technologies are well established and can be readily used to size microsatellite alleles from any organism (Song et al., 1999). SSR genotypic data from a number of loci have the potential to provide unique allelic profiles or DNA fingerprints for precisely establishing genotypic identity. Comparisons between SSR band positions against each marker in this study are shown in Table 4. The band patterns corresponding to individual variety may help to recognize the variety in question. When one primer would not distinguish individual variety from others, another primer should be considered and sometimes combination of more than one primer should be taken into account. Thus additional primer or set of primers might be needed to test to identify all expected varieties. All the ten groundnut varieties were discriminated successfully by the three SSR markers. Among 19 alleles detected, six were specific to groundnut variety. Most of the unique genotypes used for varietal identification were found at locus PM50. Locus PM50 alone discriminated four varieties (Dhaka -1, Bashanti, Tridana, Zhinga badam). One specific allele was detected in the variety Dhaka -1 (116 bp),

Bashanti (118 bp), Tridana (110 bp), Zhinga badam (137 bp) (Table 4). Besides, BARI badam 5 and BARI badam 6 could be easily identified in combination of the primer PM36 and PM50, in which PM36 showed heterozygous condition. For instance, BARI badam 7 showed a unique DNA band of 227 bp and 200 bp size for the primer PM36 and PM238 respectively. In combination of two alleles PM36215+186 and PM50112 also identified BINA Cheenabadam 1 and BINA Cheenabadam 3 respectively. On the other hand, combination of three primers identifies the variety BINA Cheenabadam 2 (Table 4). Our results represent one of the first attempts to find out a small set of microsatellite makers to discriminate groundnut varieties of Bangladesh providing meaningful data that can be enlarged by additional groundnut varieties and new microsatellite markers.

3.3 Nei's genetic distance:

The varieties Dhaka 1 vs BARI badam 6 and BARI badam 7 showed the highest genetic distance value (2.080). The minimal genetic distance (0.000) was observed between Tridana (DM 1) vs Dhaka 1, Bashanti (DG-2), Zhinga badam, BARI badam 5 and BINA Cheenabadam 1; BINA Cheenabadam 2 vs Dhaka 1, Bashanti (DG-2), Zhinga badam, BARI badam 5 and BINA Cheenabadam 3 vs Dhaka 1, Bashanti (DG-2), Zhinga badam, BARI badam 5 (Table 5). High genetic distance values between variety pairs were found due to different genetic background while least/nil values indicated they are genetically more similar.

Table 3. Size and frequency of alleles at three microsatellite loci in 10 groundnut (*Arachis hypogaea* L.) varieties in Bangladesh

Locus	Number of alleles	Allele size (bp)	Allele frequency	Gene Diversity (PIC=1-ΣXi^2)
PM36	6	241	0.150	0.81
		227	0.050	
		221	0.150	
		215	0.150	
		189	0.250	
		186	0.250	
PM36	6	137	0.100	0.76
		118	0.100	
		116	0.100	
		112	0.400	
		110	0.100	
		109	0.200	
PM238	7	219	0.100	0.82
		214	0.100	
		206	0.300	
		202	0.150	
		200	0.050	
		195	0.200	
		193	0.200	

Table 4: Analysis of three microsatellite loci for 10 Groundnut varieties

SL. No.	Cultivars	Band positions due to primers (bp)																				
		PM 36							PM 50							PM 238						
		A	B	C	D	E	F	GN	A	B	C	D	E	F	GN	A	B	C	D	E	F	G	GN
1	Dhaka-1			221			186	CF			**116**				CC			206			195		CF
2	Bashanti (DG-2)			221			186	CF		**118**					BB			206			195		CF
3	Tridana (DM 1)	241				189		AE					**110**		EE	219			202				AD
4	Zhinga badam				215		186	DF	**137**						AA			206			195		CF
5	BARI badam 5				215		186	DF						109	FF			206			195		CF
6	BARI badam 6			221		189		CE						109	FF		214					193	BG
7	BARI badam 7		**227**			189		BE				112			DD			206		**200**			CE
8	BINA Cheenabadam 1				215		186	DF				112			DD			206				193	CG
9	BINA Cheenabadam 2	241				189		AE				112			DD	219			202				AD
10	BINA Cheenabadam 3	241					186	AE				112			DD		214		202				BD

3.4 Phylogenetic dendogram

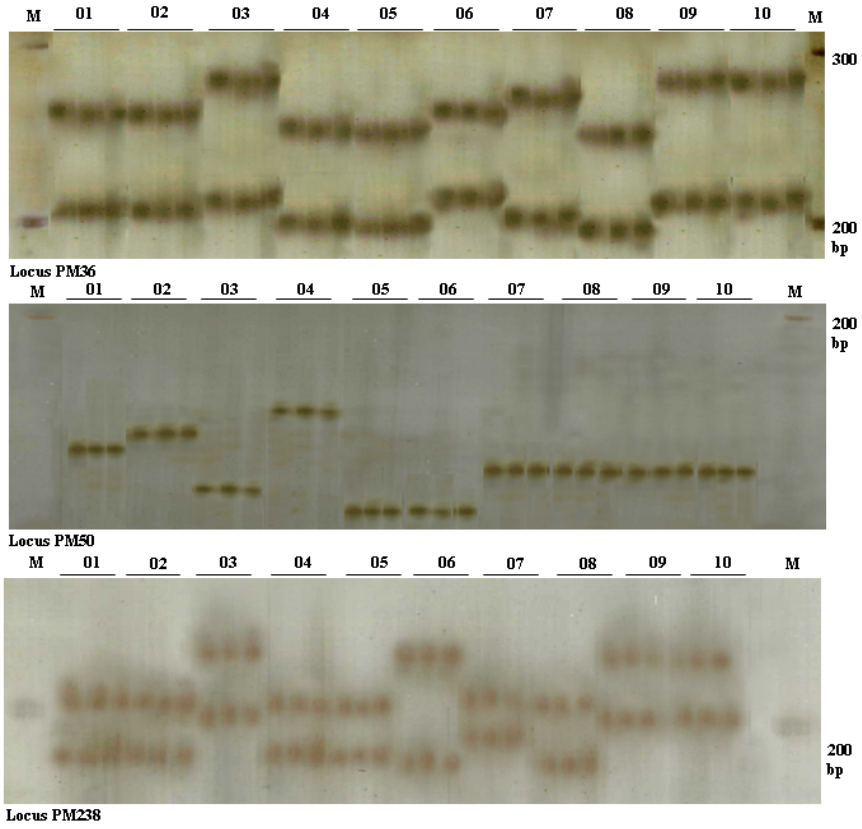
UPGMA dendogram based on Nei's (1972) genetic distance indicated segregation of 10 varieties of groundnut into two main clusters: In cluster 1 Dhaka-1, Tridana (DM-1); Bashanti, BINA cheena badam-2 and BARI badam-5, BINA cheena badam-3 grouped in sub-cluster I and zhinga badam and BARI badam 6 alone formed sub-cluster II. BARI badam 7 and BINA cheenabadam-1, grouped in cluster 2 (Fig. 2). Formation of major clusters and subclusters in the phylogenetic dendogram in our study was according to their ancestor. Dhaka 1, BINA cheena badam 1, BINA

cheena badam 2 and BARI badam 5 were formed into one cluster which might due the varieties developed through mutation breeding from the variety Dhaka 1. Upon subsequent separation, the varieties, Dhaka-1, Tridana (DM-1); Bashanti, BINA cheena badam-2 and BARI badam-5, BINA cheena badam-3 grouped together in one sub-cluster which also probably due to similar type groundnut on the basis of use and separated from other cultivars grouped in another sub-cluster because there might have effect of growth habit, peg colour, pod reticulation and seed colour in combination with SSR band patterns.

The data obtained from such kind of study can be used for IPR to protect groundnut varieties of Bangladesh in addition to the breeder's morphological information. The format can be helpful for plant variety protection act in Bangladesh with necessary information. The promising germplasm should be characterized at more loci and a set of least number of informative loci to be identified for variety identification. Genetic diversity information could be used for the breeder to develop a new variety.

Table 5. Summary of Nei's (1972) genetic diversity values between 10 Groundnut varieties for all loci

Variety	Dhaka-1	Bashanti (DG-2)	Tridana (DM 1)	Zhinga badam	BARI badam 5	BARI badam 6	BARI badam 7	BINA Cheenabadam 1	BINA Cheenabadam 2
Bashanti (DG-2)	0.693								
Tridana (DM 1)	0.000	0.000							
Zhinga badam	0.981	0.981	0.000						
BARI badam 5	0.981	0.981	0.000	0.693					
BARI badam 6	2.080	2.079	2.079	0.000	0.693				
BARI badam 7	2.080	2.079	2.079	2.079	2.079	2.079			
BINA Cheenabadam 1	1.386	1.386	0.000	0.981	0.981	2.079	0.470		
BINA Cheenabadam 2	0.000	0.000	0.693	0.000	0.000	2.079	0.470	0.693	
BINA Cheenabadam 3	0.000	0.000	0.981	0.000	0.000	1.386	0.470	0.693	0.134



Lane	Variety
01	Dhaka-1
02	Bashanti (DG-2)
03	Tridana (DM 1)
04	Zhinga badam
05	BARI badam 5
06	BARI badam 6
07	BARI badam 7
08	BINA Cheenabadam 1
09	BINA Cheenabadam 2
10	BINA Cheenabadam 3

Figure 1. Microsatellite profiles of Ten Groundnut (*Arachis hypogaea* L.) varieties at locus PM36, PM50 and PM238; M: Molecular wt. Marker (100 bp DNA ladder)

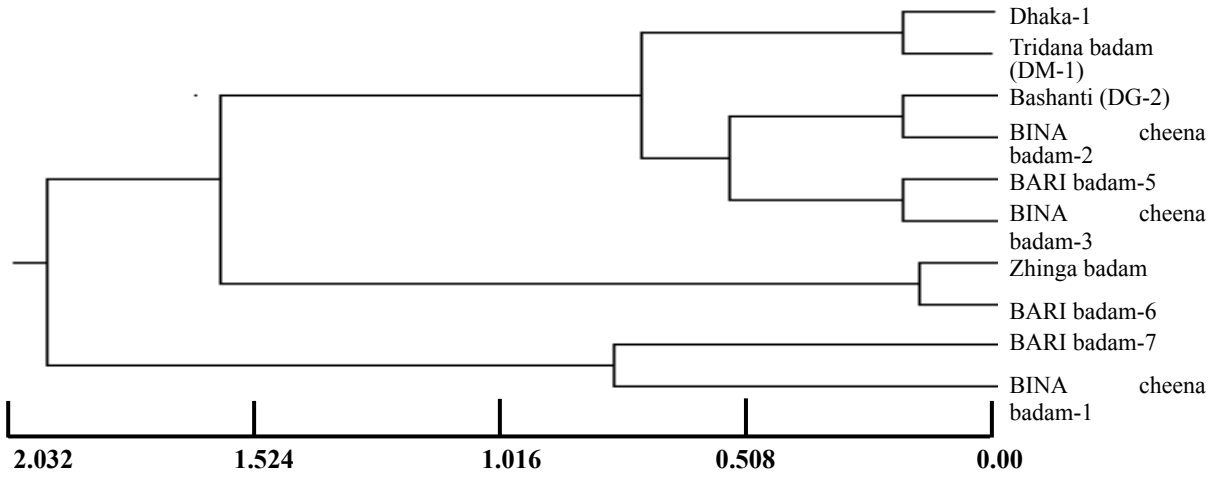


Figure 2. UPGMA dendrogram based on Nei's (1972) genetic distance, summarizing the data on differentiation between 10 Groundnut varieties according to microsatellite marker analysis.

4. Acknowledgement

This research presented here was supported by the institution, Seed Industry Development under the Ministry of Agriculture, Government of the Peoples Republic of Bangladesh with the financial assistance of Danish International Development Agency (DANIDA). The technical support provided by Prof. Dr. Md. Samsul Alam of the Department of Fisheries Biology and Genetics at Bangladesh Agricultural University, Mymensingh, is acknowledged with appreciation.

Correspondence to:
M.R. MOLLA
Scientific Officer
Plant Genetic Resources Division
Bangladesh Agricultural Research Institute
Joydebpur, Gazipur-1701, Bangladesh
E-mail: rezwanbt@gmail.com
motiar_1@yahoo.com
Cellular phone: +8801716374405

References

[1]. Gimenes MA, Lopes CR, Valls JFM. Genetic relationships among Arachis species based on AFLP. Genet Mol Biol 2002; 25:349-353.

[2]. Guohao H, Meng R, Newman M, Gao G, Pitman R and Prakash CS. Microsatellites as DNA markers in cultivafte peanut (*Arachis hypogaea* L.). BMC Pl Biol 2003; 3 (3): 1-6.

[3]. Halward TM, Stalker HT, Larue EA, Kochert G. Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome 1991; 34:1013-1020.

[4]. He G, Prakash CS. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). Euphytica 1997; 97:143-149.

[5]. Herselman L. Genetic variation among Southern African cultivated peanut (*A. hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 2003; 133:319-327.

[6]. Hilu KW, Stalker HT. Genetic relationships between peanut and wild species of *Arachis* sect. *Arachis* (Fabaceae): evidence from RAPDs. Pl Syst Evol 1995; 198:167-178.

[7]. Kochert G, Stalker HT, Gimenes M, Galgaro L, Lopes CR, Moore K. RFLP and cytogenetic evidence on the origin and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 1996; 83:1282-1291.

[8]. Lacks. G, Stalker H. Isozyme analyses of Arachis species and interspecific hybrids. Peanut Sci. 1993; 20: 76-81.

[9]. Lanham, P, Forster B, Mcnicol P, Moss J, Powell W. Seed storage protein variation in *Arachis* species. Genome, 1994; 37(3):487-496.

[10]. Mailer RJ, Scarth R, Fristensky B. Discrimination among cultivars of rapeseed (*Brassica napus* L.) using DNA polymorphism amplified from arbitrary primers. Theor App. Genet, 1994; 87: 697-704.

[11]. Marshall P, Marchand MC, Lisieczko Z, Landry BS. A simple method to estimate the percentage of hybridity in canola (*Brassica napus*) F1 hybrids. Theor. Appl. Genet, 1994; 89: 853-858.

[12]. Nash JHE. DNAfrag, Version 3.03. Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario, Canada, 1991.

[13]. Nei M. Genetic distance between populations. Am Nat, 1972; 106: 283-292.

[14]. Saghai-Maroof MA, Soliman KM, Jonsensan RA, Allard RW. Ribosomal spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc. Natl Acad Sci USA, 1984; 81: 804-8019.

[15]. Smith, JSC, Register JC. Genetic purity and testing technologies for seed quality: a company perspective. Seed Sci Res, 1998; 8: 285–293.

[16]. Song QJ, Quiley CV, Nelson RL, Carter TE, Boema HR, Strachan JL, Cregan PB. A selected set of trinucleotide simple sequence repeat markers for soybean cultivar identification. Plant varieties and seeds, 1999; 12:207-220.

[17]. Subramanian V, Gurtu S, Nageswara Rao RC, Nigam SN. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. Genome, 2000; 43:656-660.

[18]. Yeh FC, Yang RC, Boyle T. 1999. POPGENE VERSION 1.31: Microsoft Window-based free Software for Population Genetic Analysis. ftp://ftp.microsoft.com/Softlib/HPGL.EXE

10/20/2010