DOI: 10.5567/GENETICS-IK.2013.1.6

Assessment of Microsatellite Markers: An Attempt Towards Documenting Genetic Diversity of Ethiopian Coffee

K.H. Sene and B. Admassu
Ethiopian Institute of Agricultural Research, Holetta Agricultural Research Center, Agricultural Biotechnology Laboratory, Holetta, P.O. Box 2003 code 31 Holetta, Ethiopia

ABSTRACT

Background: Identification of unique and diverse coffee genotypes and documenting the entire genetic structure of coffee is an overriding task to devise breeding and conservation strategy. **Aim of the study:** Two different microsatellites coded with their known NCBI locus tag as CFGA465 and AJ308774 were selected to investigate the role of marker screening in molecular characterization. **Methods:** The microsatellite CFGA465 distinguished identical and variant genotypes while all genotypes of *Coffea arabica* were similar based on the molecular data obtained from the microsatellite AJ308774. **Results:** The result of the present study revealed that CFGA465 is a polymorphic marker and AJ308774 is a monomorphic marker. The fact that the genetic information obtained from CFGA465 is not in accordance with the genetic information obtained from AJ308774 gave insight in the need to test several microsatellites to confirm the diversity of the various genotypes of *C. arabica*. It seems that CFGA465 is informative simple sequence repeat for it showed the highest polymorphism and aided in determining the genetic distance of the coffee accessions under study. The rest microsatellites, namely, AJ250255, AJ308755 and CFGA502 are included in this study to show diversity of alleles in *C. arabica*. **Conclusion:** The efficiency of 6% polyacryl amide gel in allele separation is quite evident in case of all microsatellites under study. The need of developing successful methodologies for implementing PCR based isolation of simple sequence repeat of coffee and depicting the relative importance of the simple sequence repeats in determining polymorphism is underlined. CFGA465 is a promising marker towards identification of various coffee genotypes.

Key words: Allele, breeding, *C. arabica*, conservation, microsatellite, monomorphic polymorphic

Insight Genetics 1 (1): 1-6, 2013

INTRODUCTION

Ethiopia is the origin and centre of diversity for *C. arabica*, which is known to occur in forest, semi-forest, garden and plantation. *C. arabica* is the only allotetraploid species out of the entire diploid species in the genus *Coffea* for it is a combination of the genome of its ancestors *C.canephora* and *C. eugenioides* (http://www.coffeegenome.org). *C. arabica* and *C.canephora* are two coffee cultivars[1]. *C. arabica*, a member of the big plant family Rubiaceae, is comprised of roughly 100 taxa[2]. There existed coffee specialities and varieties of both cultivar and wild type. Also such varieties of coffee like Ghawe, Aba buna and MCH2 are hybrid coffee acquired through conventional breeding[3]. Gene flow exists within *Coffea* sp. For instance, there exists gene flow from diploid coffee species to the tetraploid *C. arabica*. A species which resembles *C. eugenoids* is speculated to be ancestor of coffee[4].

Corresponding Author: K.H. Sene, Ethiopian Institute of Agricultural Research, Holetta Agricultural Research Center, Agricultural Biotechnology Laboratory, Holetta, P.O. Box 2003 code 31 Holetta, Ethiopia

Arabica coffee seeds can't be maintained in genbank. Ex-situ conservation strategy is the sole means of conserving the coffee genetic resource besides protecting the coffee genetic resource at in situ level via protecting the forest which harboured the coffee. The genetic diversity in gene bank must be assessed. Duplicate accessions must be screened to economize resource and aid gene bank managers to plan effectively[5]. Added to this, the genotypes in breeding sites should be explored. Molecular characterization of coffee via the aid of molecular markers is of paramount importance to document the diversity of the coffee genetic resource and formulate conservation plan, management plan and breeding plan. The genetic variations within accessions of *C. arabica* should be studied using simple sequence repeat, SSR, for they are multiallelic, polymorphic, co-dominant, transferable/reproducible and informative. The need of the present study stems from the fact that several genotypes must be characterized with sufficient number of markers. This research will have a merit for devising conservation and breeding strategy. Also it will enhance our understanding about diversity of satellite DNA in *C. arabica* and their difference in extent of polymorphism across different genotypes.

© 2013 Insight Knowledge, UK

MATERIALS AND METHODS

Coffee leaf samples were collected from an Agricultural Research Center, namely, Jimma Agricultural Research Center (JARC) of Ethiopia. To transport the leaf from JARC, located in a remote area of like 400 km from the biotech lab in which the present research is conducted, a plastic tube containing silica gel was used for most of the leaf sample and a plastic tube containing CTAB buffer was used for the few leaf samples. The DNA was extracted directly from leaf dried in silical gel. Also the DNA was extracted from the leaf in CTAB buffer which was preserved overnight in deep freez (-70°C) to make the leaf dried for ease grinding. All the leaf samples were crushed in mortar and pestel except two samples which were crushed using the genogrinder. The DNA was extracted from the leaf powder following the Doyle and Doyle CTAB protocol. The concentration of the DNA was quantified in a nano drop, spectrophotometer. The extracted DNA was analyzed in horizontal agarose gel electrophoresis. Primers were screened from published articles and ordered from oligonucleotide synthesizing company, microsynth, Switzerland. Total 100 μM stock and 10 μM working primer solutions were prepared in the lab. List of primers is shown in Table 1. About 2 μL DNA, 1 μL of 10 μM forward primer, 1 μL of 10 μM reverse primer, 0.5 μL Taq DNA polymerase, 0.5 μL dNTP, 2.5 μL 10× buffer, 1.5 μL $MgCl_2$ and 16 μL nuclease free water were mixed to prepare the PCR reaction in a final volume of 25 μL. Two different PCR programs were used. The touch down PCR program 94°C 4 min, 9 cycle of 94°C 45 sec, 1 min 60 to 55°C, decreasing by 0.5°C and 1 min 30 sec at 72°C, 72°C 1 min 30 sec and 72°C and 5 min was used following[6] to amplify AJ308755 and AJ308774. 94°C 5 min, 30 cycle of 94°C 1 min, 55°C 2 min and 72°C 2 min and 72°C 5 min is the other PCR program applied for amplification of CFGA502 and CFGA465[7]. The PCR program to amplify AJ250255 was 94°C 2 min, 5 cycles of 94°C 45 sec, 60°C (decreasing by one degree per cycle) 1 min and 72°C 90 sec followed by 30 cycles of 90°C 45 sec, 55°C 1min and 72°C 90 sec and 72°C 8 min[8]. The success of the PCR product was checked in 3% agarose for all of the microsatellites. Here the agarose data for CFGA465 is shown in this paper for it had value for polymorphism study. 1×TAE buffer was used for running horizontal agarose gel electrophoresis. The DNA image was analyzed in a UVP transilluminator. The allele separation was studied using 6% Polyacryl Amide Gel Electrophoresis (PAGE), vertical electrophoresis for all of the microsatellites. Acrylamide:bisacrylamide, distilled water, 10×TBE buffer, 10% APS and TEMED used for preparing the polyacryl amide gel solution, which was poured between the glass plates. Following polymerization of polyacryl amide solution occurring in between the glass plates, the PCR product was loaded in the polyacryl amide gel wells and migrated vertically in a vertical electrophoresis for 2:30 to 3 h in a running buffer of 1X TBE buffer to separate alleles. Following disassembly of the glass plates, the polyarylamide gel was left in the lower plate and stained for 15-20 min in a staining box, which contained distilled water and ethidium bromide. UVP transilluminator was used to visualize and analyze the image of the microsatellite in the stained gel. The application of DNA PAGE and staining with ethidium bromide was universally applied in several crops and other organisms. Such articles as[9,10,11,12] are typical examples, which applied DNA PAGE and staining with ethidium bromide. One of the microsatellite CFGA465 was analyzed in 3% agarose since it exhibited polymorphism in agarose. Among others, the role of 3% agarose in polymorphism study was applied by such authors as[13,14]. Among others, polymorphism was studied based on pattern of band, size of microsatellite allele and number of alleles.

Table 1. List of primers

SSR locus tag	Repeat	No. of repeats	Primer sequence (5' to 3')	Source
CFGA502	AG	27	F:AGCCACCCAGAAAACAG CACATC R:ATTTGCTTCTC AT GTTC CCTTTCA	7
CFGA465	AG	18	F: ACCCTTTACTACTTATT TACTCTC R: ACATCCCCT TGCCA TTTCTTC	7
AJ308755	CA	20	F: CCCTCCCTCTTTCT CCTCTC R: TCTGGGTT TT CTGTG TTCTCG	6
AJ308774	CT,CA	5,7	F: GCCACAAGTTTCGT GCTTTT R: GGGTGTCGG TGTAGG TGTATG	6
AJ250255	$(GT)_5C$ $(GT)_2$ $(GT)_{12}$		F: CCCTCCCT GCCAG AAGAAGC R: AACCACCGTC CTTTT CCTCG	8

RESULTS AND DISCUSSION

The PCR results showing the amplified microsatellites are displayed in Fig. 1-4 displays the template genomic DNA of 20 genotypes just to show the extracted DNA in 0.8% agarose. Figure 1a-c shows 3% agarose gel picture for amplicon CFGA465 while Fig. 1b-d shows the 6% polyacrylamide gel picture of alleles of CFGA465. Out of the successfully amplified

© 2013 Insight Knowledge, UK

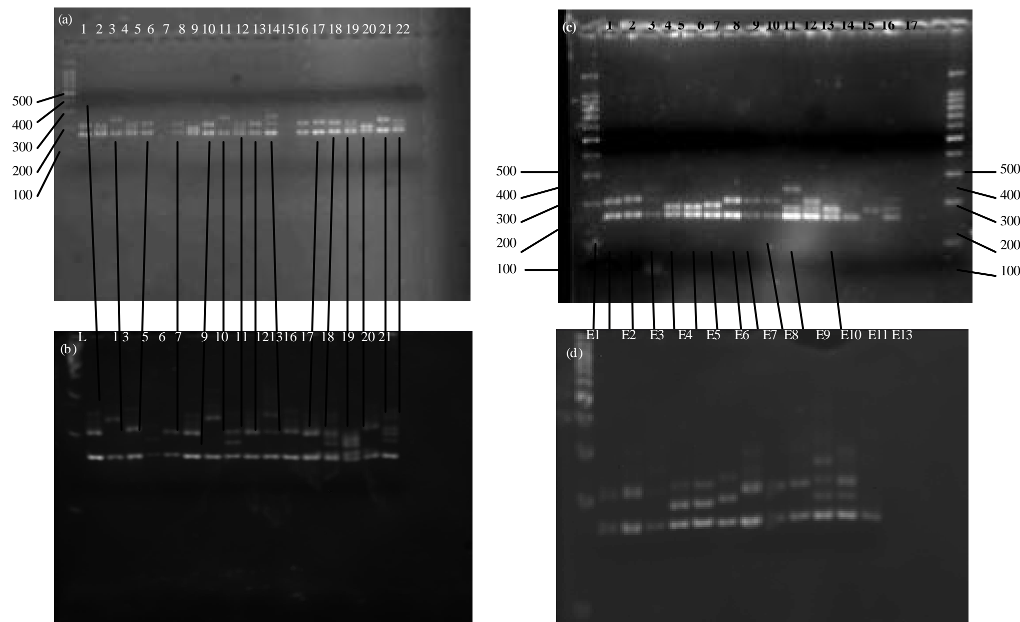

Fig. 1(a-d): (a) Characterization of 19 genotypes using microsatellite CFGA465 in 3% agarose. One lane represents one accession (b) Characterization of 16 genotypes out of the 19 genotypes displayed in "A" using 6% polyacryl amide gel. The labelling follows the same sample and number in 3% agarose and 6% polyacryl amide gel. Samples which are not loaded in 6% polyacryl amide gel are skipped in the label of 6% polyacryl amide gel. One lane represents one accession (c) Characterization of 15 genotypes using microsatellite CFGA465 in 3% agarose. One lane represents one accession and (d) Characterization of 12 genotypes out of the 14 genotypes displayed in "C" using 6% polyacryl amide gel. The labelling follows the same sample and number in 3% agarose and 6% polyacryl amide gel. Samples which are not loaded in 6% polyacryl amide gel are skipped in the label of 6% polyacryl amide gel. One lane represents one accession

CFGA465 from 19 genotypes, seven different clusters are can be identified in 3% agarose (Fig 1a). These different genotypes are eight genotypes in lane 1,5,7,9,12,15,16 and 17, four genotypes in lane 2,4,11 and18, two genotypes in lane 3 and 10, two genotypes in lane 8 and 19 and three additional unique genotypes in lane 13, lane 20 and lane 21. These polymorphic genotypes were confirmed with 6% polyacryl amide gel as displayed in Fig 1b. And the results from 3% agarose and 6% polyacryl amide are in parallel. Out of the 15 genotypes, CFGA465 showed polymorphism in 11 genotypes in 3% agarose (Fig. 1c). These different categories of genotypes are genotypes in lane 1, 2 and 7, genotypes in lane 4, 5 and 12, genotypes in lane 8 and9 and other such unique genotypes as lane 3, lane 6, lane 7, lane 10, lane 11, lane 13, lane 14 and lane 15. This result is also confirmed with 6% polyacryl amide gel (Fig. 1d). The alleles of the monomorphic microsatellite AJ308774 are shown in Fig. 2. The rest microsatellites namely, AJ250255, AJ308755 and CFGA502 are shown in Fig. 3. These microsatellites also indicate the existence of diverse SSR in coffee, which paved a new era in genotyping. All in all, the published microsatellite and the present amplicon microsatellite are compatible size wise. The previous reports of CFGA502 showed that the tetraploid *C. arabica* consists of 5 and 6 alleles for the cultivated ones and the wild, respectively[7]. Here only four alleles (2 alleles clearly separated and 2 alleles compacted together) for CFGA502 are counted (Fig. 3c). Perhaps polymorphism might have attributed to this allele number difference by one allele or two alleles. Obviously, the existence of greater than two alleles for all microsatellites indicates heterozygosity. Four alleles indicate hundred percent heterozygosity for the tetraploid coffee. Three alleles indicate, heterozygosity at one loci and homozygosity at the other loci. Five alleles counted for AJ250255 (Fig. 3a). AJ308774 is a monomorphic marker (Fig. 2). No genotypes except two genotypes are different according to the monomorphic

© 2013 Insight Knowledge, UK

Fig. 2(a-b): Microsatellite alleles of AJ308774 in 6% DNA polyacryl amide gel (DNA PAGE), (a) First lane-ladder. Lane 1-18 represents the various genotypes of *C. arabica* except one and (b) First and last lane-ladder. Lane 1-15 represents the various genotypes of *C.arabica* except one

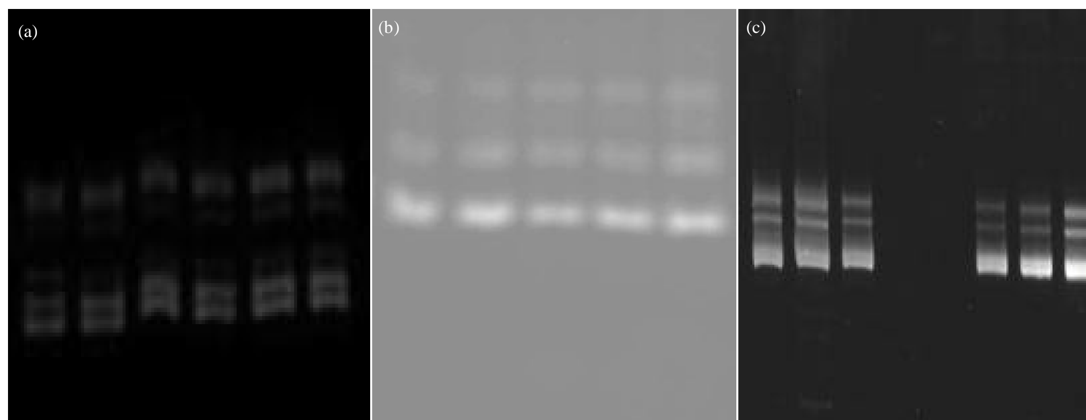

Fig. 3(a-c): Result from DNA PAGE. Microsatellite alleles of *C. arabica* in 6% polyacrylamide gel (DNA PAGE). (a) Microsatellite AJ250255 (five alleles containing microsatellite) across six genotypes, (b) AJ308755 (3 alleles containing microsatellite) across five genotypes and (c) CFGA502 (4 allele = 2+2 alleles compacted together) across six genotypes

marker, AJ308774. Two accessions of *Coffea robusta* were used in this study. And one of the unique genotype in lane 13 (Fig. 2a) is a microsatellite from *C. robusta.* This band is completely unique relative to other bands. It seems that this marker is clearly applicable in species identification like identifying *C. arabica* from *C. canephora.* However, it has no application in cultivar identification of *C. arabica* since it is a monomorphic marker for *C. arabica* as all genotypes displayed in Fig. 2 are similar. The allele number and size of AJ308774 remained the same in all genotypes of *C. arabica.* All the gentypes of *C. arabica* contained three alleles of AJ308774. AJ308774 was published to contain 2 allele only[6]. The present result indicates that there are three alleles of AJ308774. One additional allele might have acquired during gene introgression/crossing over which occured during breeding. Five alleles counted in AJ250255 and this deviates from the expected maximum four alleles of *C. arabica.* Extra one allele might have occurred through events of tandem repeat duplication or experiments of breeding.

© 2013 Insight Knowledge, UK

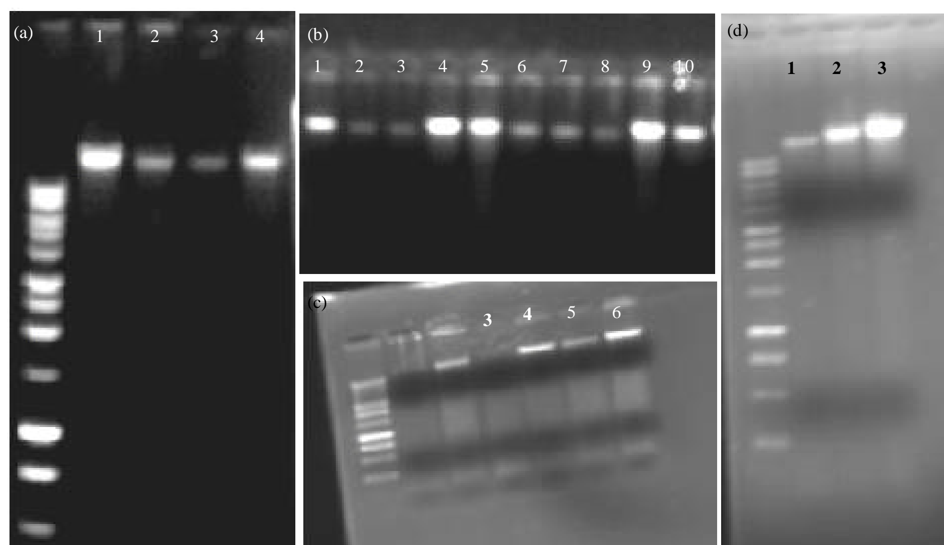

Fig. 4(a-d): Extracted genomic DNA for some of the samples, 20 coffee genotypes, used for SSR Study, (a) Extracted DNA for four samples, (b) Extracted DNA for 10 samples. (c) Extracted DNA samples for 6 samples, two failed(1 and 3) and (d) Extracted DNA for samples failed in fig. c, that is two samples for 1 and 3

CONCLUSION AND RECOMMENDATIONS

The use of the right simple sequence repeat out of candidate satellite DNA seems to be the bottom line to characterize coffee genotypes. Similar genotypes are identified via the aid of some of the microsatellites. A typical example is the identical genotypes identified with AJ308774. But diverse genotypes were identified using CFGA465. The task of marker screening seems a primary step before making decision about duplicate and unique genotypes. The knowledge of duplicates is critical. Attention must be paid for conservation not to waste resource by conserving duplicate genotypes. The tendency is to avoid redundant genotypes in conservation areas. This preliminary test indicates that screening of microsatellite is quite substantial before taking a conservation action. Still, the present study is based on two microsatellites and checking several microsatellites across quite many genotypes is recommendable to confirm similar and different genotypes. Breeding sites should focus on unique genotypes and all of these may not be relevant for breeders. And conservationist must conserve one of the genotypes among duplicates to manage resources and to have representative genotypes. The fact that CFGA465 exhibited polymorphism across the genotypes under study is indicative of the use of this microsatellite for identification of coffee cultivars. Perhaps, due to its polymorphism, CFGA465 in conjunction with other microsatellites may be sufficient to characterize coffee genotypes for CFGA465 appears to be polymorphic marker. Sequencing this microsatellite is highly recommendable. It is of considerable importance to compare the result from 3% agarose and 6% polyacryl amide in case of CFGA465. Microsatellite allele may separate in 3% agarose as detected in the scenario of CFGA465. The scenario of CFGA465 showed that the 3% agarose data is very applicable for genotype identification. But, 3% agarose data was found to be impractical and of no use for allele separation for the other microsatellites. The resolution of 6% polyacryl mide gel is crystal clear. This allows counting more number of alleles. A good example of allele number and size difference is detected in lane 4, 5 and 6 of Fig 1c-d.

Two alleles counted for genotypes in lane 4, 5 and 6 of Fig. 1c, 3% agarose while these same genotypes contained three alleles in 6% polyacryl amide gel, Fig 1d, lane 4,5 and 6. In Fig 1b, 6% polyacryl amide gel, the lower alleles with the smallest size of all the alleles showed variant size in some of the genotypes. This variant size which could be due to single or two base pair difference was not manifested in 3% agarose, Fig. 1a. In a nutshell, microsatellite data from 6% polyacryl amide gel is the most accurate, informative and helpful for the study of polymorphism of genotypes. The polyacrylamide gel result for AJ250255, AJ308755 and CFGA502 indicates that 6% polyacryl amdie gel is the most applicable technique to use these microsatellites in cultivar identification. This case study seems a cornerstone towards the need of confirmation and investigation of difference and similarity of coffee genotypes using significant number of microsatellites.

© 2013 Insight Knowledge, UK

ACKNOWLEDGMENTS

The molecular practicum which demanded ample of molecular experiments wouldn't have been realized without the existence of sophisticated molecular equipments and ideal molecular laboratory facility of the Ethiopian Institute of Agricultural Research (EIAR), the Agricultural Biotechnology Laboratory, Molecular Biology Research Laboratory. Financial cost for consumables like DNA extraction kit, PCR reagents, reagents for electrophoresis and other molecular reagents and chemicals was covered by EIAR through the regular budget of the government. Jimma Agricultural Research Center (JARC) of Ethiopia is appreciated for offering coffee leaf samples. Computer with internet facility and field vehicles were arranged by EIAR. The authors thank both the host institute, EIAR and JARC.

REFERENCES

1. Davis, A.P., R. Govaerts, D.M. Bridson and P. Stoffelen, 2006. An annotated taxonomic conspectus of the genus *Coffea* (Rubiaceae). Bot. J. Linnean Soc., 152: 465-512.
2. Cros, J., P. Lashermes, P. Marmey, F. Anthony, S. Hamon and A. Charrier, 1993. Molecular analysis of genetic diversity and phylogenetic relationships in *Coffea*. Proceedings of the 14th Colloquium of International Coffee Science Association (ASIC), July 14-19, 1991, San Francisco, USA.
3. Ramos, R., T. Wondyifraw, F. Martinez, M.E. Gonzalez, G. Endale, T. Alemayehu and A. Zerihun, 2009. Plant regeneration through somatic embryogenesis in three Ethiopian *Coffea arabica* Lin. hybrids. Biotecnologia Vegetal, 9: 19-26.
4. Herrera, C.J., M.C. Combes, H. Cortina and P. Lashermes, 2004. Factors influencing gene introgression into the allotetraploid *Coffea arabica* L. from its diploid relatives. Genome, 47: 1053-1060.
5. Karp, A., S. Kresovich, K.V. Bhat, W.G. Ayand and T. Hodgkin, 1997. Molecular tools in plant genetics resources conservation: A guide to the technologies. IPGRI Tech. Bull. No. 2. http://pdf.usaid.gov /pdf_docs/PNACB166.pdf
6. Cubry, P., P. Musoli, H. Legnate, D. Pot, F. de Bellis, V. Poncet, F. Anthony, M. Dufour and T. Leroy, 2008. Diversity in coffee assessed with SSR markers: Structure of the genus Coffea and perspectives for breeding. Genome., 51: 50-63.
7. Moncada, P. and S. McCouch, 2004. Simple sequence repeat diversity in diploid and tetraploid *Coffea* species. Genome, 47: 501-509.
8. Combes, M.C., S. Andrzejewski, F. Anthony, B. Bertrand, P. Rovelli, G. Graziosis and P. Lashermes, 2000. Characterization of microsatellite loci in *Coffea arabica* and related coffee species. Mol. Ecol., 9: 1178-1180.
9. Adawy, S.A., 2007. An evaluation of the utility of Simple Sequence Repeat loci (SSR), Expressed Sequence Tags (ESTs) and expressed sequence tag microsatellites (EST-SSR) as molecular markers in cotton. J. Applied Sci. Res., 3: 1581-1588.
10. Anas and T. Yoshida, 2004. Genetic diversity among Japanese cultivated sorghum assessed with simple sequence repeat markers. Plant Prod. Sci., 7: 217-223.
11. Ghosh, A.K. and L.N. Lukens, 2006. European and Asian pears: Simple sequence repeat-polyacrylamide gel electrophoresis-based analysis of commercially important North American cultivars. HortScience, 41: 304-309.
12. Nghia, P.T. and T.T.C. Hoa, 2007. Golden rice clustering by microsatellites for further breeding approach. Omonrice, 15: 54-62.
13. Darvishzadeh, R., M. Azizi, H. Hatami-Maleki, I. Bernousi, A.B. Mandoulakani, M. Jafari and A. Sarrafi, 2010. Molecular characterization and similarity relationships among sunflower (*Helianthus annuus* L.) inbred lines using some mapped simple sequence repeats. Afr. J. Biotechnol., 9: 7280-7288.
14. Kanagaraj, P., K.S.J. Prince, J.A. Sheeba, K.R. Biji, S.B. Paul, A. Senthil and R.C. Babu, 2010. Microsatellite markers linked to drought resistance in rice (*Oryza sativa* L.). Curr. Sci., 98: 836-839.

© 2013 Insight Knowledge, UK