Advances in Biological Research 4 (3): 185-192, 2010
ISSN 1992-0067

Malaysian Medicinal, Aromatic and Ornamental Plants Database Supplemented with DNA Markers

[1]Nurul A.H., [1,3]Subhash J. Bhore and [1,2]Farida H. Shah

[1]Molecular Biology Division, Melaka Institute of Biotechnology, Lot 7,
Melaka International Trade Center City, Ayer Keroh, Melaka, Malaysia
[2]Novel Plants Sdn. Bhd., 27C Jln Petaling Utama 12, 7.5miles Old Klang Road, 46000 Petaling Jaya, Malaysia
[3]Department of Biotechnology, Faculty of Applied Sciences, AIMST University,
Bedong-Semeling Road, Bedong, 08100, Kedah, Malaysia

Abstract: Plant biodiversity richness and endemism in Malaysia is of global importance. Malaysia is known to have more than 15,000 species of higher plants. Currently, there is no database on medicinal, aromatic and ornamental plants (MAOPs) supplemented with DNA markers. The aim of this research project is to develop a fully functional MAOPs database supplemented with DNA markers, which could serve as a tool in identification of MAOPs. In total, 688 plants (or samples) were collected from different parts of Peninsular Malaysia and RAPD markers (or profiles) has been developed for all 688 plants accessions. The database contains cascade to put in information about plants local name, scientific name, morphological traits, habitat location, isolated DNA concentrations, PCR primers used, primer sequences, gel images, schematic diagrams, informative DNA markers, marker DNA bands size (bp), DNA marker sequences, botanical characteristics and bioactive compounds. Work is being carried out to populate the database. This MAOPs database supplemented with DNA markers will be useful not only in identification and cataloging of MAOPs but also in authentication of herbal raw material for quality control. As of July 14, 2010, this database contains entries for 688 accessions, which represents 347 genus and 395 species. The current status of the MAOPs database supplemented with DNA markers is being reported in this paper.

Key words: Aromatic plants · DNA markers · Medicinal plants · Ornamental plants · Plant database · RAMS · RAPD

INTRODUCTION

Biodiversity refers to biological variety and it is divided into three levels which are called ecosystem diversity, species diversity and genetic diversity. Malaysia is recognized as one of the 12 mega-diversity countries out of the 25 identified biodiversity hotspots in the world. Malaysia is also one of the global biodiversity hotspots [1]. National Policy of Malaysia on Biological Diversity states that biodiversity has particular significance for nation and peoples economic benefits, food security; environmental stability, national biological heritage, scientific, educational, recreational values and biosafety [2]. By realizing a huge importance and potential of plant resources this research project was initiated to establish the database for medicinal, aromatic and ornamental plants (MAOPs).

With the current advances in deoxyribonucleic acids (DNA) technology and their applications as precise identification tools, it is highly recommended to utilize DNA marker based methods for unambiguous germplasms and cryptic plants species identification [3-6]. The identification of plants genotype and its genetic variability is important in plant taxonomic study and in conservation of rare or threatened plant species in forests. Moreover, MAOPs genetic resources are also the main source of raw materials in various types of herbal preparations. Authentication of raw material is a crucial step before it can be used in manufacturing process. Traditionally, medicinal plants are being collected from

Corresponding Author: Farida H. Shah, Molecular Biology Division, Melaka Institute of Biotechnology, Lot 7,
Melaka International Trade Center City, Ayer Keroh, Melaka, Malaysia,
E-mail: shahf2@yahoo.com / subhashbhore@gmail.com

their wild habitats based on only their morphological traits and authentication to prevent adulteration or to get the right variety are not in practice. Thus, accurate verification of taxa/variety is important to prevent adulteration, to confirm raw material and to protect intellectual property rights (IPR). The DNA markers are very useful in genetic fingerprinting and precise plant identification [3-4, 6]. Random Amplification of Polymorphic DNA (RAPD) and Random Amplified Microsatellites (RAMS) techniques are economical and frequently used in genetic mapping and DNA fingerprinting [6-8].

RAPD is a technique that utilizes short synthetic oligonucleotides (8-12 mer) of random sequences as primers to amplify nanogram amounts of DNA fragments from total nuclear or sub-cellular genomic DNA by PCR [5-7]. RAPD produces DNA profiles of varying complexity, depending on the primer and template used [9-10]. Each amplification product is expected to result from the existence of two annealing sites in inverted orientations, 3' ends facing each other within amplifiable distance. DNA polymorphisms could be caused by differences in nucleotide sequences or by structural arrangements [10]. Other approach (RAMS) of using microsatellites primers in PCR is also used. RAMS technique involves PCR amplification with oligonucleotide primers that are complementary to specific simple sequence repeats (SSRs) [11-13]. Synthetic oligonucleotides, each representing a specific SSR, are used as single PCR primers in RAMS [8]. If two inversely oriented microsatellites are present within an amplifiable distance from each other, the inter-repeat sequence is amplified. The amplified products generate distinct banding patterns that can be separated on low-resolution agarose gels using EtBr staining [14-15].

The use of molecular markers in herbal research are increasingly used for screening of germplasms to study genetic diversity, identify redundancies in the collections, test accessions stability and integrity and to resolve taxonomic relationship. Molecular (DNA) markers are also used in studying genetic resources to correlate their genetic makeup and beneficial active compounds produced by genotype. There are number of DNA marker applications in agricultural biotechnology and herbal industry. A well established MAOPs database with DNA markers will greatly assist agricultural sector and herbal industry. The research work is in progress to populate the database and output of this research project is expected to help in establishing a functional and reliable MAOPs database supplemented with DNA markers. This paper reports the current status of the database.

MATERIALS AND METHODS

Plant Materials: Plants or plant materials for targeted plants were collected from the field or from their natural habitats in Peninsular Malaysia. The collected plants are being maintained in our medicinal plants collection as a reference. Morphological traits of collected plants were recorded to document their characteristics. The collected fresh leaves were washed systematically and stored at -80°C for genomic DNA extraction.

Genomic DNA Extraction: Total genomic DNA was isolated from fresh leaves using manual method, a method described by Sambrook *et al.* [16] with some minor modifications. Young leaves were washed with plenty of running tap water and then in 70 % ethanol for 5 min followed by washing with sterile deionized water for 2 min to avoid surface contamination. The samples were ground separately into powder with the help of liquid nitrogen (LN) using mortar-pestle. Two (2) gram of the leaves tissue powder was added to 15 ml extraction buffer (0.1 M Tris HCl, 0.05 M EDTA, 0.5 M NaCl, 1% PVP, 1.4% SDS and 10 mM 2-mercaptoethanol). The suspension was incubated in a water bath at 37°C for 1 h. Potassium acetate was added to each tube and again incubated on ice for 1 h. The phases were separated by centrifugation and isopropanol was added to the new suspension. The mixtures were then allowed to precipitate for 1 h at -20°C. After centrifugation, pellets were dried and dissolved in TES buffer. RNase (10 mg/ml) was added to each tube and incubated for 1h at 37°C. DNA was purified by using phenol/chloroform/isoamyl-alcohol extraction method, followed by precipitation with isopropanol and sodium acetate. Finally, pellet was washed with 70 % ethanol twice. Then, the pellet was dissolved in 1x TE buffer. The quantitative estimation was carried out using spectrophotometer. DNA samples were stored at -80°C till its use in RAPD reactions.

RAPD Analysis: In a preliminary testing, 100 random primers from Kit A, B, C, D and E (Proligo LLC) were used in screening. All primers were ten bases long with 60-70 % G+C content. From the screening tests, selected primers which produced reproducible band fragments were taken as markers. Each selected primer was further tested twice or thrice to verify the consistency of the observed banding patterns. Primers which produced consistent patterns were used in subsequent RAPD analysis.

After optimization of PCR conditions, RAPD-PCR was performed in a total volume of 25 µl mix containing 25 ng template nuclear genomic DNA, 1x PCR buffer, 4 mM

$MgCl_2$, 0.4 mM dNTP mix, 1 U of *Taq* polymerase (Promega) and 0.6 μM primer (Proligo). Amplifications were performed in a programmable thermal cycler (MJ Research). RAPD-PCR cycles were as follows; 40 cycles of 95°C for 30 sec, 38.5°C for 1 min and 72°C for 1 min. A final step of extension was carried out at 72°C for 10 min. Out of 25 μl, 15 μl amplified products from each completed RAPD-PCR were separated on 1% agarose gel in 1x TAE buffer [5, 7]. The gels were stained with ethidium bromide and photographed under UV light using Gel Documentation System. For reference, 1 kb DNA ladder (Promega) marker was used as a molecular standard. PCR results representing bands were tested for reproducibility and consistency at least three times.

RAMS Analysis: For RAMS analysis, PCR amplification were tested selectively using 11 microsatellites primers; $(GGAT)_4$, $(GACA)_4$, $(CA)_8$, $(GT)_8$, $(GAA)_6$, $(GTG)_5$, $(GAC)_5$, T3, T7, M13, and $(GATA)_4$. The reactions were performed in a total volume of 25 μl mix consisting 25 ng template (nuclear genomic DNA) 1x PCR buffer, 4 mM $MgCl_2$, 0.4 mM dNTP mix, 1 U of *Taq* polymerase and 0.9 μM primer. The amplification was done in a programmable thermal cycler. PCR cycles were as follows; 40 cycles of 94°C for 30 sec, vary from 43.5°C - 60.8°C for 1 min and 72°C for 20 sec. A final step of extension was carried out at 72°C for 5 min. Out of 25 μl, 10 μl of PCR products were analyzed by electrophoresis in 1.8% agarose gel electrophoresis in TAE buffer and stained with ethidium bromide.

Cloning and Sequencing of RAPD Fragments: RAPD band pattern of each accession was analyzed in order to select the most prominent and informative bands for cloning. Selected RAPD fragments were excised from the agarose gel and purified using Qiagen Gel Extraction-kit (Qiagen, Germany). Purified RAPD fragments were cloned in pGEM-T Easy PCR-cloning vector (Promega) [17]. Ligated products were used in transformation of *E. coli* DH5-α strain. Plasmid DNA was isolated using DNA purification Kit (Promega). In order to confirm the insert length, pDNA samples were digested with *Eco*RI restriction enzyme. After confirmation of insert length, bacterial cells were cultivated for respective recombinant clones and purification of the pDNAs was carried out. Sequencing of the RAPD fragments was carried out from both ends by using SP6 and T7 primers.

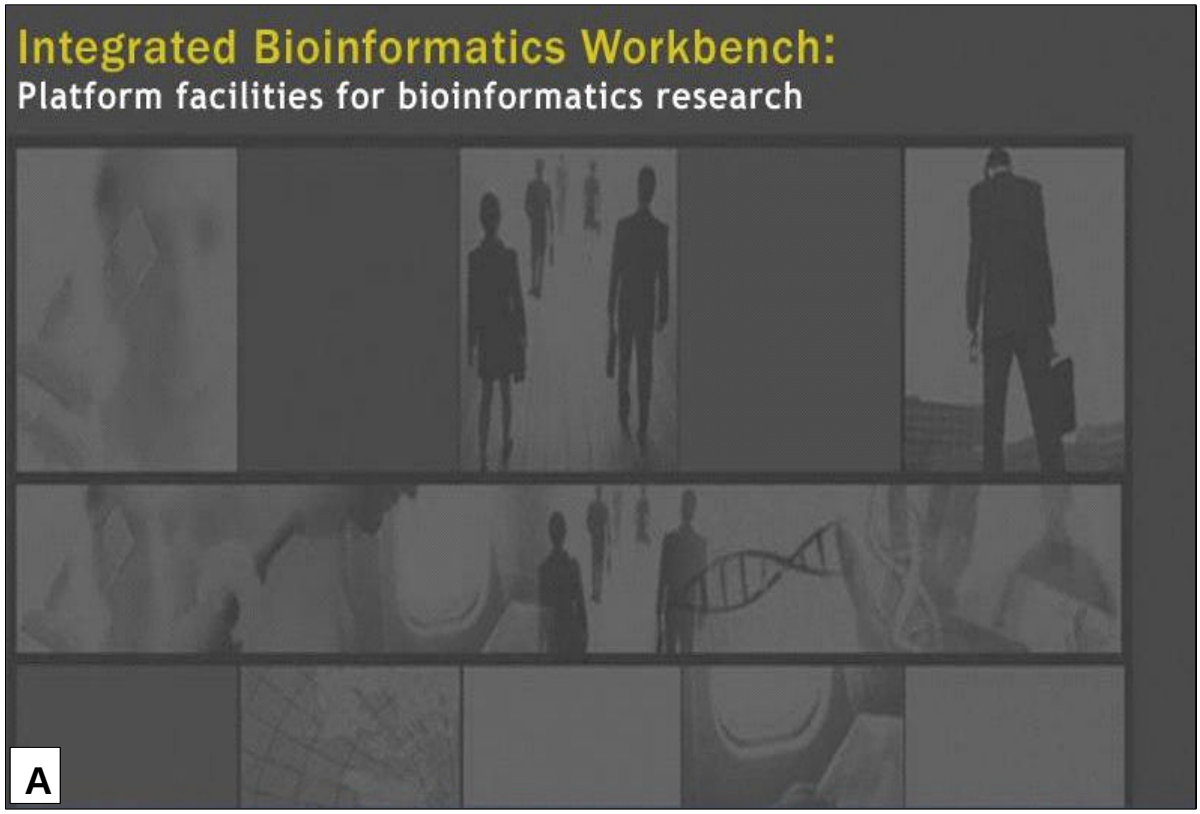
Data Collection and Data Analysis: From RAPD-PCR gel-electrophoresis images, RAPD profiles were analyzed. From DNA fingerprint profiles, PCR band was scored (1) for the presence or (0) for the absence of a band-pairwise distance (similarity matrices). It was computed based on *Jaccard's coefficient* of similarity, using NTSYS-pc software. Dendrograms were constructed using unweighted pair group method with arithmetic averages (UPGMA) cluster analysis using NTYSYS-pc v2.0 software [18]. The informative markers from generated RAPD bands were selected and marked by ID Gel Analysis software. These markers were used as DNA markers to differentiate plants species or varieties from other accessions. To identify the DNA markers, data was collected from sequenced RAPD fragments. Sequences of RAPD marker fragments were compared with DNA sequences of other organism available in NCBI database. Annotation of the DNA marker sequences was carried out using online softwares such as blastn and blastx at NCBI (http://www.ncbi.nlm.nih.gov/) and other online Bioinformatics tools available at JustBio (http://www.justbio.com/). The whole data including botanical source, morphological traits of accessions and other characteristics were added into MAOPs database.

RESULTS AND DISCUSSION

As of July 14, 2010, 688 plant accessions have been collected in total. Total nuclear genomic DNA is isolated from all accessions leaf sample tissues. The yield of the nuclear genomic DNA from 2 gram fresh leaf tissue was in range of 8 to 520 μg. The RAPD primer screening steps resulted in selection of 10-decamer primers which detected reasonable DNA polymorphism. RAPD-PCR amplification was carried out in triplicate for all 688 accessions to make sure all bands are reliable and consistent (data is not shown). Similarly, all 11 microsatellites primers; $(GGAT)_4$, $(GACA)_4$, $(CA)_8$, $(GT)_8$, $(GAA)_6$, $(GTG)_5$, $(GAC)_5$, T3, T7, M13 and $(GATA)_4$ were used to generate RAMS markers for the accessions in our collection. RAPD-based and microsatellites-based DNA marker profiles have been developed for 688 plants accessions in our collection. This 688 plant accessions comprises plant species from 347 genus and 395 species. The summary of the collection is depicted in Table 1.

Table 1: Total number of plant accessions, genus and species in medicinal, aromatic and ornamental plants (MAOPs) database

Grouping	Total Number
Accessions	688
Genus	347
Species	395
Medicinal plants	547
Aromatic plants	102
Ornamental plants	39



Delete	Edit	View Details	Species Name	Common Name	Category	No of Accession	No of Information Marker	Morphological Image
☐	Edit	View Details	Cymbopogun citratus Stapf.	Lemon Grass,Serai makan/ Serai betul	Medicinal,Aromatics	3	1	
☐	Edit	View Details	Oroxylum indicum L.	Bonglai kayu	Medicinal	1	0	
☐	Edit	View Details	Hydrangea macrophylla	Bigleaf hydrangea,Bunga 3 bulan,thunberg	Medicinal	1	0	
☐	Edit	View Details	Canna indica Linnaeus	Bunga tasbih	Medicinal	1	0	
☐	Edit	View Details	Lagerstroemia floribunda	Bungor,Crape myrtle,Crepe myrtle,Kedah bungor	Medicinal	1	0	
☐	Edit	View Details	Amaranthus spinosus	Bayam berduri,Pigweed,Prickly Amaranth,Spiny amaranth	Medicinal	1	0	

B

Fig. 1: Medicinal, aromatic and ornamental plants (MAOPs) database platform. (A) Integrated Bioinformatics workbench (homepage). (B) Window displaying the summary of accessions/entries in the MAOPs database. (C) Window displaying the individual plant accession entry details. (D) Window displaying the DNA marker sequence, gel image, sequenced band and its size in base pairs.

Authentic identification of taxa is necessary to ensure protection of intellectual property rights (IPR). The traditional method for identifying species by morphological characters is now being replaced by DNA profiling because of some limitations of morphological data [19]. Evidently, DNA marker technique is a rapid and sensitive, which can be used to estimate relationship between closely and more distantly related species and groups. The RAPD and RAMS markers generated in this study will be useful in authentic identification of respective taxa. The RAPD markers generated for some popular medicinal plants are deposited in GenBank/DDBJ/EMBL. For instance, twelve (12) RAPD sequence markers developed for two medicinally and commercially important medicinal plants namely, Mistletoe Fig (*Ficus deltoidea* Jack) and Kacip Fatimah (*Labisia pumila* Benth and Hook f) could be useful in different applications [5, 7].

The standardized PCR using single, arbitrarily small oligonucleotides, known as RAPD and microsatellites primers allows the random amplification of DNA sequences throughout the genome. The PCR-products of these amplifications are highly polymorphic and unique bands and band patterns can be used as genetic fingerprints to precisely identify the medicinal plants for their commercial exploitation as well as conservation. The MAOPs database supplemented with DNA markers is a collection of MAOPs images, medicinal properties of the plants, morphological traits of the plants along with the RAPD and RAMS markers for each individual plant in the database. The database is divided into several sections according to plant species and primers utilized in the PCR reaction. Database contains information about local name, scientific name, morphological variations, localities, DNA concentrations (μg/μl), primers used, primer sequences, gel images, schematic diagrams, informative markers, DNA marker bands size (bp), DNA sequences and botanical characteristics. From distinctive DNA band patterns, it provides important baseline data for conservation and collection strategies for respective plants. The study showed that RAPD and RAMS analysis is rapid, economical and a reliable method for precise identification of plants.

The MAOPs database is designed to allow the quick and easy retrieval of a maximum amount of information associated with each accession. This database will be made available on World wide web (www) for easy, efficient and effective way of sharing the database information. This database will help and encourage information transfer from highly protected species to more researchers. The DNA samples for all accessions in our collection are stored at -80°C with the equivalent vouchers stored in a herbarium. Information on stored genomic DNA stocks with associated data on the source plant species is also available in the database. DNA is very stable and can be stored for years [20]. DNA and DNA based markers from plants provides a wealth of information [21]. Researches can take advantage of available data on DNA markers and morphological traits to utilize it in various studies including phylogenetic, phylogeographic and population studies as well as gene discovery and marker development.

The MAOPs database developed using Bioinformatics platform (Figure 1), aims at development of unique database covering systematic laboratory information management system (LIMS), development of new Bioinformatics tools, analyzing, storing, retrieving, editing and organizing scientific data. This project also covers the data mining integration between the existing data and relevant pre-selected databases to identify and create commercially viable knowledgebase using high performance computing. Perhaps, this MAOPs database supplemented with DNA markers is the first of its kind.

There are some similar types of plant databases developed by researchers in other countries, but without DNA markers for each entry. One example to quote is of University of Connecticut (UConn) plant database of trees, shrubs and vines [22]. This database contains a good collection of photographic and textual information on trees, shrubs and vines but lack DNA marker information for individual plants in the database. Popular plant databases available online includes USDA plant database [23], the Dave's Garden plants database [24] and Tropical plant database [25]. All plant databases available online are user friendly though they represent data in different versions and forms. For instance, the USDA plant database provides standardized information about the vascular plants, mosses, liverworts, hornworts and lichens of the U.S. and its territories. We have developed MAOPs database supplemented with DNA markers, but it does not preclude the importance and applications of other databases available either online or offline.

DNA markers have vast applications in biotech industry. As we know that plant DNA barcoding is crucial in biodiversity research especially in biodiversity regions and countries like Malaysia and we are exploiting it for the benefit to community. In this plant barcoding and database development project, future work will focus on integration of information from chloroplast DNA in

addition to existing nuclear genomic DNA based barcode. The informative and or species-specific DNA markers will be used to distinguish the plant species. Developed DNA markers will also be used to catalogue the MAOPs in correlation with their secondary metabolites content.

CONCLUSIONS

This project is successful in establishing MAOPs database supplemented with DNA markers. The results from this study indicate that the RAPD and RAMS techniques are useful and reliable tools for identification of germplasms and genetic relationship between and within the plants species. RAPD and RAMS markers can be used as a marker ID for individual plant species. MAOPs database will be helpful in i) plant species conservation programs, ii) resolving biopiracy issues, iii) plant forensic investigations, iv) proper documentation of natural resources at national/regional level. The reasonable diversity observed at DNA level in this study may be exploited for various applications in agriculture sector and herbal industry. Nevertheless, this MAOPs database supplemented with DNA markers will be useful in selected plants information management.

ACKNOWLEDGEMENT

The authors are grateful to the Melaka State Government, Melaka, Malaysia for research funding (Project code; No: 3001-5).

REFERENCES

1. HIS, 1998. Humane society international and World wide fund for nature Australia, 1998. From words to action: a preliminary review of progress to implement the national strategy for the conservation of Australia's biological diversity, world wide fund for nature Australia, Sydney.
2. MNPBD, 1998. Malaysia's national policy on biological diversity, official declaration: Thursday, April 16 1998, Kuala Lumpur, Malaysia, Ministry of Science, Environment and Technology.
3. Kumar, L.S., 1999. DNA markers in plant improvement: An overview. Biotechnology Advances, 17: 143-182.
4. Craft, K.J., J.D. Owens and M.V. Ashley, 2007. Application of plant DNA markers in forensic botany: Genetic comparison of *Quercus* evidence leaves to crime scene trees using microsatellites. Forensic Science International, 165: 64-70.
5. Bhore, S.J., A.H. Nurul and F.H. Shah, 2009a. Genetic variability based on randomly amplified polymorphic DNA in Mistletoe Fig (*Ficus deltoidea* Jack) collected from peninsular Malaysia. J. Forest Sci., 25: 57-65.
6. Waldron, J., C.P. Peace, I.R. Searle, A. Furtado, N. Wade, I. Findlay, M.W. Graham and B.J. Carroll, 2002. Randomly amplified DNA Fingerprinting: A culmination of DNA marker technology based on arbitrary-primed PCR amplification. J. Biomed. Biotechnol., 2: 141-150.
7. Bhore, S.J., A.H. Nurul and F.H. Shah, 2009b. Genetic variability based on randomly amplified polymorphic DNA in Kacip Fatimah (*Labisia pumila* Benth and Hook f) collected from Melaka and Negeri Sembilan States of Malaysia. J. Forest Sci., 25: 93-100.
8. Hantula, J., M. Dusabenyagasani and R.C. Hamelin, 2007. Random amplified microsatellites (RAMS)-A novel method for characterizing genetic variation within fungi. Euro. J. Forest Patho., 26: 3159-166.
9. Welsh, J. and M. McClelland, 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res., 18: 7213-7218.
10. Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., 18: 6531-6535.
11. Meyer, W., T.G. Michell, E.Z. Freedman and R. Vilgalys, 1993. Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. J. Clin. Biol., 31: 2274-2280.
12. Perring, T.M., A.D. Cooper, R.J. Rodriguez, C.A. Farrar and T.S. Bellows, 1993. Identification of a white fly species by genomic and behavioural studies. Science, 259: 74-77.
13. Gupta, M., J. Chyi, J. Romero-Severson and J.L. Owen, 1994. Amplification of DNA markers from evolutionary diverse genomes using single primers of simple-sequence repeats. Theor. Appl. Genet., 89: 998-1006.
14. Sharma, P.C., B. Huttel, P. Winter, G. Kahl, R.C. Gardner and K. Weising, 1995. The potential of microsatellites for hybridizationand polymerase chain reaction-based DNA fingerprinting of chickpea (*Cicer arietinum L.*) and related species. Electrophoresis, 16: 1755-1761.

15. Weising, K., H. Nybom, K. Wolff and W. Meyer, 1995. DNA Fingerprinting in Plants and Fungi. CRC Press, Boca Raton, Florida.
16. Sambrook, J., E.F. Fritsch and T. Maniatis, 1989. *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, Cold Spring Harbor Laboratory Press, New York.
17. Zhou, M.Y., S.E. Clark and C.E. Gomez-Sanchez, 1995. Universal cloning method by TA strategy. Biotechniques, 19: 34-35.
18. Rohlf, F.J., 1998. NTSYS-pc, Numerical taxonomy and multivariate analysis system, Exeter Software, Applied Biostatics, New York.
19. Virk, P.S., B.V. Ford-Lloyd, M.T. Jackson and H.J. Newbury, 1995. Use of RAPD for the study of diversity within plant germplasm collections. Heredity, 74: 170-179.
20. Kogoma, T., 1997. Stable DNA replication: interplay between DNA replication, homologous recombination and transcription. Microbiol. Mol. Biol. Rev., 61: 212-238.
21. Strand, A.E., J. Leebens-Mack and B.G. Milligan, 2003. Nuclear DNA-based markers for plant evolutionary biology. Mol. Eco., 6: 113-118.
22. Brand, M.H., 2010. University of Connecticut (UConn) plant database of trees, shrubs and vines. Available online at http://www.hort. uconn. edu/ plants/, verified on June 23, 2010.
23. USDA, 2010. Plant Database: USDA, Natural resources conservation services. Available online at http://plants.usda.gov/, verified on June 23, 2010.
24. Dave's Garden, 2010. PlantFiles: The Dave's Garden plants database (an Internet Brands company), Available online at http://davesgarden. com/ guides/pf/, verified on June 23, 2010.
25. Raintree Nutrition Inc., 2010. Tropical Plant Database. Available online at http://www.rain-tree.com/plants. htm, verified on June 23, 2010.