

Full Length Research Paper

Estimation of genetic diversity of the Kenyan yam (*Dioscorea* spp.) using microsatellite markers

Z. K. Muthamia^{1*}, F. E. Morag, A. B. Nyende², E. G. Mamati² and B. W. Wanjala¹

Kenya Agricultural Research Institute (KARI), Jomo Kenyatta University of Agriculture and Technology (JKUAT), International Institute of Tropical Agriculture (IITA), Nairobi, Kenya.

Accepted 16 September, 2013

Yam landraces in Kenya have not been fully characterized both at morphological and molecular level. Application of molecular markers can overcome this bottleneck. 187 accessions comprising of 166 yam landraces and 21 Yam DNA samples from IITA, Nigeria were extracted from leaf samples grown at Muguga and genotyped at BeCA. DNA was extracted using CTAB method. Twelve (12) primer pairs were used for genotyping and PCR products detected on ABI-3730 capillary system. Data was analyzed for genetic diversity, ordination and analysis of molecular variance with GenAlEx software. A total of 131 alleles were amplified with a minimum of two alleles and a maximum of 13 alleles per primer with a minimum allele size of 64 bp and a maximum of 368 bp. Accessions from Eastern province had the highest number of unique alleles. Shannon's information index (I) was 0.1444 for West African samples and 0.2366 for Central province. Accession dispersion revealed four clusters with no distinct geographical pattern. Dense clustering of accessions was an indication of genetic relatedness. Analysis of molecular variance revealed that most variation of 88% ($P < 0.010$) was found within populations or provinces. The simple sequence repeats (SSR) markers were polymorphic and were able to discriminate local yam landraces.

Key words: Genetic diversity, microsatellite, yam, Kenya.

INTRODUCTION

Dioscorea spp., belong to the family Dioscoreaceae and the genus *Dioscorea*. The genus had been reported to have 600 species but has recently been estimated as comprising less than 200 species (Ayensu and Coursey, 1972). Classifications by Hutchinson (1959), Burkill et al. (1960), and Ayensu and Coursey (1972) place the family under the order Dioscoreales.

The taxonomy of yams is complex and further groupings could emerge based on recent molecular biology techniques. For example, controversy exists in the relationships among various species (Terauchi et al., 1992). Kenya's yam diversity is represented by a number of species including *Dioscorea minutiflora* Engl, *Dioscorea bulbifera* L., and *Dioscorea dumetorum* (Kunth) Pax. that

are grown for food by mainly elderly farmers in the Eastern, Central, Western and Coastal regions of the country (Maundu et al., 1999).

There is need for a concise phylogeny investigation to fully understand the local landraces in Kenya vis vis those in the region (Wilkin, P, personal communication). Mwirigi et al. (2009), used morphological markers on 43 Kenyan local landraces and identified four major clusters. Morphological and agronomic features cannot be relied upon and is the major cause of inconsistency in identification. An understanding of the diversity of the local yam germplasm provides a good baseline for further improvement of the crop.

Molecular markers are useful for cultivar identification,

*Corresponding author. E-mail: zkmathamia@yahoo.com. Tel: +254722352305.

because they are not influenced by variable environmental conditions or plant phenology, and are a basis for discriminating among cultivars with similar morphological characteristics (Beebe et al., 2000). Genetic markers are also valuable for determining the phylogenetic relationships among accessions and for true-to-type plant identification (Aggarwal et al., 2008). Examples of molecular markers that have been utilized in phylogenetic analysis include: restriction fragment length polymorphism (RFLPs), and PCR-based markers such as random amplified polymorphic DNA (RAPDs), amplified fragment length polymorphism (AFLP), microsatellites/simple sequence repeats (SSRs) and sequence characterized amplified regions (SCARs).

Good DNA markers should be: highly polymorphic, co dominant, abundant in the genome, evenly distributed throughout the genome, easy and fast to assay, high reproducibility and easy to exchange data between laboratories (Weising et al., 1995). No DNA marker technology currently fulfills all of these criteria. However, amplified fragment length polymorphism (AFLP) and microsatellites or simple sequence repeats (fulfill most of these requirements. Use of molecular markers not only help in genetic diversity and phylogenetic analysis but also enhance genetic gain in selection in a plant breeding program (Spooner, 2005).

Molecular markers such as RFLP, RAPD, AFLP and SSR have been applied in yams for taxonomic, phylogenetic, diversity and mapping studies (Terauchi et al., 1992; Terauchi and Kanoma, 1994; Asemota et al., 1996; Mignouna et al., 1998, 2002a, 2003; Egesi et al., 2006; Tostain et al., 2007).

A study of 45 yam cultivars from eight West African countries using nine SSR loci showed that all the cultivars could be distinguished with three SSR loci (Mignouna et al., 2003). These authors found that the SSR-based classification differed from that based on morphotypes and no geographical structure was observed. In this study, the objective was to describe the diversity of the yam landraces in Kenya using SSRs.

MATERIALS AND METHODS

DNA extraction

Young leaf tissues of 5 g were harvested from 166 yam landraces and yam wild relatives planted at Muguga field genebank. In addition, 21 DNA samples from IITA, Nigeria were sourced making a total of 187 accessions. These were transported in cool boxes with dry ice to BecA for molecular analysis. The accessions from IITA represented *D. alata* L., *D. rotundata* Poir., *D. cayenensis* L., *D. esculenta* (Lour.), *D. bulbifera* and *D. dumetorum*. DNA was extracted following the CTAB method (Doyle and Doyle, 1987) with some modifications (Wanjala et al., 2013).

Optimization of PCR conditions

A total of 12 primer pairs were used (Table 1). PCR amplifications were performed to a final volume of 10 µL consisting of 1X buffer,

2 mM MgCl₂, 0.2 mM dNTPs, 0.08 pmoles /µl each of the forward and reverse primers, 0.375U Taq Polymerase, 4.525 µl of dd H₂O and 25 ng of DNA. GeneAmp PCR system 9700 thermal cycler was used with the following programme: denaturation at 94°C for 3 min, followed by 35 cycles of primer extension stage of 94°C for 30 s, 56°C for 1 min (annealing temperature depending on marker) and 72°C for 2 min, with final extension step of 72°C for 10 min with a final hold at 4°C. PCR optimization was done for all markers and the optimum conditions identified.

Fragment analysis

The PCR products were detected on ABI-3730 capillary system using the LIZ-500 as internal size standard. Data capture was done using the genscan[®] software (Applied Biosystems) and the resulting fragments analysed the alleles scored using the Genemapper[®] software ver4.1. Size matching/ calling was based on Local Southern Method algorithm with reference to a defined standard range, GS75-500 (-250) Liz base pairs. Alleles were sorted according to size and tolerance level of 0.4 base pairs as minimum distance between adjacent bins in base pairs. This process was repeated for each marker until all alleles were binned, from the smallest allele size to the largest sized alleles. A new bin was created each time the threshold tolerance between two sequentially sized alleles was attained. The mean and ranges for all bins were calculated and bin labels rounded up to the nearest whole number was assigned to each group. This data was then fed to the software for bin adjustment. Scored results were exported to an Excel matrix.

SSR data analysis

Alleles were scored and converted to Binary data by use of ALS Binary software. Data was analyzed on 187 accessions for genetic diversity, ordination analysis and Analysis of Molecular Variance (AMOVA) with GenAlEx software (Peakall and Smouse, 2009). In addition, cluster analysis of the cultivars was performed using population genetic analysis software PopGen32 (Yeh et al., 1997). Diversity among and within populations was determined. Genetic distance, matrix dendrogram, F statistics, gene diversity over loci, proportion of polymorphic loci, Shannon index and gene frequency were analyzed (Nei, 1987). To show the relationship between populations and individuals, Principal Co-ordinate Analysis (PCoA) was generated.

RESULTS

Allele frequency

A total of 131 alleles were amplified by 12 SSR primer pairs analyzed in 187 accessions with a minimum of two alleles (Da1F08) and a maximum of 13 alleles (Da1A01 and Dab2D08) and minimum allele size of 64 bp and a maximum of 368 bp. Eastern population had the highest number of private bands (bands unique to a single population) while Coast population had none (Figure 1).

Population genetic diversity

Percentage of polymorphic loci of genotypes from Eastern province was 76.3%, while those from West

Table 1. Yam microsatellite markers and allele sizes detected in 187 genotypes.

Marker name	Min size detected (bp)	Max size detected (bp)	Number of alleles detected	Allele sizes detected	Forward primer sequence (5'-3')	Dye	Reverse primer sequence (5'-3')
YM13	113	254	9	113,184,188,193,209,215,232,235,254	TATAATCGGCCAGAGG	VIC	TGTTGGAAGCATAGAGAA
Dpr3F04	64	148	10	64, 66,71, 127,130,132,135,138,145,148	CCCATGCTTGTAGTTGT	PET	TGCTCACCTCTTTACTTG
Da1A01	141	221	13	141,147,152,155,157,160,164,170,172,180,188,202,221	TGTAAGATGCCACATT	VIC	TCTCAGGCTTCAGGG
Dab2D08	124	368	13	124,146,162,224,236,259,268,279,287,296,338,349,368	ACAAGAGAACCGACATAGT	6-FAM	GATTTGCTTTGAGTCCTT
YM15	180	298	12	180,201,205,208,212,219,221,230,232,234,236,298	TTGAACCTTGACTTTGGT	PET	GAGTTCCTGTCCTTGGT
Dab2DO6	123	180	6	123,129,132,148,162,180	AACATATAAAGAGAGATCA	PET	ATAACCCTTA ACTCCA
Dpr3B12	127	170	8	127,133,135,138,142,147,159,170	CATCAATCTTTCTCTGCTT	NED	CCATCACACAATCCATC
Dpr3F12	130	190	11	130,135,141,148,151,153,156,161,165,172,190	AGACTCTTGCTCATGT	PET	GCCTTGTTACTTTATTC
Dab2CO5	103	221	11	103,146,151,163,173,179,187,194,201,217,221	TCCCCATAGAAACAAAGT	NED	TCAAGCAAGAGAAGGTG
Dab2EO7	104	183	12	104,108,111,128,137,139,142,144,151,162,168,173,183	TTCCCTAATTGTTCTCTTGTG	PET	GTCCTCGTTTTCCCTCTGTGT
Dab2EO9	97	144	9	97, 105,110,115,119,121,124,127,144	TACGGCCTCACTCCAACACTA	VIC	AAAATGGCCACGTCTAATCCTA
Da1FO8	152	198	2	152,198	AATGCTTCGTAATCCAAC	PET	CTATAAGGAATTGGTGCC

Africa had 32.8% (Table 2). The genetic diversity coefficients based on Nei's genetic diversity (H) was lowest for West African populations at 0.0872 while those from Central province had 0.1551. Shannon's information index was 0.1444 for West African samples and 0.2366 for Central province.

To construct the genetic relationships between the six regions, the values of genetic distance were subjected to hierarchical clustering by

unweighted paired-grouping method with arithmetic averages (UPGMA). The cluster analysis separated the sub-populations into two distinct groups A and B which did not reflect their geographic origins (Figure 2). This was corroborated by dendrogram (figure not shown) that gave four clusters.

The accessions dispersion (Table 2) revealed four groups. Clade A was made up of accessions

from all the six regions, with most accessions from Meru in Eastern region and the Central regions clustering together. Five accessions from West Africa representing *D. cayenensis*, *D. esculenta*, *D. rotundata*, *D. alata* and *D. bulbifera* formed part of this clade. Clade B was dominated by accessions from Embu in Eastern province. Clade C represented accessions from the six regions with dense clustering in the upper part of the

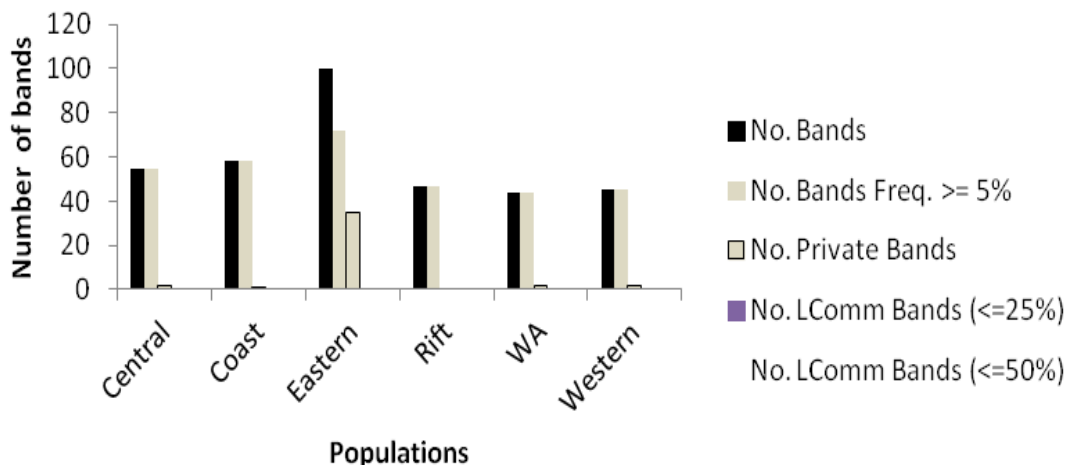


Figure 1. Allele frequency in six regions. No. Bands = number of different bands; No. Bands Freq. $\geq 5\%$ = number of different bands with a Frequency $\geq 5\%$; No. Private Bands = number of bands unique to a single population; No. LComm Bands ($\leq 25\%$) = number of Locally Common Bands (Freq. $\geq 5\%$) Found in 25% or Fewer Populations; No. LComm Bands ($\leq 50\%$) = number of locally common bands (Freq. $\geq 5\%$) Found in 50% or fewer populations.

Table 2. Genetic diversity indices.

Population	H	I	Number of polymorphic loci	Percent polymorphic loci
Central	0.1551	0.2366	54	41.2
Coast	0.1112	0.1870	58	44.2
Eastern	0.1232	0.2086	100	76.3
Rift Valley	0.1407	0.2142	47	35.8
West Africa	0.0872	0.1444	43	32.8
Western	0.1174	0.1824	44	33.5

Mean diversity estimates (H) and Shannon's information index (I).

clade.

Three West African accessions clustered in this clade. Eight accessions from West Africa were represented in clade D together with local landraces. Accessions from West Africa were found in all the four clades.

In clade A, Ac 3079 (*D. bulbifera* from West Africa) clustered closely with Ac 10 and Ac 23 which are both local accessions of species *D. bulbifera*. The clustering of Ac 164 (*D. odoratissima*) with West African *D. rotundata* and *D. cayenensis* suggests close relationship between the three species. Accessions from West Africa were distinct and did not form strong clusters in the four clades as compared to most of the local landraces. Due to the large number of genotypes used, individual groupings within the clades were not properly visualized. There was no clear geographical separation of accessions.

Analysis of molecular variance

Analysis of molecular variance (AMOVA) based on six

populations indicated that most variation was found within populations (88% $P < 0.010$) (Table 3). On the other hand, variation among populations was 12% ($P < 0.001$) among all the major populations studied.

DISCUSSION

A total of 131 alleles were detected using 12 SSR markers. Kolesnikova (2007) reported 281 alleles in 391 accessions of six economically important yam species using 23 SSR markers. Otoo et al. (2009) detected 27 loci using 13 SSR markers in a study of pona complex yam in Ghana. The higher number of alleles detected with some of the markers means they can be used more universally than the ones with low allele detection. The range of allele sizes detected is similar to that obtained by Otoo et al. (2009) and Kolesnikova (2007). From the current study, alleles were detected outside the range suggested by Otoo et al. (2009) and Kolesnikova (2007). This is new information reported in this finding.

tage of polymorphic loci ranged from 32.8% for West Africa to 76.3% for Eastern population. PIC is a measure of marker informativeness related to expected heterozygosity and is calculated from allele frequencies (Bolstein et al., 1980; Hearne et al., 1992). The lower percent of polymorphic loci for West Africa, western Kenya and Rift Valley have to do with the small number of samples obtained from these regions. The genetic diversity coefficient for West Africa population was 0.0872 and that for the Central population 0.1551 while the Shannon's information index, (I) for West Africa 0.1444 and for the Central population 0.2366. Eastern region had the highest percent of unique bands suggesting that some of these landraces may have been domesticated from wild species. Tostain et al. (2007) observed some unique alleles in some local landraces of *D. rotundata* suggesting a possible domestication from *D. burkilliana* or *D. mangenotiana*.

The principal coordinate analysis revealed wide scattering of accessions from the six populations. Accessions from West Africa did not cluster together but scattered among the local landraces in the four clades. The clustering pattern suggests a genetic relationship among the accessions collected from the same regions. Mignouna et al. (2003) reported no geographical structure in a collection of cultivars from eight West African countries but Tostain et al. (2007) reported genetic diversity organized according to cropping regions. The clustering of the local *D. alata* and *D. bulbifera* with similar known checks from West Africa means that the two are correctly identified.

The study indicates that among the local yam landraces there are those closely related to the West Africa yam species and yet there are others that are distantly related. The close clustering of most accessions collected from the same locality suggests possible duplication. Muluneh (2006), using AFLPs showed that Ethiopian accessions were closer to *D. cayenensis* and *D. rotundata* than all the other species studied suggesting involvement of wild species in the process of domestication of yams in Ethiopia. Their finding also showed some clustering together of local landraces distinct from the West African materials. The relationship between *D. cayenensis* and *D. rotundata* has been a subject of debate for close to half a century. On the basis of phylogenetic studies using RFLP analysis in chloroplast and nuclear ribosomal DNA, it was suggested that *D. cayenensis* is a variety of *D. rotundata* (Terauchi et al., 1992). More recent studies based on isozyme (Dansi et al., 2000a) and molecular markers (Ramser et al., 1997; Mignouna et al., 1998, 2005) support the separate identity of the two species.

Though there was no general geographical pattern of accessions dispersion there was evidence of genetic relatedness of many landraces collected from Eastern region. Other workers have not reported geographical relationships with diversity (Muluneh, 2006; Mignouna et

al., 2005). The high diversity within the regions (88%, $P < 0.010$), suggest little or no germplasm exchange between regions. There was low variation among the populations studied (12%, $P < 0.001$). This finding is supported by the fact that yams are exclusively transplanted using corms which are usually bulky and cumbersome to transport thus limiting transfer of propagation material from one region to another.

Conclusion and recommendation

The SSR markers used were polymorphic and were able to discriminate the landraces. More taxonomic work needs to be done to clearly place the local landraces and especially *D. minutiflora* which seems to have many related cultivars. The SSR markers did not group together all the West African species as expected but there was dispersion among the accessions. Landraces from the eastern part of the country presented the widest diversity and highest number of private alleles suggesting a possible centre of dispersal and domestication of yams in Kenya. The study will help eliminate duplicates in the genebank collections thereby cutting maintenance costs.

It is recommended that more phylogeny studies be conducted to elucidate the taxonomy of *D. minutiflora* and its related cultivars.

ACKNOWLEDGEMENTS

This study was financially supported by Swedish Development Agency through Eastern Africa Plant Genetic Resources (EAPGREN). The authors thank the Director KARI for allowing time off to conduct the research work. They also acknowledge the farmers who provided planting materials for the study. We are grateful to the technical staff at IITA, Nairobi for their input in the lab work.

REFERENCES

- Aggarwal M, Shrivastava N, Padh H (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* 27:617-631.
- Asemota HN, Ramser J, Lopez-Peralta C, Weising K, Kahl G (1996). Genetic ariation and cultivar identification of Jamaican yam germplasm by random amplified polymorphic DNA analysis. *Euphytica* 92:341-351.
- Ayensu ES, Coursey DG (1972). Guinea yams. The botany, ethnobotany, use and possible future of yams in West Africa. *Econ. Bot.* 26:301-318.
- Beebe S, Sckroch PW, Tohme J, Duque MC, Pedraza F, Nienhuis J (2000). Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Sci.* 40:264-273.
- Bolstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32:314-331.
- Burkill HM (1960). Organography and the evolution of *Dioscoreaceae* the family of yams. *Bot. J. Linn. Soc.* 56:319-412.
- Dansi A, Mignouna HD, Zoundjehkpon J, Sangare A, Asiedu R,

- Ahoussou N (2000a). Using isozyme polymorphism to assess genetic variation within cultivated yams (*Dioscorea cayenensis* / *Dioscorea rotundata* complex) of the Republic of Benin. *Genet. Resour. Crop Evol.* 46:371-383.
- Doyle JJ, Doyle JL (1987). A rapid DNA isolation procedure for small quantities of fresh leaf material. *Phytochem. Bull.* 19:11-15.
- Egesi CN, Asiedu R, Ude G, Ogunyemi S, Egunjobi JK (2006). AFLP marker diversity in water yam (*Dioscorea alata* L.). *Plant Genet. Resour: Characterization and Utilization* 4:181-187.
- Hearne CM, Ghosh S, Todd JA (1992). Microsatellites for linkage analysis of genetic traits. *Trends Ecol. Evol.* 8:288-294.
- Hutchinson, IJ (1959). The families of flowering plants. II. Monocotyledons. 2nd edition. Clarendon Press, Oxford.
- Kolesnikova-Allen M (2007). Molecular characterization of tier 2 orphan crops-yam. Generation Challenge Project. Final Technical Report, IITA.
- Maundu PM, Ngugi GW, Kabugi CH (1999). Traditional Food Plants of Kenya. National Museums of Kenya, Nairobi.
- Mignouna HD, Abang M M, Fagbemi SA (2003). A comparative assessment of molecular marker assays (AFLP, RAPD and SSR) for white yam (*Dioscorea rotundata*) germplasm characterization. *Ann. Appl. Biol.* 142:269-276.
- Mignouna HD, Abang MM, Wanyera NW, Chikaleke VA, Asiedu R, Thottapilly G (2005). PCR-marker based analysis of wild and cultivated yams (*Dioscorea* spp.) in Nigeria: genetic relationships and implications for ex-situ conservation. *Genet. Resour. Crop Evol.* 52:755-763.
- Mignouna HD, Dansi A, Zok S (2002a). Morphological and isozymic diversity of the cultivated yams (*Dioscorea cayenensis* / *D. rotundata* complex) of Cameroon. *Genet. Resour. Crop Evol.* 49:21-29.
- Mignouna HD, Ellis NTH, Knox M, Asiedu R, Ng QN (1998). Analysis of genetic diversity in Guinea yams (*Dioscorea* spp.) using AFLP fingerprinting. *Trop. Agric. (Trinidad)* 75:224-229.
- Muluneh T (2006). Assessing Diversity in Yams (*Dioscorea* spp.) from Ethiopia based on Morphology, AFLP Markers and Tuber Quality, and Farmers' Management of Landraces. PhD thesis, Georg-August University Göttingen, Germany.
- Mwirigi PN, Kahangi EM, Nyende AB, Mamati EG (2009). Morphological Variability within the Kenyan yam (*Dioscorea* spp.). *J. Appl. Biosci.* 16:894-901.
- Nei M (1972). Genetic distance between populations. *Am. Nat.* 106:283-292.
- Otoo E, Akkromah R, Kolesnikova-Allen M, Asiedu R (2009). Ethnobotany and morphological characterization of the yam pona complex in Ghana. In: *Proceedings of the African Crop Science* 9:407-414.
- Peakall R, Smouse PE (2006). GENALEX 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6:288-295.
- Ramser J, Weising K, Lopez-Peralta C, Terhalle W, Terauchi R, Kahl G (1997). Molecular marker based taxonomy and phylogeny of Guinea yam (*Dioscorea rotundata*-*D. cayenensis*). *Genome* 40:903-915.
- Spooner D, R van Treuren, de Vicente MC (2005). Molecular markers for genebank management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute, Rome, Italy.
- Terauchi R, Chikaleke VA, Thottapilly G, Hahn SK (1992). Origin and phylogeny of Guinea yams as revealed by RFLP analysis of chloroplast DNA and nuclear ribosomal DNA. *Theor. Appl. Genet.* 83:743-751.
- Terauchi R, Konuma A (1994). Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. *Genome* 37:794-801.
- Tostain S, Agbangla C, Scarcelli N, Mariac C, Dainou O, Berthaud J, Pham JL (2007). Genetic diversity analysis of yam cultivars (*Dioscorea rotundata* Poir.) in Benin using simple sequence repeats (SSR) markers. *Plant Genet. Resour: Characterization and utilization* 5(2):71-81.
- Tostain S, Scarcelli N, Pham JL, Noyer JL, Marchand JL, Brottier P (2006). Development of DNA microsatellite markers in tropical yam (*Dioscorea* sp.). *Molecular Ecology Notes* 6:173-175.
- Wanjala BW, Obonyo M, Wachira FN, Harvey J, Muchugi A, Mulaa M, Harvey J, Skilton RA, Proud J, Hanson J (2013). Genetic diversity in Napier grass (*Pennisetum purpureum*) cultivars: implications for breeding and conservation. *AoB PLANTS* 5:plt022;doi:10.1093/aobpla/plt022.
- Weising K, Nybom H, Wolff K, Meyer W (1995). DNA Fingerprinting in Plants and Fungi (ed. Arbor, A.) CRC Press, Boca Raton. pp. 1-3.
- Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX (1997). POPGENE, the User-Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.