African Journal of Biotechnology Vol. 11(18), pp. 4140-4157, 1 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB10.1738
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Molecular characterisation of *Musa* L. cultivars cultivated in Malawi using microsatellite markers

Wisdom Changadeya[1]* Emmanuel Kaunda[2] and Aggrey J. D. Ambali[3]

[1]Department of Biology, University of Malawi, Chancellor College, P.O. Box 280, Zomba, Malawi.
[2]Department of Aquaculture and Fisheries Science, University of Malawi, Bunda College of Agriculture, P.O Box 219, Lilongwe, Malawi
[3]NEPAD African Biosciences Initiative, Policy Alignment and Programme Development Directorate, NEPAD Agency, c/o CSIR Building 10F, Meiring Naude Road, Brummeria, Pretoria, Republic of South Africa, 0001.

Accepted 16 February, 2012

Genetic diversity and relationships were assessed in 141 locally named banana cultivars growing in five districts of Malawi at 12 microsatellite loci. High allelic diversity (174 alleles) attributable to high frequency of duplicated alleles was observed. Primers discriminating power was high with mean polymorphism information content (PIC) of 0.74. Low mean genetic diversity contrary to actual allele count was detected by Nei's gene diversity index (h, 0.12) and Shannon information index (I, 0.18). The low genetic diversity estimation could be due to loss of co-dominance by simple sequence repeats (SSR) loci in polyploids which leads to underestimation of allelic relationships in populations. The results also reveal that the genetic diversity of bananas from Chitipa, Karonga, Nkhata Bay, Thyolo, Mulanje, BRS local collection and BRS gene bank is not significantly different. Pooled northern region population (Chitipa, Karonga and Nkhata Bay) is as genetically diverse as pooled southern region population (Thyolo, Mulanje, Local collection and Gene bank). Cluster analysis showed that most cultivars were dissimilar probably due to multiplicity of mutations generated by high rate of cultivar multiplication by farmers. However, genetic relationships among some cultivars showed some possible synonyms.

Keywords: Bananas and plantains cultivars, genetic diversity, microsatellite markers.

INTRODUCTION

Bananas also referred to as *Musa L.*, are perennial monocotyledonous herbs characteristically divided into dessert bananas, cooking bananas and plantains and beer bananas (Robinson, 1996; Strosse et al., 2006). Their major genomic groups include diploids (AA, AB and BB), triploids (AAA, AAB, ABB and BBB) and tetraploids (AAAA, AAAB, AABB and ABBB) (INIBAP, 2003). Bananas and plantains were derived from two species of genus Musa*, Musa acuminata* (AA) and *Musa balbisiana* (BB) through inter or intra-specific hybridization (Stover and Simmonds, 1987; Pollefeys et al., 2004). *Musa L.* are thought to have originated from the South-East Asia and Western Pacific regions where their inedible, seed-bearing diploid ancestors can still be found in the wild (Simmonds, 1966). The Asian–Pacific region is the primary centre of diversity for bananas while humid lowlands of West and Central Africa and East Africa Highlands are known to be secondary centres of diversity for plantains and cooking bananas respectively (Pillay et al., 2004). African bananas are grouped into three categories including East African (mainly dessert) bananas (AA, AAA, ABB and AB), the African plantain bananas (AAB) grown mainly in central and West Africa and the East African Highland Banana (AAA), used for cooking and beer brewing (Karamura, 1998).

Bananas and plantains are the fourth most important crop in the developing world after rice, maize and wheat in terms of gross value of production (Ortiz and Vuylsteke, 1996; FAOSTAT, 2007). Estimates put world banana and plantain production at about 100 million

*Corresponding author. E-mail: wchanga@chanco.unima.mw, wisdomchangadeya@gmail.com.
Abbreviation: SSR, Simple sequence repeats.

tonnes (FAOSTAT, 2008). Dessert bananas, usually eaten fresh when ripe, and cooking bananas, which are starchier when ripe and are boiled, fried or roasted constitute 43 and 57% of the world production respectively (Jones, 2000). Cavendish bananas, which are dessert bananas constitute the most commercially important component of world banana production, accounting for 47% of the global banana production (Arias et al., 2003).

In Malawi, Bananas and plantains are ranked sixth in their contribution to both food security and income generation, after maize, rice, groundnuts, vegetables and beans in their decreasing order (Malawi Government, 1995). Though important, banana and plantain production is lower compared to other countries in the region such as Tanzania, Rwanda and Burundi. According to Food and Agriculture Organization (FAO) production data of 2001, Tanzania produced 1,452,678 metric tones, Rwanda (1,572,661), Burundi (1,548,897) while Malawi produced 293,000 tonnes of bananas and plantains. Yet per capita consumption of bananas and plantains of Malawi (23.3 kg/ca/yr) in 2000 was higher than Tanzania (14.5 kg/ca/yr), an indication that there is sufficient consumptive demand for bananas in Malawi (FAO, 2000). Yields, as low as 5 t ha $^{-1}$ per year against the potential yield of 60 to 120 t ha $^{-1}$ per year, are common (Banda and Mwenebanda, 2000). The below potential crop performance is attributed to several factors such as inferior cultivars, pests and diseases, deteriorating soil fertility and poor management practices.

However, there is enormous potential for increasing yields in *Musa* crops through genetic improvement (de Vries et al., 1967). Considerable advances have been made in understanding the genetic basis of specific traits in *Musa* elsewhere (Ortiz, 1995; 1997), which have become the basis for genetic improvement programs. Consequently, there was a need to properly and accurately identify banana cultivars growing in Malawi before engaging in genetic improvement programs. A better understanding of the extent and patterns of genetic diversity of bananas and plantains from ecologically different banana growing districts in Malawi is essential for banana and plantain conservation and utilisation.

Molecular markers are powerful tools for fingerprinting, assessing genetic variation and studying relatedness among cultivars of many species thereby improving the efficiency of *Musa* breeding and conservation. This is due, in part, because they are not influenced by the environment (Martínez et al., 2006). *Musa* genetic studies have applied different molecular markers including randomly amplified polymorphic DNA (RAPD) markers which Pillay et al. (2001) used to investigate genetic diversity and phylogenetic relationships of 29 East African banana cultivars and two outgroup taxa, *M. acuminata* Calcutta 4 and Agbagba. The results reveal that East African bananas are closely related with a narrow genetic base. A dendrogram derived from the RAPD data was divisible into a cluster composed of all AAA highland banana cultivars and Agbagba (AAB) and a minor cluster consisting of Kisubi (AB), Kamaramasenge (AB) and Calcutta 4 (AA). However, RAPD data did not separate beer and cooking banana cultivars both of AAA genome composition (Pillay et al., 2001). Ude et al. (2002) reported that using amplified fragment length polymorphism (AFLP) marker on 28 accessions of *M. acuminata* (AA) Colla and *M. balbisiana* (BB) Colla, Neighbour-joining and principal co-ordinate (PCO) analyses using Jaccard's similarity coefficient produced four major clusters that closely corresponded with the genome composition of the accessions (AA, BB, AAB, and ABB). The AFLP data distinguished between wild diploid accessions and suggested new subspecies relationships in *M. acuminata* complex that are different from those based on morphological data. The data suggests three subspecies within the *M. acuminata* complex (ssp. *burmannica* Simmonds, *malaccensis* Simmonds and *microcarpa* Simmonds. A study by Creste et al. (2004) that used microsatellites markers on *Musa* germplasm in Brazil, detected a large number of alleles suggesting a large genetic variability present and evidence of the multi-allelic nature of microsatellite. The study also reveals similarity between diploid and triploid accessions implying possibility of potential crosses to maximize the recovery of typical fruit qualities required in Brazil's (AAB, Pome and Silk dessert banana). Among all these markers, microsatellites have become the most popular and are considered the most powerful in terms of their ressolving power (Buhariwalla et al., 2005; Creste et al., 2004).

In Malawi, accurate classification and identification of banana cultivars has not been done. Hence several cultivars under cultivation can not be accurately identified because they have not been genetically analyzed and identified. Such cultivars were identified using traditional system of classification and identification of banana cultivars which is primarily based on morphological and quantitative traits (Stover and Simmonds, 1962). Therefore, this study was undertaken to test the null hypothesis that assumed that different vernacular names assigned to various banana cultivars in different districts mirrored genetic differences among the cultivars. Microsatellite markers were applied to assess genetic diversity and relationships among various banana cultivars known by their local names growing in different agro-ecological zones of Malawi.

MATERIALS AND METHODS

Collection of banana plant specimens

Banana leaf specimens were collected from five major banana-producing districts in Malawi from November 2005 to January 2006 (Table 1 and Figure 1). Three to four young and tender banana leaf discs were collected into 2 ml micro-centrifuge tubes by punching the discs directly into the tubes, using the lid as a punch. A total of

Table 1. Banana sample collection sites and sample composition.

Banana growing district	Name of Village/place where specimen were collected	Number of different banana cultivars (local names) collected	n
Mulanje (MJ) (South)	Khumbanyiwa Village (Limbuli)	17	85
	Bondo Village	5	25
	Subtotal	**22**	**110**
Thyolo (TO) (South)	Ntanangale Village	4	20
	Nkolokosa Village	9	45
	Thekelani	5	25
	Nandewe	2	10
	Molele	2	10
	Chamela	1	5
	Bvumbwe Research Station (BRS)		
	Local Colllection (LC)	19	95
	Gene Bank (GB)	18	90
	Subtotal	**60**	**300**
Nkhata Bay (NB) (North)	Kalowa Village (Dwambadzi area)	5	25
	Chinyakula Village (Kawalazi area)	10	50
	Chighachang'ombe Village (Lwazi area)	5	25
	Subtotal	**20**	**100**
Karonga (KA) (North)	Mwahimba Village (Karonga Boma)	12	60
	Kwiyula Village (near Karonga Teachers College)	4	20
	Mwandambo Village (Songwe River area)	5	25
	Mwakaboko Village (Iponga Area)	4	20
	Mwakasebwe Village (Iponga Area)	2	10
	Subtotal	**27**	**135**
Chitipa (CP) (North)	Misuku Hill area	12	60
	Total	**141**	**705**

five plants (individuals) growing next to each other were collected representing each locally named banana cultivar. Specimens were randomly collected in major banana growing areas of the district in collaboration with smallholder banana farmers of the area. A total of 705 specimens were collected under 141 local names in the five districts (Table 1; Supplementary data Appendix 1). The leaf samples were transferred to Chancellor College, Molecular Biology and Ecology Research Unit, DNA laboratory in Zomba, Malawi for analysis. Seven different banana populations were recognized and used in this study namely; Mulanje (MJ), Thyolo (TO), Local collections at Bvumbwe Agricultural Research Station in Thyolo (LC), Gene Bank at Bvumbwe Agricultural Research Station in Thyolo (GB), Nkhata Bay (NB), Karonga (KA) and Chitipa (CP).

DNA extraction and primer amplification

Total DNA was extracted following cetyltrimethylammonium bromide (CTAB) protocol as described by Gawal and Jarret (1991) with minor modifications. Briefly, three or four fresh leaf discs were ground in the presence of Carborudnum in a 2.0 ml microcentrifuge tube. A total of 400 µl of extraction buffer (1.5% CTAB); 100mM Tris-HCl; 20 mM EDTA; 1.4 mM NaCl; 0.2% ß-mercaptoethanol) at 60°C was added and the mixture incubated at 60°C for 60 min in an automated Advantec water bath. After incubation, 400 µl of chloroform: isoamylalcohol (24:1) was added to the microcentrifuge tube and the homogenate mixed on a shaker for 20 min at room temperature. This was followed by centrifugation at 15,000 rpm for 10 min in a Tomy high speed micro-centrifuge. Subsequently, 350 µl of supernatant was transferred to a 2.0 ml microcentrifuge tube and the DNA in the supernatant was precipitated in 210 µl of cold isopropanol at -20°C for 60 min. The precipitated DNA was separated from the suspension by centrifugation at 10,000 rpm for 10 min, the supernatant was decanted and the resultant pellet was washed in 500 µl of 70% cold ethanol and centrifuged again for 5 min. The ethanol was decanted and the DNA pellets air dried for 10 to 15 min before dissolution in 50 µl Tris-EDTA (TE).

DNA amplification

Microsatellite amplification reactions were done using 12 primers (Table 2) that produced distinct reproducible amplification products. The reactions were completed in final polymerase chain reaction (PCR) volumes of 13.11 µl consisting of 5.7 µl PCR grade water, 1 µl of 10 mM dNTP mix, 1.25 µl of 10 X PCR buffer, 1.6 µl of 25 mM magnesium chloride ($MgCl_2$), 0.75 µl of 15 pmol of both forward and reverse banana microsatellite primers, 0.06 µl of 5 u/µl *Taq* DNA

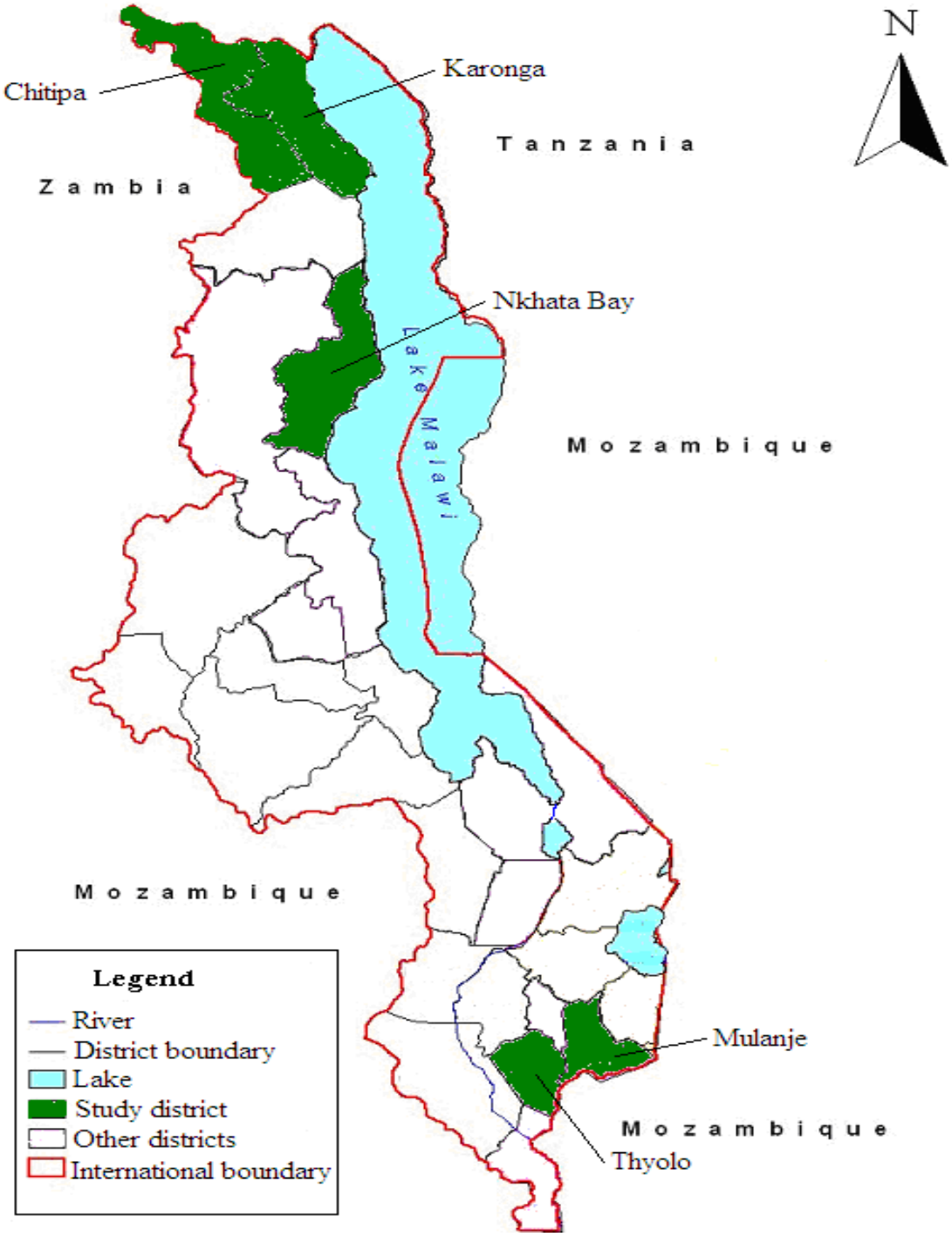


Figure 1. Sampled major banana producing districts.

polymerase in storage buffer A (Promega, USA) and 2 µl of 25 ng/ul template DNA. The amplifications were carried out in a Mastercycler gradient 5331 Eppendorf Version 2.30.31-09 with the following PCR conditions: denaturation step at 94°C for 30 s, followed by 30 amplification cycles of denaturing at 89°C for 30 s, annealing at an optimal temperature for a specific primer pair for 15 s (Table 2) and elongation at 72°C for 30 s. The final extension was at 65°C for 20 min followed by a soaking temperature of 4°C.

Detection of PCR products using silver staining technique

PCR products were resolved using 6% polyacrylamide gel electrophoresis. The 6% polyacrylamide gel was poured in BIORAD Sequi-Gen® GT nucleic acid electrophoresis cell. A total of 6 µL of STR 3X loading solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanol FF) was added to the PCR products in 0.2 ml PCR tubes and denatured in mastercycler

Table 2. Microsatellite (SSR) primers used in this study.

Locus		Primers sequences (5'-3')	Primer source	T_{ann} (°C)	Allele size range observed
STMS1FP/1RP	F	TGAGGCGGGGAATCGGTA	Kaemmer et al., 1997	58 °C	110-154
	R	GGCGGGAGACAGATGGAGTT			
STMS7FP/7RP	F	AAGAAGGCACGAGGGTAG-	Kaemmer et al., 1997	54 °C	234-252
	R	CGAACCAAGTGAAATAGCG			
STMS10FP/10RP	F	ATGATCATGAGAGGAATATCT	Kaemmer et al., 1997	53 °C	116
	R	TCGCTCTAATCGGATTATCTC			
AGMI93/94	F	AACAACTAGGATGGTAA-TGTGTGGAA	Lagoda et al., 1998	57.5°C	124-158
	R	GATCTGAGGATGGTTCTGTT-GGAGTG			
Ma 1/16	F	TTTGCCTGGTTGGGCTGA	Crouch et al., 1998	58.5°C	140-166
	R	CCCCCCTTTCCTCTTTTGC			
Ma 1/17	F	AGGCGGGGAATCGGTAGA	Crouch et al., 1998	57.7 °C	100-156
	R	GGCGGGAGACAGATGGAGT			
Ma 1/18	F	TTTGCCTGGTTGGGCTGA	Crouch et al., 1998	57.5 °C	138-182
	R	CCCCCCTTTCCTCTTTTGC			
Ma 1/24	F	GAGCCCATTAAGCTGAACA	Crouch et al., 1998	54°C	154-168
	R	CCGACAGTCAACATACAATACA			
Ma 1/27	F	TGAATCCCAAGTTTGGTCAAG	Crouch et al., 1998	54°C	114-132
	R	CAAACACATGTCCCCATCTC			
Ma 3/2	F	GGAACAGGTGATCAAAGTGTGA	Crouch et al., 1998	56.2°C	204-238
	R	TTGATCATGTGCCGCTACTG			
Ma 3/90	F	GCACGAAGAGGCATCAC	Crouch et al., 1998	53.2 °C	128-162
	R	GGCCAAATTTGATGGACT			
Ma 3/103	F	TCGCCTCTCTTTAGCTCTG	Crouch et al., 1998	54°C	138-160
	R	TGTTGGAGGATCTGAGATTG			

F, Forward; R, reverse, T_{ann} (°C), annealing temperature.

gradient at 95°C for 5 min. Then 6 µl of denatured PCR products was loaded on the 6% polyacrylamide gel and ran at 50 W. The gel plates were fixed, stained and developed following the direction reported in the Promega Silver Sequence™ DNA Sequencing System Technical Manual. The microsatellites bands were scored over a light box using pGem DNA marker (Promega, USA) and ϕ X174 DNA/*Hinf* 1 (Promega, (USA) as band size standard markers.

Data analysis

Statistical analysis of loci polymorphism

Polymorphism information content (PIC), a variability measure of each locus, was calculated as described by Saal and Wracke (1999):

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} p_i^2$$

Where p_i is the frequency of the *i*th allele out of the total number of alleles at simple sequence repeats (SSR) locus, and *n* is the total number of different alleles for that locus.

Analysis of genetic variation

Estimating the exact number of copies of individual alleles is difficult

among polyploidy species hence data is often analyzed as a binary data matrix and SSR markers are treated as dominant markers (Mengoni et al., 2000; Lian et al., 2003). Consequently, the presence or absence of each PCR amplification product was scored as "1" or "0", respectively and data matrix was generated. Using the data matrix, POPGENE Version 1.31 freeware (Yeh et al., 1999) was employed to compute several measures of genetic variation within and between banana cultivar populations. The following variables were computed: observed number of alleles (na), mean number of alleles (A) (Kimura and Crow, 1964), Nei's genetic diversity (h) (Nei, 1973) and Shannon's information index (I) (Lewontin, 1974). Significance of various analyses was determined by the formula (Mean±2SE). Means were compared at 95% level of significance using unpaired *t* test with Welch's correction performed by GraphPad Prism version 3.00 for Windows, GraphPad Software, San Diego, California, USA, www.graphpad.com. Genetic differentiation (Gst) (Nei, 1973), which measures among population component of genetic variance was calculated to determine the proportion of total variation that was due to differences between population allele frequencies. Other population variance measurements such as total heterozygosity (Ht) and gene diversity of individuals relative to their population (Hs) were also determined (Nei, 1973). Genetic distances between pooled northern and southern populations were calculated as described by Nei and Li (1979) using the following equation:

$$GD_{nor\text{-}sou} = 1 - \left[\frac{2N_{nor\text{-}sou}}{(N_{nor} + N_{sou})}\right]$$

Where N_{nor} and N_{sou} are the numbers alleles in northern and southern pooled populations respectively, and $N_{nor\text{-}sou}$ is the number of alleles shared by the two regional populations.

Analysis of cultivar similarity

The PCR amplified products from individual plants were scored as either present (1) or absent (0) to create binary matrices, which were used to determine genetic similarity among banana cultivars. The binary matrices were analysed using NTSYSpc version 2.11c (Rolhf, 2001). Pair wise similarity matrices were computed using simple matching coefficient (SM) (Sokal and Michener, 1958). The similarity matrices were used to construct dendrograms from the Sequential Agglomerative Hierarchical and Nested (SAHN) clustering method using the unweighted pair-group method with arithmetical averages (UPGMA) (Sneath and Sokal, 1973). Bootstrap analysis, which is a method for determining confidence limits of clusters produced by UPGMA-based dendrograms, was performed using Win Boot program (Yap and Nelson, 1996). In order to obtain statistically accurate bootstrap *p* values at 95% level, dendrograms for all banana cultivars (Appendix 1, Supplementary data), northern region banana cultivars (Figure 2) and southern region banana cultivars (Figure 3) had boostrap replications of 230, 950 and 580, respectively.

RESULTS AND DISCUSSION

Polymorphism of microsatellite markers

PIC of microsatellite primers averaged 0.74 with a range of 0.00 (STMS10FP/10RP)-0.88 (Ma 1/17), implying that the primers exhibited high polymorphism and discriminatory power. Botstein et al. (1980) described that a PIC value >0.5 indicates highly polymorphic locus while a value <0.25 refers to a lower polymorphic locus. Due to the high discriminatory power of primers used in the study, the consequential cultivar identifications and molecular relationships as shown in dendrograms are strongly supported. Venkatachalam et al. (2008) showed in their genetic study of Indian bananas that there is a strong linear relationship between discriminating power of microsatellite primers and their ability to distinguish genotypes.

Allelic diversity

Allele frequency distribution plots (data not shown) indicated a lot of allelic diversity especially at loci Ma1/17, Ma1/18, STMS 1 FP/1RP, Ma3/90 and Ma 3/2, which were the most polymorphic loci (Table 3). A total of 174 alleles with allele mean number of 14.5 and a range of one (STMS10FP/RP) to 30 (Ma 1/17) per locus were generated. The total number of alleles in the pooled south population (156 alleles) was larger than that of pooled north population (139 alleles). The mean numbers of alleles per locus between the two regions were similar, 13 and 12 alleles, respectively. A relatively larger number of alleles were observed in the present study compared to earlier reports indicating the existence of high genetic variability among Malawian bananas. For instance, Creste et al. (2004) in their study of assessing genetic diversity in 58 Brazilian *Musa* L. accessions using nine SSRs found 115 alleles while Oriero et al. (2006) reported 23 alleles using nine SSRs in 40 accessions and Ning et al. (2007) obtained 92 alleles from 10 SSRs in a study of 216 banana accessions. The high allelic numbers among Malawian banana cultivars may be attributable to the high frequency of duplicated alleles or duplicated chromosomal regions, which are common in *Musa L*. polyploid genomes (Creste et al., 2003).

A summary of mean observed number of alleles (na), Nei's genetic diversity (h) measure and Shannon's information index (I) for all populations is presented in Table 4. The results reveal that na, h and I were not significantly different in all populations ($p \geq 0.05$). When northern region populations (Chitipa, Karonga and Nkhata Bay) and southern region populations (Thyolo, Mulanje, Gene Bank and Local collection) are pooled together into separate regional groupings, their mean observed number of alleles, Nei's gene diversity and Shannon's Information index are not significantly different between the two regional populations ($p \geq 0.05$).

Shannon's information index (I) ranged from 0.15 to 0.21 while Nei's gene diversity (h) varied from 0.10 to 0.14. Both indices revealed low diversity among banana populations which is possibly due to the loss of co-dominancy of SSR loci in polyploid species like *Musa L*., leading to either overestimation or underestimation of allelic relationships in a population (Creste et al., 2004). The co-dominant nature of microsatellite markers is one

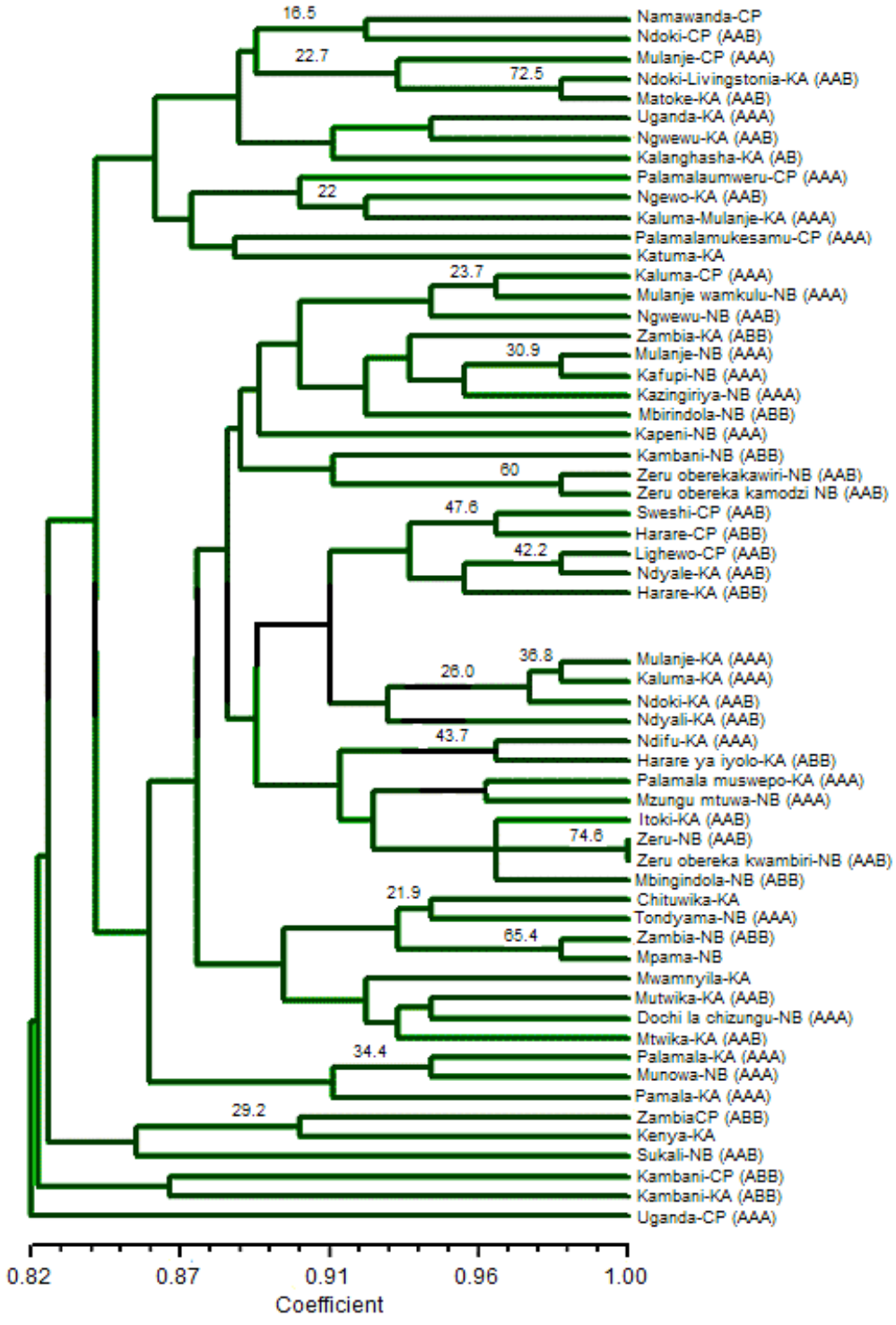


Figure 2. Genetic relationships among northern region (Chitipa, Karonga and Nkhata Bay) banana cultivars analysed with microsatellites. Number at node or side indicates bootstrap values in percentages.

of the major advantages ascribed to these markers that allow an absolute estimation of allelic relationship between individuals. However even if high proportion of alleles is revealed by SSR markers in polyploid species, they do not render identification of actual allelic relationships between individuals possible as they are treated as dominant markers. The situation is aggravated by the fact that *Musa L.* germplasm have

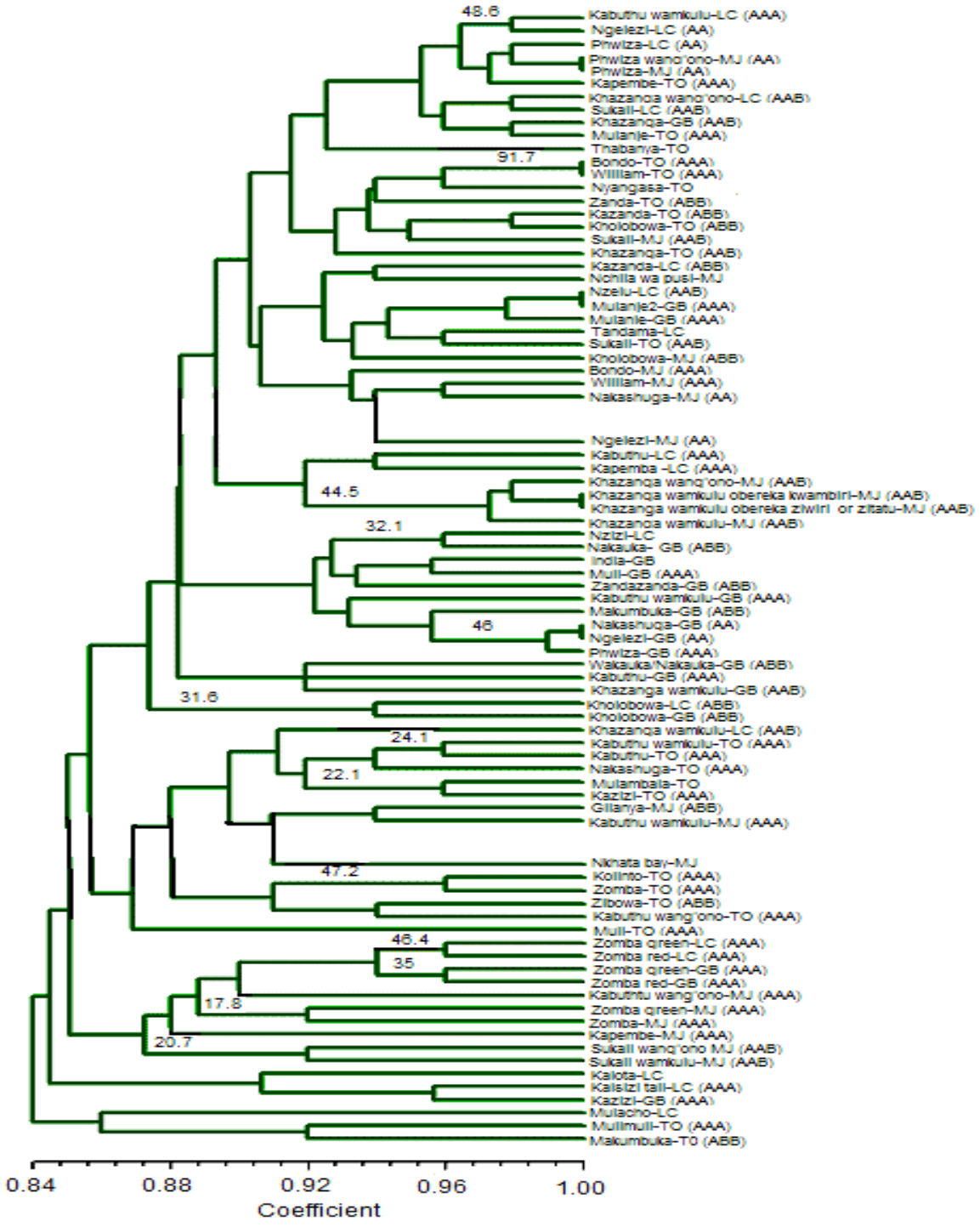


Figure 3. Genetic relationships among southern region (Thyolo, Mulanje, local collection and Gene Bank) banana cultivars analysed with microsatellites. Number at node or side indicates bootstrap values in percentages.

complex genomic nature in all of the polyploidy levels, involving much chromosomal duplication, which limits the use of SSR markers as co-dominant markers (Crouch et al., 1999; Creste et al., 2004). Such findings are not limited to bananas, similar results have been found in sexually reproducing bean local variety 'Flor de Junio' of Mexico (h = 0.06, I = 0.10) although reduction in genetic diversity in this case would be due to self pollination and inbreeding within single landraces (Zizumbo-Villarreal et al., 2005; Kumar et al., 2008).

Table 3. Total number of alleles (A) and allele size range (SR) in base pairs (bp) in *Musa L.* populations at 12 loci.

Population	STMS1FP1RP		STMS7FP7RP		STMS10FP10RP		AGMI93/94		Ma 1/16		Ma 1/17	
	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR
Local collection (LC)	11	116-150	5	242-250	1	116	7	124-144	10	142-164	9	110-134
Gene Bank (GB)	9	116-150	9	234-252	1	116	7	124-144	11	142-166	6	106-128
Chitipa (CP)	8	116-146	3	246-250	1	116	4	124-144	11	140-162	9	108-140
Karonga (KA)	14	116-154	4	244-250	1	116	8	124-150	13	140-166	24	100-156
Nkhata bay (NB)	13	116-148	4	246-252	1	116	9	124-150	11	142-166	19	104-154
Thyolo (TO)	12	116-146	6	242-252	1	116	11	124-152	8	144-158	20	106-154
Mulanje (MJ)	12	110-150	8	238-252	1	116	10	124-158	10	142-162	17	106-154
All	19	110-154	10	234-252	1	116	14	124-158	14	140-166	30	100-156
Mean primer PIC		0.85		0.69		0.00		0.78		0.86		0.88

Population	Ma 1/18		Ma 1/24		Ma 1/27		Ma 3/2		Ma 3/90		Ma 3/103	
	A	SR	A	SR	A	SR	A	SR	A	SR	A	SR
Local collection (LC)	13	138-176	6	156-166	7	116-128	11	204-228	13	128-162	8	140-158
Gene Bank (GB)	10	150-178	4	158-164	4	116-126	5	216-224	11	130-156	7	140-152
Chitipa (CP)	13	144-174	8	154-168	5	116-126	7	220-236	7	132-144	5	142-150
Karonga (KA)	10	144-172	6	154-164	7	116-132	9	222-238	12	132-158	8	140-154
Nkhata Bay (NB)	11	142-182	5	156-164	7	116-128	6	222-234	7	132-146	11	138-158
Thyolo (TO)	12	140-172	7	154-166	6	116-128	8	220-234	14	132-160	8	140-160
Mulanje (MJ)	14	138-166	6	156-166	9	114-130	5	222-234	8	132-150	7	140-152
All	22	138-182	8	154-168	10	114-132	17	204-238	17	128-162	12	138-160
Mean primer PIC		0.87		0.76		0.76		0.80		0.85		0.82

BRS, Bvumbwe Agricultural Research Station, Thyolo; PIC, polymorphism information content.

The average Nei's gene diversity for this study was 0.12 which was threefold lower than values observed in banana landraces from Myanmar (Burma) (h = 0.32) and in cultivars obtained from the International Network for the Improvement of Banana and Plantain (INIBAP) (h = 0.34). The studies involving banana cultivars from Myanmar and INIBAP used P450-based analog (PBA) markers, which are derived from specific multi-gene family sequences (Yamanaka et al., 2003). Myanmar is a country among Southeast Asian countries from where banana originated and also represents the primary centre of banana diversity. It is common practice among banana breeders to use Myanmar landraces in breeding programme to broaden the genetic diversity of their working collection and revive the lost diversity (Wan et al., 2005). Although the genetic diversity observed among Malawian banana cultivars is three times lower than that of the primary centre of banana diversity, some considerable genetic diversity has been maintained or generated through mutation in Malawian bananas.

Population structure of banana cultivars

The statistics on population structure and heterozygosity are summarised in Table 5. The Gst values indicated a

Table 4. Mean genetic variability for *Musa L.* populations at 12 microsatellite loci. Genetic diversity index values are averages±SD of the mean.

Population	n	Number of cultivars collected	na	h	I
LCBTO(LC)	95	19	1.41±0.50 ns[b]	0.11±0.17 ns[b]	0.18±0.24 ns[b]
GBBTO (GB)	90	18	1.31±0.47 ns	0.10±0.16 ns	0.15±0.24 ns
Chitipa (CP)	60	12	1.35±0.48 ns	0.13±0.19 ns	0.19±0.28 ns
Karonga (KA)	135	27	1.43±0.50 ns	0.14±0.18 ns	0.21±0.27 ns
Nkhata Bay (NB)	100	20	1.37±0.49 ns	0.11±0.17 ns	0.17±0.25 ns
Thyolo (TO)	115	23	1.37±0.49 ns	0.11±0.17 ns	0.17±0.25 ns
Mulanje (MJ)	110	22	1.37±0.49 ns	0.11±0.17 ns	0.17±0.25 ns

na, observed number of alleles; h, Nei's gene diversity; I, Shannon's information index; BTO, Bvumbwe Agricultural Research Station in Thyolo. [b]Means within the columns followed by ns were not significantly different at P=0.05 based on unpaired t test with Welch's corrections.

Table 5. Mean gene diversity in subdivided populations at 12 loci.

Population	n	Ht±SD	Hs±SD	Gst
Northern region (N)	295	0.14±0.03* [a]	0.13±0.03*[a]	0.11
Southern (S) region	407	0.12±0.03 *	0.11±0.02*	0.12
All populations (All)	701	0.13±0.03*	0.12±0.02*	0.14

Northern region = CP, KA and NB; Southern region = LC, GB, TO and MJ; Gst, genetic differentiation; Ht, gene diversity over all groups/total heterozygosity; Hs, gene diversity of individual relative to their population/heterozygosity in each generation; [a] means within the columns followed by *were significantly different at P=0.05 based on unpaired t test with Welch's corrections.

total population differentiation of 14%, with the northern populations being slightly less differentiated (11%) than the southern populations (12%). Ht and Hs was significantly higher in the northern populations than in the southern populations ($p \leq 0.05$).

The differentiation value (Gst = 0.14) observed in the present study is lower than the average value reported for crop species (Gst = 0.34) (Oriero et al., 2006). The differentiation value is also lower than those reported in studies of wild *M. balbisiana* Colla that used cpDNA PCR-RFLP (Gst = 0.77) and SSR data (Fst=0.35) (Ge et al., 2005). However, it is slightly higher than that observed in the Brazilian diploid and triploid banana accessions using SSR data (Rst=0.105) (Creste et al., 2004). The lower differentiation observed among domesticated bananas than in wild species is expected due to reduction in genetic diversity through domestication process. Domestication is aided by human interventions and artificial selection of preferred banana cultivars resulting in the disappearance of genes in the course of time, thereby reducing the total genetic diversity of the crop species.

The low genetic differentiation among banana accessions in Malawi was further confirmed by a low genetic distance ($GD_{nor\text{-}sou}$=0.17) coupled with an average of 10 common/shared alleles (data not shown) observed in the pooled data of north and south populations.

The pooled analysis indicated that Ht and Hs of the northern region was significantly higher than that of the southern region ($p \leq 0.05$), although the total number of alleles recorded was lower (139) in the northern population than southern population (156). This indicates the presence of more rare alleles in the northern populations allowing greater gene combinations and potential adaptability to climatic changes, pests and disease pressures such as *Fusarium wilt* in Chitipa and Karonga and Black Sigatoka and Banana Bunchy Top disease in Nkhata Bay. Proximity to East Africa, which is a secondary centre of banana diversity, could also explain the possibility of the observed heterozygosity and probable germplasm admixture resulting in the origin of the rare alleles in the northern banana populations. Besides, genetic variability among and within plant populations reflects the interactions of different processes including the long-term evolutionary history of the species (shifts in distribution, habitat fragmentation and population isolation) mutation, genetic drift, mating system, gene flow and selection (Zhao et al., 2007).

Cluster analysis and genetic relationships among banana cultivars

The dendrograms showing similarity among northern and southern region banana populations are provided in Figures 2 and 3, respectively. Cluster analysis dendrogram for all cultivars is in Appendix 2 (Supplementary data). The significance level for each cluster generated through bootstrap analysis was below 95%. According to Felsenstein (1985) clusters are

Table 6. Local banana cultivars of the same genome with (100%) similarity in genetic relationship.

Local name	Another local name	Genome	Bootstrap *p* value	Reference dendrogram	Similarity level (%)
Zeru-NB	Zeru obereka kwambiri-NB	AAB	74.6	Figure 2	100
Phwiza wang'ono-MJ	Phwiza-MJ	AA		Figure 3	100
Bondo-TO	William-TO	AAA	91.7	Figure 3	100
Khazanga wamkulu wobereka kwambiri-MJ	Khazanga wamkulu obereka ziwiri or zitatu-MJ	AAB	44.5	Figure 3	100
Nakashuga-GB	Ngelezi-GB	AA	46	Figure 3	100

considered significant only if bootstrap *p* values are 95% or greater.

The UPGMA cluster analysis could not group the genotypes according to their genome composition (AA, AAA, AAB and ABB) or geographical origin. The genotypes were mixed up in all the clusters except for few cases where clusters were dominated by genotypes of a particular genome or geographical origin. Lack of clear clustering based on geographical origins among the banana cultivars can be due to introductions followed by naturalization of cultivars in areas away from their initial origins and also signify that a common genetic basis characterizes these cultivars despite their phenotypic and geographical isolation (Zhao et al., 2007; Jbir et al., 2008). Similar results where relationship between geographical origin and genetic similarity could not be clearly demonstrated have been reported in other crops such as mulberry (*Morus* L.), teasel gourd (*Momordica dioica* Roxb ex Willd) and Tunisian pomegranate (*Punica granatum L.*) (Sharma et al., 2000; Vijayan et al., 2004; Rasul et al., 2007; Jbir et al., 2008).

Clustering of crop species is based normally on common origins of cultivars or shared mutations (Dantas et al., 1997). High level of genetic similarity is expected among cultivars from the same subcluster because of common ancestry or mutations (Creste et al., 2003). Among vegetative propagated crops like bananas, variations within each cluster is mainly dependent on differences in genotype and genome differences arising from mutations whose frequency is dependent on how often a clone has been multiplied and planted (Jenny et al., 1999). The exact mechanisms of somatic mutations that result in genetic diversity in bananas remain unknown (Nsabimana and Staden, 2007). The involvement of retrotransposons in spontaneous mutations has not been ruled out because retrotransposons have been reported in bananas (Pillay et al., 2003; Ashalatha et al., 2005). Retrotransposons are abundant and comprise over 50% of nuclear DNA content in many species (Ashalatha et al., 2005). They are known to insert themselves into the genome and act as mutagenic agents, thereby providing a potential source of gene diversity through mutation (Heslop-Harrison, 2000).

Cluster analyses further showed that most genotypes were different from each other implying high degree of polymorphism and dissimilarity. This denotes that the different vernacular names of bananas reflect genetic differences among the cultivars. However, a few cultivars with the same genome composition clustered at 100% similarity (Table 6). Some of the cultivars that showed 100% similarity (Table 6) and clustered together within a subcluster had obvious differences in their morphology. For example, 'Khazanga wamkulu obereka kwambiri'(MJ, AAB group) (French plantain with a big bunch of many hands) and 'Khazanga wamkulu obereka ziwiri or zitatu'(MJ and AAB group) (French plantain whose bunch has three or four fingers) (Figure 3), represent similar local names with similar genome composition and 100% similarity, however, show morphological differences to group them into separate groups. This indicates that the environment and domestication plays a role in morphological differences among different cultivars available in Malawi.

Even though the genetic similarity revealed by SSR markers in polyploid heterozygous species like bananas is known to have limited level of correspondence with real genetic similarity due to loss of co-dominance of SSR markers in polyploids which result in either overestimation or underestimation of allelic relationship, the occurrence of multiple alleles or duplicated chromosomal regions makes SSR markers powerful tools for fingerprinting and discrimination of different genotypes (Crest et al., 2003). The SSR primers used in the present study also clustered some genotypes of the same genome, similar morphological and fruit quality traits and are possible synonymous cultivars such as 'Nakashuga' (GB) and 'Ngelezi'(GB), 'Bondo' (TO) and 'William'(TO), 'Zeru' (NB) and 'Zeru obereka kwambiri' (NB), 'Phwiza wang'ono' (MJ) and 'Phwiza' (MJ), 'Khazanga wamkulu obereka kwambiri'(MJ) and 'Khazanga wamkulu obereka ziwiri or zitatu' (MJ) (Table 6). Molecular markers such as RAPD and SSR (Ulanovsky *et al.,* 2002) and SSR (Martínez *et al*., 2006) have been used elsewhere to detect synonymies in other crops such as grapevines cultivars (*Vitis vinifera* L.).

Some cultivars in the red green group (AAA) which had similar morphological and fruit quality characteristics did not cluster closely and were shown to be genetically distant. These cultivars were 'Palamala umweru' (CP),

'Palamala mukesamu' (CP), 'Palamala muswepo' (KA), 'Mzungu mtuwa' (NB), 'Dochi la chizungu' (NB), 'Palamala' (KA), 'Pamala' (KA), 'Zomba' (TO), 'Kolinto' (TO), 'Zomba red' (LC,GB) and 'Zomba green' (LC, GB). Members of the red green group are known to be soma clones resulting from somatic mutation. The genetic variations among members of the red green group could be due to multiplicity of mutations in each clone as it is locally multiplied and planted by farmers. Somatic mutations resulting in somaclonal variants are known to occur even within a cultivar due to genetic changes that occur during tissue culture process. Ramage et al. (2006) reported such somaclonal variants in Williams (*Musa L.* AAA; subgroup Cavendish). Thomas et al. (2006) showed that morphologically, field grown somatic tea plants had no identical character aligning with the parent plants. Using ISSR markers data and Pearson's coefficient correlation, only 9.2% of the soma clones exhibited significant similarity at genetic level indicating the existence of wide genetic variation among soma clones.

Establishment of genetic relationships among cultivars especially between diploids and triploids is useful in banana breeding programmes, since most programmes aim at developing superior tetraploids derived from crosses between commercial triploid cultivars and diploid genotypes with favourable traits like disease resistance (Crouch et al., 1999). Identifying the most similar diploid genotypes to triploid cultivars based on genetic relationship revealed by molecular markers could allow re-synthesis of triploids by chromosome duplication of diploids followed by crossing with other diploids (Jenny et al., 1999; Careel et al., 2002).

Conclusions

Musa L. cultivars genetic diversity, mostly of the north, should be conserved *in situ* and *ex situ*, and serve as a source of desirable genes in either conventional banana hybridization or genetic transformation programmes. Breeding programmes can benefit from more genotypes of bananas of the north than the south.

The Bvumbwe Agricultural Research Station (BRS) banana field gene bank should be augmented with more wide-ranging banana cultivars from the north and south so that its collection is more genetically diverse and representative since this study has shown that diversity of cultivars names mirror genetic diversity.

ACKNOWLEDGEMENT

The study was funded by the International Foundation for Science (IFS). We are sincerely grateful for the funding. We would like also convey our gratitude to technicians and banana farmers who provided the banana leaf samples analyzed in the study.

REFERENCES

Arias D, Dankers C, Liu P, Pilkauskas P (2003). The World Banana Economy 1985-2002. Food and Agriculture Organisation of the United Nations, Rome, Italy.

Ashalatha SN, Teo CH, Schwarzacher T, Pat HH (2005). Genome classification of banana cultivars from South India using IRAP markers. Euphytica, 144: 285-290.

Banda DLN, Mwenebanda BML (2000). Evaluation of local banana and plantins cultivars for yield potential in Malawi. In: Craenen K, Ortiz R, Karamura EB, Vuylsteke DR (eds) Proceedings of the First International Conference on Banana and Plantain For Africa, pp. 107-108.

Botstein 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.

Buhariwalla HK, Jarret RL, Jayashree B, Crouch JH, Ortiz R (2005). Isolation and characterisation of microsatellite markers from *Musa balbisiana*. Mol. Eco. Not. 5:327-330.

Chizala, C.T., Braunworth, W.S., Banda, D. L. N. (1995). Banana and plantain cultivar diversity in Malawi. Afr. Crop Sci. Conference Proc. 2: 57-60.

Carreel F, Leon DG, Lagoda P, Lanaud C, Jenny C, Horry JP, Montcel, HT (2002). Ascertaining maternal and parental lienage within *Musa* by chloroplast and mitochodrial DNA RFLP analyses. Genome, 45: 679-692.

Creste S, Neto AT, de Oliveira Silva S, Figueira A (2003). Genetic characterisation of banana cultivars (*Musa* spp.) from Brazil using microsatellite markers. Euphytica, 132: 259-268.

Creste S, Neto AT, Vencovsky R, Silva SO, Figueira A (2004). Genetic diversity of *Musa* diploid and triploid accessions from the Brazilian banana breeding programme estimated by microsatellite markers. Genet. Res. Crop Evol. 51: 723-733.

Crouch HK, Crouch JH, Jarret RL, Cregan PB, Ortiz R (1998) Segregation at microsatellite loci in haploid and diploid gametes of *Musa*. Crop Sci. 38: 211-217.

Crouch JH, Crouch HK, Tenkouano A, Ortiz R (1999).VNTR-based diversity analyses of 2 x and 4x full-sib Musa hybrids. EJB Electr. J. Biotechnol. 15: 99-108.

Dantas JLL, Shepherd K, Silva SO, Soares-Filho WS (1997). Classificacao botanica origem, evolucao e distribuicao geografica. In: Alves EJ (ed) A Cultura da Banana: aspectos tecnicos, socioeconomicose agroindustriaia. Brasilia Embrapa, pp. 27-34.

De Vries AC, Ferwerda JD, Flach M (1967). Choice of food crops in relation to actual and potential production in the tropics. Netherlands J. Agric. Sci. 15: 241-248.

FAO (2000). Consumption of bananas and plantains. FAO, Rome, Italy.

FAO Statistical database, FAOSTAT (2007) http://faostat.fao.org/.

FAO Statistical database, FAOSTAT (2008) http://faostat.fao.org/.

Felsenstein J (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39: 783-791.

Gawal NJ, Jarret RL (1991). Cytoplasmic genetic diversity in banana and plantain. Euphytica 52: 19-23.

Ge XJ, Liu MH, Wang WK, Schaal BA, Chiang TY (2005). Population structure of wild bananas, *Musa balbisiana*, in China determined by SSR fingerprinting and cpDNA-RFLP. Mol. Ecol. 14: 933-944.

GraphPad Prism Version 3.00 (1999). GraphPad Software Incorporated, San Diego, California, USA, www.graphpad.com

Heslop-Harrison JS (2000). Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. Plant Cell, 12: 617-636.

INIBAP (2003). Banana-Food and wealth, Http://www.inibap.org.

Jbir R, Hasnaoui N, Mars M, Marrakchi M, Triff M (2008). Characterisation of Tunisian pomegranate (*Punica granatum L.*) cultivars using amplified fragment length polymorphism analysis. Scientia Horticulturae, 115: 231-237.

Jenny C, Carreel F, Tomekpe K, Perrier X, Dubois C, Horry JP, Montcel HT (1999). Les bananiers. In: Centre de Cooperation Internationale en Recherche Agronomique pour le Developpment. Diversite Genetique des plantes Tropicales, CIRAD, Montpellier, France. pp. 113-129.

Jones DR (2000). Introduction to banana, abaca and enset. In: Jones

DR (ed) Diseases of Banana, Abaca and Enset, CAB International, Wallingford, UK.
Karamura DA (1998). Numerical taxonomic studies of the Eastern African highland bananas (*Musa* AAA-East Africa) in Uganda, INIBAP, France.
Kaemmer D, Fisher D, Jarret RL, Baurens FC, Grapin A, Dambier D, Noyer JL, Lanaud C, Khal G, Lagoda PJL (1997) Molecular breeding in the genus *Musa*: a strong case for STMS marker technology. Euphytica, 96: 49-63.
Kimura M, Crow JK (1964). The number of alleles that can be maintained in a finite population. Genetics, 49: 725-738.
Kumar V, Sharma S, Kero S, Sharma S, Sharma AK, Kumar M, Bhat KV (2008). Assessment of genetic diversity in common bean (*Phaseolus vulagris* L.) germplasm using amplified fragment length polymorphism (AFLP). Scientia Horticulturae 116: 138-143.
Laisnez, I. (2005). *Musa spp*. in Malawi: Morphological characterisation of local varieties. MSc. Thesis. Departement Biotechnologische Wetenschappen, Landschapsbeheer en Landbouw, Gent, Belgium.
Lagoda PJL, Noyer JL, Dambier D, Baurens FC, Grapin A, Lanaud C (1998) Sequence tagged microsatellite site (STMS) markers in *Musaceae*. Mol. Ecol. 7:657-666.
Lewontin RC (1974). The genetic basis of evolutionary change. Columbia University Press, New York.
Lian CL, Oishi R, Miyashita N, Nara K, Nakaya H, Wu B, Zhou ZH, Hogetsu T (2003). Genetic structure and reproductive dynamics of *Salix reinii* during primary sucession of Mount Fuji, as revealed by nuclear and chloroplast microsatellite analysis. Mol. Ecol. 12: 609-618.
Malawi Government (1995). Agricultural Research Master Plan. Department of Agricultural Research, Ministry of Agriculture, Lilongwe.
Martínez LE, Cavagnaro PF, Masuelli RW, Zůňiga M (2006). SSR-based assessment of genetic diversity in South American *Vitis vinifera* varieties. Plant Sci. 170: 1036-1044.
Mengoni A, Gori A, Bazzicalupo M (2000). Use of RAPD and microsatellite (SSR) variation to assess genetic relationship among populations of tetraploid alfafa, *Medicago sativa*. Plant Breed, 119: 311-317.
Nei M (1973). Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70: 3321-3323.
Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonuclease. Proc. Natl. Acad. Sci. USA. 5269-5273.
Ning SP, Xu LB, Lu Y, Huang BZ, Ge XJ (2007). Genome composition and genetic diversity of *Musa* germplasm from China revealed by PCR-RFLP and SSR markers. Scientia Horticulturae, 114: 281-288.
Nsabimana A, Van Staden J (2007). Assessment of genetic diversity of highland bananas from the National Banana Germaplasm Collection at Rubona, Rwanda using RAPD markers. Scientia Horticulturae, 113: 293-299.
Oriero CE, Odunola OA, Lokko Y, Ingelbrecht I (2006). Analysis of B-genome derived simple sequences repeat (SSR) marker in *Musa* spp. Afr. J. Biotechnol. 5 (2): 126-128.
Ortiz R (1995). *Musa* genetics. In: Gowen S (ed) Banana and plantains, Chapman and Hall, London, UK. pp. 84-109
Ortiz R (1997). Morphological variation in *Musa* germplasm. Genet. Res. Crop Evol. 44: 393-404.
Ortiz R, Vuylsteke D (1996). Recent advances in *Musa* genetics, breeding and biotechnology. Plant Breed. Abstr,66: 1355-1363.
Pollefeys, P., Sharrock, S, Arnaud, E. (2004). Preliminary analysis of the literature on the distribution of wild Musa species using MGIS and DIVA-GIS Montpellier, France, INIBAP.
Pillay M, Dimkpa C, Ude G, Makumbi D, Tushemereirwe W (2003). Genetic variation in banana cultivar 'Sukali ndizi' grown in different regions of Uganda. Afr. Crop Sci. J. 11: 87-95.
Pillay M, Ogundiwin E, Nwakanma DC, Ude G (2001). Analysis of genetic diversity and relationships in East African banana germplasm. Theor. Appl.Genet.102: 965-970.
Pillay M, Tenkouano A, Ude G, Ortiz R (2004). Molecular characterisation of genomes in *Musa* and its applications. In: Jain SM, Swennen R (eds) Banana Improvement: Cellular, Molecular Biology, and Induced Mutations. Science Publishers, Enfied, NH, USA, pp. 271-286.
Ramage CM, Borda AM, Hamill SD, Smith MK (2006). A simplified PCR test for early detection of dwarf off-types in micropropagated Cavendish bananas (*Musa spp*. AAA). Scientia Horticulturae, 103: 145-151.
Rasul MG, Hiramatsu M, Okubo H (2007). Genetic relatedness (diversity) and cultivar identification by randomly amplified polymorphic DNA (RAPD) markers in teasle gourd (*Momordica dioica* Roxb). Scientia Horticulturae, 111: 271-279.
Rholf JR (2001). NTSYSpc Version 2.11c Numerical Taxonomy and Multivariate Analysis System, Exeter Software, New York.
Robinson JC (1996). Bananas and plantains. Wallingford, UK, CAB International, p. 233.
Saal B, Wricke G (1999). Development of simple sequence repeat markers in rye (*Secale cereale* L.) Genome, 42: 964-972.
Sharma A, Sharma R, Machii H (2000). Assessment of genetic diversity in a Morus germplasm collection using fluorescence-based AFLP markers. Theor. Appl. Genet. 101: 1049-1055.
Simmonds NW (1966). Bananas. 2nd ed., Tropical Agricultural Series, Longman, New York.
Sneath PHA, Sokal RR (1973). Numerical taxonomy. Freeman, San Francisco,p. 573.
Sokal RR, Michener CD (1958). A statistical method for evaluating systematic relationships. Univ. Kansas Sci. Bull. 38: 1409-1438.
Stover RH, Simmonds NW (1962). Bananas. 3rd ed. Longman, New York, USA.
Stover RH, Simmonds NW (1987). Bananas. Longmans, London, UK.
Strosse H, Schoofs H, Panis B, Andre E, Reyniers K, Swennen R (2006). Development of embroyogenic cell suspensions from shoots meristematic tissue in bananas and plantains (*Musa spp*). Plant Sci. 170:104-112.
Thomas J, Vijayan D, Joshi SD, Joseph-Lopez S. Raj-Kumar R (2006). Genetic integrity of somaclonal variants in Tea (*Camellia sinensis* (L) O Kuntze) as revealed by inter simple sequence repeats. J. Biotechnol. 123: 149-154.
Ude G, Pillay M, Nkwakanama D, Tenkouano A (2002). Genetic diversity in *Musa acuminata* Colla and *Musa balbisiana* Colla and some of their natural hybrids using AFLP markers. Theor. Appl. Genet. 104: 1246-1252.
Ulanovsky S, Gogorcena Y, Martinez de Toda F, Ortiz JM (2002). Use of molecular markers in detection of synonymies and homonymies in grapevines (*Vitis viniferaL).*Scientia Horticulturae, 92: 241-254.
Venkatachalam L, Sreedhar RV, Bhagyalakshmi N (2008). The use of genetic markers for detecting DNA polymorphism, genotypes identification and phylogentic relationships among banana cultivars. Mol. Phylogenet. Evol. 47: 974-985.
Vijayan K, Srivastava PP, Awasthi AK (2004). Analysis of phylogenetic relationship among five mulberry (*Morus*) species using molecular markers. Genome, 47 (3): 439-448.
Wan Y, Watanabe JA, Yi SS, Htaik T, Win K, Yamanaka S, Nakamura I, Watanabe KN (2005). Assessment of Genetic diversity among the major Myanmar banana landraces. Breed. Sci. 55: 365-369.
Yamanaka S, Suzuki E, Tanaka M, Takeda Y, Watanabe JA, Watanabe KN (2003). Assessment of cytochrome P450 sequences offers a useful tool for determining genetic diversity in high plants species. Theor. Appl. Genet. 108: 19-24.
Yap IV, Nelson RJ (1996). WinBOOT: A Programme for performing bootstrap analysis of binary Data to determine the confidence limits of UPGMA-based dendrograms. International Rice Research Institute (IRRI), Manila 1099, Philippines,
Yeh FC, Yang R, Boyle T (1999). POPGENE VERSION 1.31, Microsoft Window-based Freeware for Population Genetics Analysis. Quick User Guide, University of Alberta Centre Int. Forest. Res.
Zhao W, Wang Y, Chen T, Jia G, Wang X, Qi J, Pang Y, Wang S, Li Z, Huang Y, Pan Y, Yang YH (2007). Genetic structure of mulberry from different ecotypes revealed by ISSRs in China: An implication for conservation of local mulberry varieties. Scientia Horticulturae, 115: 47-55.
Zizumbo-Villarreal D, Colunga-GarciaMarin P, Payro dela Cruz E, Delgado-Valerio P, Gepts P (2005). Population structure and evolutionary dynamics of wild-weedy-domesticated complexes of common bean in a Mesoamerican region. Crop Sci. 45: 1073-1083.

Supplementary data

Appendix 1. Local names of banana and plantain cultivars collected in this study.

S/N	Local name	Other common name	Genome group	Use	Source/District
1	Bondo		AAA	D	Thyolo, Mulanje
2	Chituwika				Karonga
3	Ghewo/Nghewo		AAB	C	Karonga
4	Gilanya	Similar to plantains	ABB	C	Mulanje
5	Harare	Bluggoe	ABB	C/D	Chitipa, Karonga
6	Harare ya iyolo		ABB	C/D	Karonga
7	India				GBBTO
8	Itoki		AAB	C	Karonga
9	Kabuthu	Dwarf Cavendish	AAA	D	LCBTO, GBBTO, Thyolo
10	Kabuthu wa mkulu		AAA	D	LCBTO, GBBTO, Thyolo, Mulanje
11	Kabuthu wang'ono		AAA	D	Thyolo, Mulanje
12	Kafupi		AAA	D	Nkhata Bay
13	Kalaghasya/Kalanghasha	Kisubeini (Uganda)	AB	D	Karonga
14	Kalota				LCBTO
15	Kaluma	Dwarf Cavendish	AAA	D	Chitipa, Karonga
16	Kaluma Mulanje		AAA	D	Karonga
17	Kambani		AB or ABB or AAB	D	Chitipa, Karonga, Nkhata Bay
18	Kapemba		AAA	D	LCBTO
19	Kapembe		AAA or AA	D	Thyolo, Mulanje
20	Kapeni		AA or AAA	D	Nkhata Bay
21	Katsizi tall		AAA	D	LCBTO
22	Katuma				Karonga
23	Kazanda	Pisang Awak	ABB	C/D	LCBTO, Thyolo
24	Kazingiriya		AAA	D	Nkhata Bay
25	Kazizi		AAA	D	GBBTO, Thyolo
26	Kenya			C/D	Karonga
27	Khazanga	French plantain	AAB	C	GBBTO, Thyolo
28	Khazanga wa mkulu	False horn plantain	AAB	C	LCBTO, GBBTO
29	Khazanga wa mkulu (wobeleka kwambiri)		AAB	C	Mulanje
30	Khazanga wa ng'ono	French plantain	AAB	C	LCBTO, Mulanje
31	Khazanga wamkulu		AAB	C	Mulanje
32	Khazanga wamkulu wobereka ziwiri kapena zitatu		AAB	C	Mulanje
33	Kholobowa	Bluggoe	ABB	C/D	LCBTO, GBBTO, Thyolo, Mulanje
34	Kolinto		AAA	D	Thyolo
35	Lighegho/Lighewo	False horn plantain	AAB	C	Chitipa
36	Makumbuka	Silver Bluggoe	ABB	C/D	GBBTO, Thyolo
37	Matoke		AAB	C	Karonga
38	Mbingindola		ABB	C/D	Nkhata Bay
39	Mbirindola		ABB	C/D	Nkhata Bay
40	Mpama				Nkhata Bay

D, Dessert; C, cooking; LCBTO, local collection Bvumbwe Research Station, Thyolo; GBBTO, Gene Bank Bvumbwe Research Station, Thyolo. Some common names, genome groups and uses adapted from Chizala et al., 1995 and Laisnez, 2005.

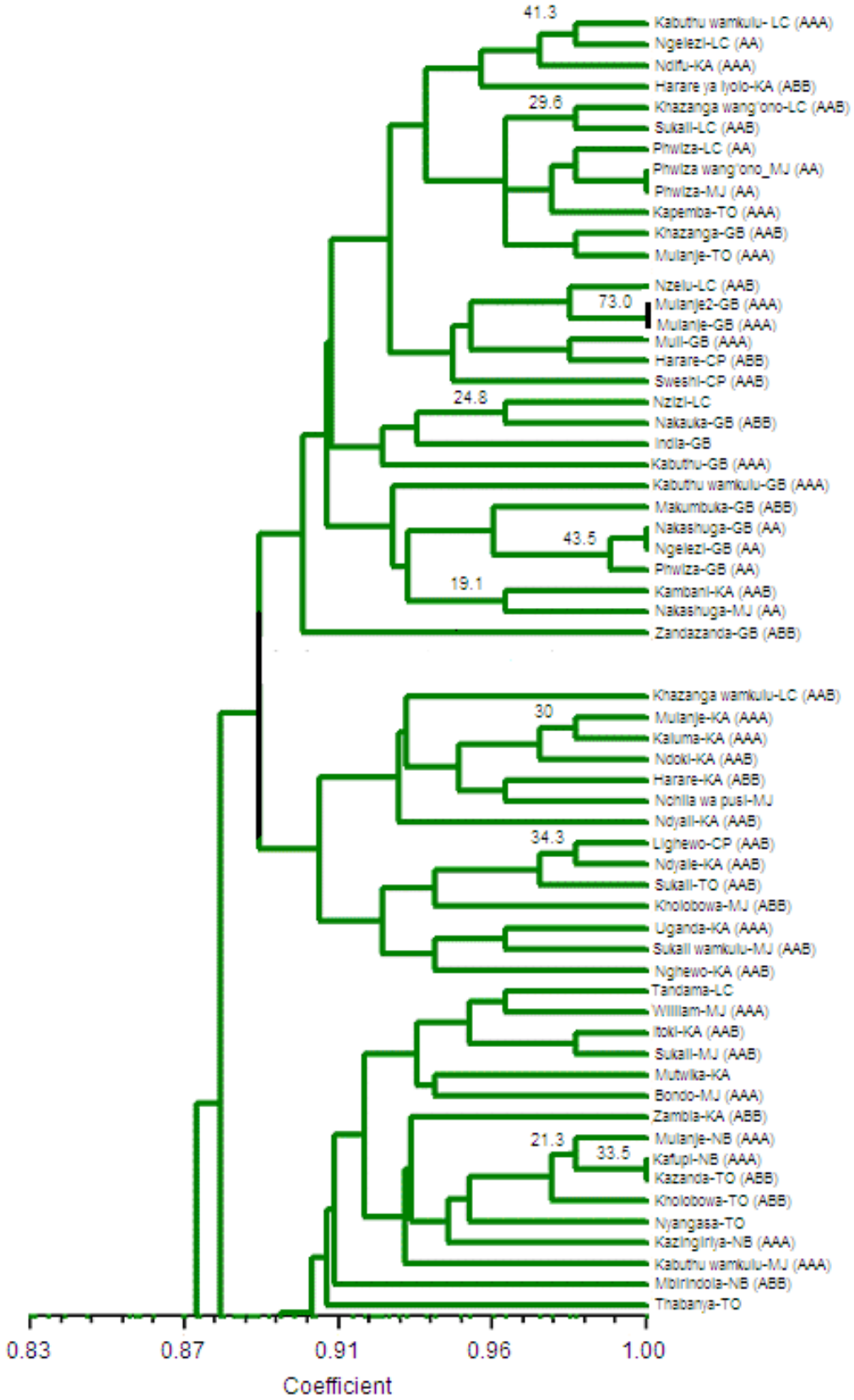
Appendix 1. Contd.

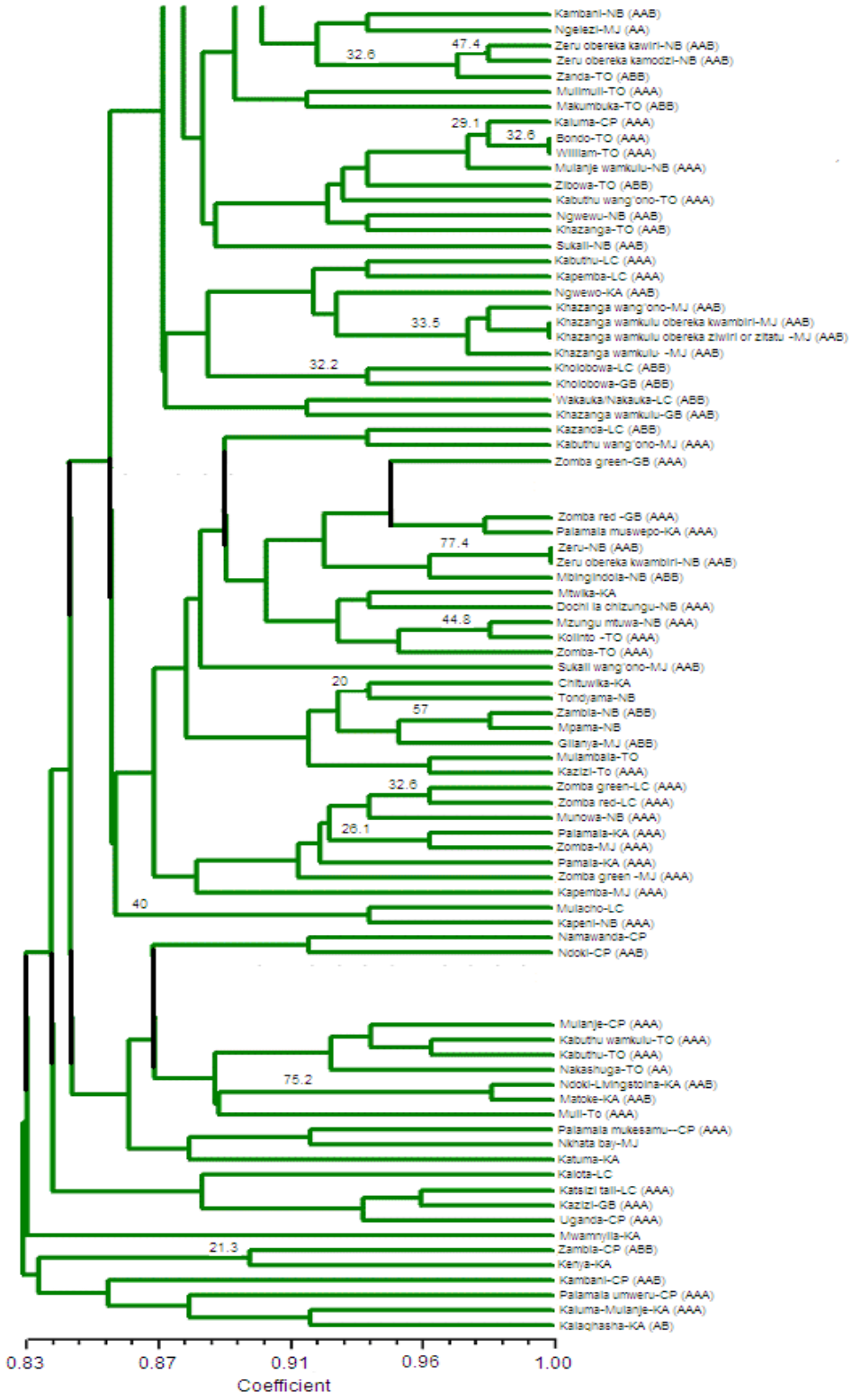
S/N	Local name	Other common name	Genome group	Use	Source/District
41	Mtwika				Karonga
42	Mulacho				LCBTO
43	Mulambala				Thyolo
44	Mulanje	Giant Cavendish	AAA	D	GBBTO, Chitipa, Karonga, Nkhata Bay, Thyolo
45	Mulanje 2	Giant Cavendish	AAA	D	GBBTO
46	Mulanje wa mkulu		AAA	D	Nkhata Bay
47	Muli		AAA	D	GBBTO, Thyolo
48	Muli bwanji		AAA	D	Thyolo
49	Munowa		AAA	D	Nkhata Bay
50	Mutwika				Karonga
51	Mwamnyila				Karonga
52	Mzungu mtuwa		AAA	D	Nkhata Bay
53	Nakashuga		AA	D	GBBTO, Mulanje, Thyolo
54	Nakauka/wakauka	Pisang Awak	ABB	C	LCBTO, GBBTO
55	Namawanda				Chitipa
56	Nchila wa pusi				Mulanje
57	Ndifu		AAA	D	Karonga
58	Dochi la chizungu		AAA	D	Nkhata Bay
59	Ndoki	French plantain	AAB	C	Chitipa, Karonga
60	Ndoki-Livingstonia		AAB	C	Karonga
61	Ndyale		AAB	C/D	Karonga
62	Ndyali	Mchale (Tanzania)	AAB	C/D	Karonga
63	Ngelezi		AA	D	LCBTO, GBBTO, Mulanje
64	Ngewo	False horn plantain	AAB	C	Karonga
65	Ngwewu	False horn plantain	AAB	C	Nkhata Bay
66	Nkhata Bay				Mulanje
67	Nyangasa				Thyolo
68	Nzelu		AAB		LCBTO
69	Nzizi				LCBTO
70	Palamala		AAA	D	Karonga
71	Palamala muswepo	Green red	AAA	D	Karonga
72	Pamala	Red/Giant Red	AAA	D	Karonga
73	Pamala mukesamu	Green red	AAA	D	Chitipa
74	Pamala umweru		AAA	D	Chitipa
75	Phwiza		AA	D	LCBTO, GBBTO, Mulanje
76	Phwiza wang'ono		AAA or AA	C/D	Mulanje
77	Sukali	Silk, Sukari	AAB	D	LCBTO, Nkhata Bay, Thyolo, Mulanje
78	Sukali wamkulu		AAB	D	Mulanje
79	Sukali wang'ono		AAB	D	Mulanje
80	Sweshi		AAB	D	Chitipa

Appendix 1. Contd.

S/N	Local name	Other common name	Genome group	Use	Source/District
81	Tandama				LCBTO
82	Thabanya			D	Thyolo
83	Tondyama		AAA	D	Nkhata Bay
84	Uganda	Lacatan (Jamaican)/ African Highland banana	AAA	D	Chitipa, Karonga
85	William	Giant Cavendish	AAA	D	LCBTO, Thyolo, Mulanje
86	Zambia	Pisang Awak	ABB	C/D	Chitipa, Karonga, Nkhata Bay
87	Zanda	Pisang Awak	ABB	C/D	Thyolo
88	Zandazanda	Pisang Awak	ABB	C/D	GBBTO
89	Zeru	French plantain	AAB	C	Nkhata Bay
90	Zeru obereka kamodzi	French plantain	AAB	C	Nkhata Bay
91	Zeru wobereka kawiri	French plantain	AAB	C	Nkhata Bay
92	Zeru wobereka kwambiri	French plantain	AAB	C	Nkhata Bay
93	Zibowa	Bluggoe	ABB	C/D	Thyolo
94	Zomba	Green red	AAA	D	Thyolo, Mulanje
95	Zomba green	Green red	AAA	D	LCBTO, GBBTO, Mulanje
96	Zomba red	Green red	AAA	D	LCBTO, GBBTO

Appendix 2.


41.3
Kabuthu wamkulu- LC (AAA)
Ngelezi-LC (AA)
Ndifu-KA (AAA)
Harare ya lyolo-KA (ABB)
29.6
Khazanga wang'ono-LC (AAB)
Sukali-LC (AAB)
Phwiza-LC (AA)
Phwiza wang'ono_MJ (AA)
Phwiza-MJ (AA)
Kapemba-TO (AAA)
Khazanga-GB (AAB)
Mulanje-TO (AAA)
Nzelu-LC (AAB)
73.0
Mulanje2-GB (AAA)
Mulanje-GB (AAA)
Mull-GB (AAA)
Harare-CP (ABB)
Sweshi-CP (AAB)
24.8
Nzizi-LC
Nakauka-GB (ABB)
India-GB
Kabuthu-GB (AAA)
Kabuthu wamkulu-GB (AAA)
Makumbuka-GB (ABB)
Nakashuga-GB (AA)
43.5
Ngelezi-GB (AA)
Phwiza-GB (AA)
19.1
Kambani-KA (AAB)
Nakashuga-MJ (AA)
Zandazanda-GB (ABB)
Khazanga wamkulu-LC (AAB)
30
Mulanje-KA (AAA)
Kaluma-KA (AAA)
Ndoki-KA (AAB)
Harare-KA (ABB)
Nchila wa pusi-MJ
Ndyali-KA (AAB)
34.3
Lighewo-CP (AAB)
Ndyale-KA (AAB)
Sukali-TO (AAB)
Kholobowa-MJ (ABB)
Uganda-KA (AAA)
Sukali wamkulu-MJ (AAB)
Nghewo-KA (AAB)
Tandama-LC
William-MJ (AAA)
Itoki-KA (AAB)
Sukali-MJ (AAB)
Mutwika-KA
Bondo-MJ (AAA)
Zambia-KA (ABB)
21.3
Mulanje-NB (AAA)
33.5
Kafupi-NB (AAA)
Kazanda-TO (ABB)
Kholobowa-TO (ABB)
Nyangasa-TO
Kazingiriya-NB (AAA)
Kabuthu wamkulu-MJ (AAA)
Mbirindola-NB (ABB)
Thabanya-TO
0.83
0.87
0.91
0.96
1.00
Coefficient



Appendix 2. Genetic relationships among all banana cultivars in study. Number at node or side indicate bootstrap values in percentages.