Australian Journal of Crop Science AJCS
AJCS 8(1):35-43 (2014) ISSN:1835-2707

# Genetic diversity in *ex-situ* conserved sorghum accessions of Botswana as estimated by microsatellite markers

**Tiny Motlhaodi[1, 2,*], Mulatu Geleta[1], Tomas Bryngelsson[1], Moneim Fatih[3], Stephen Chite[2], Rodomiro Ortiz[1]**

**[1]Swedish University of Agricultural Sciences, Department of Plant Breeding, P. O. Box 101, SE-230 53 Alnarp, Sweden**
**[2]Department of Agricultural Research, Private Bag 0033, Gaborone, Botswana**
**[3]Swedish Biodiversity Centre, P. O. Box 91, Pomonavägen 2, SE-230 53 Alnarp, Sweden**

**[*]Corresponding author: tinymphom@gmail.com**

**Abstract**

Simple sequence repeat (SSR) markers were used to characterize genetic diversity in 30 sorghum [*Sorghum bicolor* (L.) Moench] accessions conserved in the national gene bank of Botswana. This collection represents the three sorghum-growing agro-ecological zones in the country. Ten microsatellite primer pairs were used, and these generated a total of 53 alleles with three to seventeen alleles per locus across the 30 accessions. A high level of genetic variation was observed among the accessions (70% of the variation), and 30% of the total genetic variation was observed within accessions. The analysis of molecular variance on sorghum accessions grouped based on agro-ecological region, ethnicity, and sorghum race revealed a non-significant variation, indicating that the alleles are distributed all over the country. Cluster analysis also did not group the accessions according to agro-ecological region, race or ethnicity. The results of this study have indicated that substantial genetic diversity exists among sorghum accessions conserved in Botswana, but that agro-ecological region, ethnicity and race did not influence the distribution of this diversity.

**Keywords:** accessions**;** genetic diversity; simple sequence repeats; sorghum.
**Abbreviations:** AA_Among accessions; AAWG_Among accessions within groups; AG_Among groups; AMOVA_Analysis of molecular variance; He_expected heterozygosity; Hloci_Nei's gene diversity; Ho_observed heterozygosity; I_Shannon's information index; Na_number of alleles; Ne_effective number of alleles; NPGRC_National Plant Genetic Resources Centre; OFS_Observed fragment size; PCoA_Principal co-ordinate analysis; Pl_Polymorphic loci; RAPD_Random amplified polymorphic DNA; SSR_ Simple sequence repeats; UPGMA_Unweighted pair group with arithmetic average; WA_Within accessions.

## Introduction

Sorghum (*Sorghum bicolor* L. Moench.) is an important cereal, particularly in the world's semi-arid tropics. It is a major food crop in sub-Saharan Africa and South Asia and is the staple food for the most food insecure people in the world (Bibi et al., 2010). It is the world's fifth most important cereal, after wheat, rice, maize and barley (FAO, 2004; Kumar et al., 2011). The crop is predominantly autogamous and has a chromosome number of $2n = 2x = 20$, a nuclear DNA content of 1.6 pg (Thudi and Fakrudin, 2011) and a genome size of 735 Mbp (Dillon et al., 2007). More than 105 countries in Africa, Asia, Oceania and the Americas grow sorghum on 40 million hectares (Kumar et al., 2011), and 60% of this land is in Africa (FAO, 2008), where it continues to play an important food security role (Mutegi et al., 2011). In Africa, grain sorghum is mainly used as a food crop. The sweet sorghum stalk is commonly consumed like sugar cane, and some farmers use it to make good quality silage for feeding livestock, especially dairy cattle. In the American continent the grain is commonly used to make animal feed. In semi-arid and arid countries like Botswana, the sorghum stover is an important dry season livestock feed. Once the grain has been harvested livestock graze the stalks or they are cut and chopped to feed the animals. The centre of origin for sorghum seems to be the northeastern quadrant of Africa (Doggett, 1988) and it has spread to other parts of the world over centuries. In southern Africa sorghum cultivation could have started more than 1000 years ago (Nkongolo and Nsapato, 2003) though it is unclear when it was introduced to Botswana. Over the years, maize has replaced sorghum as staple food in some countries in southern Africa, but in Botswana sorghum still serves as the country's dietary staple and is widely produced by smallholders and commercial farmers. Besides its use as one of the country's important food crops, it is also used for forage as hay or silage, and for production of local beer. There are four races of cultivated sorghum: durra, bicolor, guinea and kafir. All four races are cultivated in different agro-ecological regions in Botswana. A significant number of accessions of the races are held in the national gene bank for *ex-situ* conservation. This collection was assembled mainly to make available a broad genetic base for crop improvement for food and agriculture. Hence a diversity study of sorghums in the national gene bank is important for designing appropriate germplasm conservation and utilization programs. Diversity studies are important in conservation programs for crop genetic resources (Kumar et al., 2011). The use of molecular markers has proven to be a robust technology for assessment of genetic diversity. These are used to assess questions of identity in order to identify putative duplicate accessions, and also to establish relationships and structure with the main aim of determining

how variation is distributed among individuals (Westman and Kresovich, 1997). DNA markers have an apparent advantage over morphological markers as cultivar descriptors since they are not affected by environmental factors (Bowditch et al., 1993). Information obtained is then used to design germplasm management strategies (Dje et al., 1999). Because of the importance of sorghum in Africa, there has been considerable interest in characterizing its genetic diversity. Several studies of sorghum genetic diversity in Africa have been reported for different countries (Dje et al., 2000 Ghebru et al, 2002; Nkongolo and Nsapato, 2003; Uptmoor et al., 2003; Casa et al., 2005; Mutegi et al., 2011; Ng'uni et al., 2011), but not for Botswana. This research has demonstrated that substantial genetic diversity exists between and among accessions in African gene banks, and diversity often differs among geographic regions within countries. In this study, simple sequence repeats (SSR), also known as microsatellites were used to estimate the level of genetic diversity in sorghum accessions conserved *ex-situ* in the national gene bank of Botswana. Microsatellites have been found to be more reliable and useful than other markers such as random amplified polymorphic DNA (RAPD; Vierling et al., 1994). In addition, microsatellites are highly polymorphic (Uptmoor et al., 2003) even among closely related cultivars, which demonstrates that they are highly informative. These markers are co-dominant and can be analyzed by a rapid, technically simple, specific and inexpensive polymerase chain reaction (PCR) based assay that requires only small amounts of DNA. Fourthly, they are abundant (Varshney et al., 2005) and uniformly dispersed in the plant genome. A genetic diversity study of local landraces in Botswana is important as the sorghum accessions conserved in the gene bank have not been characterized before. The study can indicate which landraces are suitable for rescue, agronomic evaluation and for short- and long-term breeding strategies in the country. The objectives of the present study were (1) to characterize the genetic diversity of sorghum accessions that have been conserved *ex-situ* at the National Plant Genetic Resource Centre (NPGRC) in Botswana, (2) to assess the pattern of the accessions' genetic diversity in relation to their racial classification and agro-ecological zones as well as the ethnicity of the local population, and (3) to evaluate the implications of the existing genetic diversity for conservation and breeding.

**Results**

***The SSR loci***

All the SSR loci investigated in this study were polymorphic, and a total of 53 alleles were obtained across the ten loci. The number of alleles observed per locus ranged from 2 (*Sb 6-34*) to 15 (*Xtxp 285*) with an average of 5.3 alleles per locus (Table 2). The observed heterozygosity ranged from 0 (Sb AGB 02)) to 0.48 (Xcup 49) with an average of 0.09 (Table 2).

***Heterozygosity and allelic diversity in sorghum accessions***

Genetic diversity analysis within 30 sorghum accessions revealed that the number of polymorphic loci (Pl) ranged from 0 to 8, with an average of 3.3 (Table 1). Twelve out of the 30 accessions had above average Pl, while eighteen were below average. Accession 2273 recorded the highest number of polymorphic loci, whereas all loci were monomorphic in accessions 2245 and 2266. The genetic diversity for each population, which is the average gene diversity over all polymorphic loci for each population ($H_{loci}$) ranged from 0 to 0.39 (Table 1). Accession 2211 recorded the highest gene diversity in six of the ten loci. The mean number of alleles (Na) per accession across all loci ranged from 1 (accessions 2245 and 2266) to 2.4 (accession 2300), with an average of 1.6 alleles per accession (Table 1). Eleven accessions had above average number of alleles, 15 accessions below average and four accessions had the average number of alleles. The observed heterozygosity (Ho) for each accession across all loci ranged from 0 to 0.39 with an average of 0.18. Accession 2266 was homozygous across all loci (Table 1) while 15 accessions had above average Ho. The expected heterozygosity (He) across all loci ranged from 0 (accession 2266) to 0.28 (accession 2211) with an overall average value of 0.15 (Table 1). Accession specific rare alleles were detected in 7 of the 30 sorghum accessions studied. Accession 2250 was most unique, as it recorded unique alleles at two different loci: a 128 bp allele at *Sb AGB 02* and a 218 bp allele at *Xtxp 285*. Accession 2250 also shared a 284 bp allele at locus *Sb 1-10*, a 174 bp allele at *Undhsbm 24*, a 165 bp allele at *Xcup 49* and a 215 bp allele at locus *Xtxp 285* only with accession 2273, as these alleles were not detected in other accessions. Accession 2300 had a 288 bp allele at *Sb 1-10* and shared a 103 bp rare allele at *Sb AGB 02* with accession 2309, as well as a 240 bp allele at *Dsenhsbm 99* shared with accession 2250, each with a frequency of 0.1 Another rare allele of 250 bp was detected in accession 2255 at locus *Xtxp 285* with a frequency of 0.1. Two other accessions bearing unique alleles were 2211 (124 bp), and 2225 (121 bp), both at locus *Sb AGB 02*. Accession 2219 also recorded a 220 bp unique allele at locus Xdhsbm 1005. Analysis of molecular variance (AMOVA) of the SSR genetic data for the 30 sorghum accessions revealed significant differentiation among accessions ($F_{ST}$ = 0.7, $P$ < 0.001; Table 3). Genetic differentiation among accessions accounted for 70% of the total variation while the within accession variation accounted for the remaining 30%. AMOVA, however, revealed no significant variation among groups of sorghum accessions grouped according to agro-ecological zones ($P$ = 0.634). The same was true when grouping was done based on ethnicity of the local populations of the collection sites of the accessions (Bakwena, Bakgatla, Barolong, Bangwato, Batawana and Bakalaka) and when the grouping was done according to races of the accessions (Bicolor, Durra, Guinea, Kafir; Table 3).

***Cluster, structure and principal coordinate analysis***

The UPGMA cluster analysis based on Nei's standard genetic distance grouped the 30 sorghum accessions into six clusters with low bootstrap support (Fig. 2). Accession 2250 formed an out group, with a bootstrap support of 100%, indicating that it was the most differentiated accession. The second group also consisted of only one accession, 2225, which was supported by a bootstrap value of 64%. The four remaining clusters had very weak bootstrap support. Cluster 3 contained 13 accessions, the bulk of which (11) came from the hardveld, and two from the sandveld. Ten of the accessions in this cluster were of the kafir race, with two bicolor and one durra race accessions. The fourth cluster consisted of three accessions from the hardveld, two from the sandveld and one from the wet sandveld. Each of the remaining two clusters had four accessions, each with three hardveld accessions and one wet sandveld accession. All clusters had accessions from more than one race. Thus there was no clear clustering

**Table 1. (**1) District, ecological zone and geographical coordinates of collection sites, and (2) number of polymorphic loci (Pl), mean number of alleles (Na), mean observed heterozygosity (Ho) mean expected heterozygosity (He) and gene diversity ($H_{loci}$) for 30 sorghum accessions representing four sorghum races and four agro-ecological zones in Botswana.

| Accession | Race | District | Zone | Latitude | Longitude | Pl | Na | Ho | He | $H_{loci}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2207 | Bicolor | Kweneng | h | 24°28'70"S | 25°53'32"E | 3 | 1.5 | 0.06 | 0.05 | 0.06 |
| 2208 | Bicolor | Kweneng | h | 24°28'70"S | 25°53'32"E | 5 | 1.6 | 0.29 | 0.22 | 0.28 |
| 2211 | Kafir | Kweneng | h | 24°28'70"S | 25°53'32"E | 6 | 1.9 | 0.39 | 0.28 | 0.39 |
| 2212 | Bicolor | Kweneng | h | 24°28'70"S | 25°53'32"E | 2 | 1.5 | 0.27 | 0.21 | 0.17 |
| 2214 | Kafir | Kweneng | h | 24°28'70"S | 25°53'32"E | 1 | 1.4 | 0.10 | 0.09 | 0.03 |
| 2219 | Kafir | Kgatleng | h | 24°15'00"S | 26°30'00"E | 3 | 1.5 | 0.12 | 0.10 | 0.06 |
| 2225 | Kafir | Southern | h | 25°00'00"S | 25°00'00"E | 6 | 1.8 | 0.23 | 0.18 | 0.21 |
| 2226 | Kafir | Southern | h | 25°00'00"S | 25°00'00"E | 2 | 1.4 | 0.11 | 0.10 | 0.05 |
| 2227 | Kafir | Southern | h | 25°00'00"S | 25°00'00"E | 4 | 1.6 | 0.30 | 0.23 | 0.29 |
| 2244 | Durra | Kgatleng | h | 24°15'00"S | 26°30'00"E | 2 | 1.4 | 0.11 | 0.10 | 0.09 |
| 2245 | Durra | Kgatleng | h | 24°15'00"S | 26°30'00"E | 0 | 1.0 | 0.21 | 0.17 | 0.00 |
| 2246 | Kafir | Kgatleng | h | 24°15'00"S | 26°30'00"E | 2 | 1.5 | 0.08 | 0.07 | 0.08 |
| 2247 | Kafir | Kgatleng | h | 24°15'00"S | 26°30'00"E | 2 | 1.6 | 0.17 | 0.14 | 0.10 |
| 2248 | Kafir | Kgatleng | h | 24°15'00"S | 26°30'00"E | 5 | 1.6 | 0.24 | 0.19 | 0.19 |
| 2250 | Durra | Central | h | 23°06'00"S | 26°48'00"E | 5 | 1.9 | 0.25 | 0.20 | 0.21 |
| 2255 | Durra | Central | h | 23°06'00"S | 26°48'00"E | 3 | 1.8 | 0.21 | 0.17 | 0.17 |
| 2259 | Kafir | Central | h | 23°06'00"S | 26°48'00"E | 2 | 1.2 | 0.11 | 0.10 | 0.10 |
| 2264 | Durra | Central | h | 23°06'00"S | 26°48'00"E | 2 | 1.2 | 0.02 | 0.02 | 0.03 |
| 2266 | Durra | Central | h | 23°06'00"S | 26°48'00"E | 0 | 1.0 | 0.00 | 0.00 | 0.00 |
| 2273 | Guinea | Central | h | 22°00'00"S | 26°00'00"E | 8 | 2.3 | 0.27 | 0.21 | 0.21 |
| 2299 | Durra | Central | h | 22°00'00"S | 26°00'00"E | 3 | 1.5 | 0.09 | 0.08 | 0.10 |
| 2300 | Bicolor | Central | h | 22°00'00"S | 26°00'00"E | 7 | 2.4 | 0.36 | 0.27 | 0.27 |
| 2301 | Guinea | North west | w | 19°30'00"S | 23°30'00"E | 2 | 1.3 | 0.11 | 0.10 | 0.11 |
| 2306 | Guinea | North west | w | 19°45'43"S | 23°40'37"E | 3 | 1.8 | 0.16 | 0.13 | 0.23 |
| 2309 | Durra | North west | w | 19°45'43"S | 23°40'37"E | 4 | 1.4 | 0.10 | 0.09 | 0.13 |
| 2313 | Kafir | Kweneng | s | 24°25'00"S | 25°32'00"E | 4 | 1.7 | 0.24 | 0.19 | 0.24 |
| 2314 | Bicolor | Kweneng | s | 24°25'00"S | 25°32'00"E | 6 | 2.0 | 0.31 | 0.24 | 0.26 |
| 2315 | Kafir | Kweneng | s | 24°25'00"S | 25°32'00"E | 2 | 1.7 | 0.20 | 0.17 | 0.11 |
| 2316 | Kafir | Kweneng | s | 24°25'00"S | 25°32'00"E | 5 | 1.9 | 0.31 | 0.24 | 0.27 |
| 2329 | Kafir | Kgatleng | h | 24°15'00"S | 26°30'00"E | 2 | 1.5 | 0.11 | 0.10 | 0.10 |
| Mean | | | | | | 3.3 | 1.6 | 0.18 | 0.15 | 0.15 |
| StDev | | | | | | 1.9 | 0.3 | 0.10 | 0.07 | 0.09 |

Note: h = hardveld, w = wet sandveld, and s = sandveld, StDev = standard deviation

pattern of accessions according to ethnicity, races, or agro-ecological regions. Similar to cluster analysis, PCoA placed accessions 2250 and 2225 as stand-alone groups (Fig. 3). The other groupings were also similar to those revealed by clustering, though with the PCoA cluster 3 was broken into two groups, separating all the kafirs from the Kweneng and Southern districts from the Kgatleng and Central Districts durras and kafirs. Accession 2266 and 2273 were placed under one group in PCoA, and were clustered very closely in the cluster analysis regardless of the fact that they belong to different races.

## Discussion

### *Microsatellites and allelic richness*

The characterization of genetic diversity in crop germplasm is essential for rational utilization and conservation of genetic resources (Thudi and Fakrudin, 2011). The ten microsatellite markers used in this study were able to discriminate between the 30 accessions studied, indicating the value of microsatellites for characterizing genetic diversity among closely related individuals (Uptmoor et al., 2003; N'guni et al., 2011). The allele sizes from each locus obtained in this study had a wide range. For the loci *Sb 1-10*, *Sb 6-34*, *Undhsbm 24*, *Xdhsbm 1005* and *Xtxp 285* the allele size range was in agreement with the range reported in various previous publications on sorghum cultivars and landraces (Brown et al., 1996; Shehzad et al., 2009; N'guni et al., 2011; Reddy et al., 2011). The number of alleles obtained in this study at *Sb 1-10* and *Sb AGB 02* loci was similar to that obtained in previous studies (Dean et al., 1999; N'guni et al., 2011; N'guni et al., 2012). Our study revealed more alleles at the *Xtxp 285* and *Xcup 50* loci and fewer alleles at the *Sb 6-34* and *Xcup 49* loci when compared to the number of alleles reported in N'guni et al. (2011 and 2012). Some population specific alleles were also observed in this study. The total number of alleles was, however, lower than previously noted. For example, Ghebru et al. (2002) detected a total of 208 alleles with 15 microsatellites in 28 accessions of Eritrean landraces, while Folkertsma et al. (2005) reported 123 alleles using 21 microsatellites in 100 sorghum accessions from 10 African countries and India. Likewise, Deu et al. (2008) reported 292 alleles with 28 microsatellites in 484 cultivars and landraces from Niger while Thudi and Fakrudin (2011) noted 228 alleles using 30 microsatellites among 42 sorghum genotypes from India. These studies, however, contained sorghum accessions from different geographic areas and the number of microsatellites used was higher when compared with the current study. Comparable results were obtained by N'guni et al. (2011) who observed a total number of 44 alleles in 27 Zambian sorghum accessions using 10 microsatellites. The low allelic richness in this study could be due to the fact that the collection was done from a narrower geographic area when compared to the other studies. In Botswana, geographical patterns of sorghum distribution do not really exist as farmers exchange seeds all over the country, and they maintain diversity among themselves as a

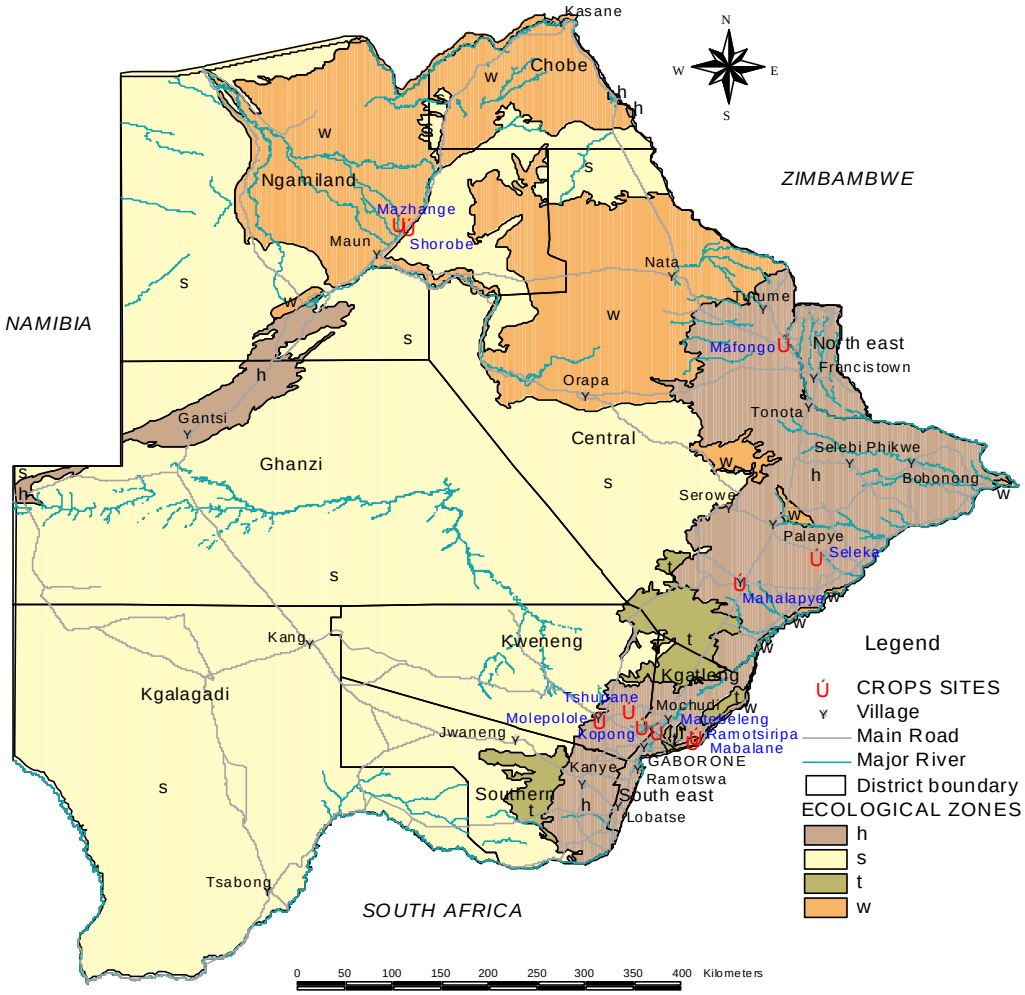


**Fig 1.** Map of Botswana showing collection sites for sorghum accessions used in this study. The sites are indicated by "U" shape red curves. The areas h (hardveld), s (sandveld), t (transition sandveld - hardveld) and w (wet sandveld) are the country's agro-ecological zones based on soil type, vegetation and annual rainfall

risk management strategy to overcome biotic and abiotic stresses, especially drought and infestation by the parasitic weed *Striga asiatica*. The use of sorghum is also quite similar between ethnic groups and geographical areas in the country. The $F_{ST}$ value observed in this study (0.70) was similar to the $F_{ST}$ value of 0.68 reported in Dje et al., (2000) for sorghum accessions from Africa, Asia and America using five microsatellite loci, but higher than the $F_{ST}$ value of 0.50 reported in Ghebru (2002) for 28 sorghum Eritrean landraces. Observed heterozygosity was also quite low in this study (0.09), which could be primarily due to the predominantly inbreeding nature of sorghum. However, Ghebru et al. (2002) reported a higher observed heterozygosity among landraces from Eritrea, in East Africa –one of the centers of origin and diversity for sorghum (Doggett, 1988). According to Ghebru (2002), east African sorghum accessions also have a reduced level of inbreeding, which leads to high levels of heterozygosity. New alleles may be generated because of outcrossing and recombination (Ghebru et al., 2002).

***Genetic variation in sorghum accessions***

In this study, higher genetic variation was observed among accessions than within accessions. During the selection process, farmers select particular landraces; hence there has been no continuous exchange of genes among populations. Previous research showed similar results, with a higher variation among accessions than within accessions. N'guni et al. (2011) observed quite a high variation among accessions (82%) in a study that used 27 Zambian sorghum accessions and 10 microsatellites. Ghebru et al. (2002) also observed significant genetic variation among 28 Eritrean sorghum accessions in which among accession variation accounted for 50.4% and within accessions 49.6% of observed variation. The higher among accession variation could be due to the selection practice of local farmers, where each farmer keeps and maintains more than one landrace for various uses. This practice of separately maintaining several landraces increases the total sorghum genetic diversity within a given geographic area but does not increase the within population genetic diversity, due to the inbreeding nature of sorghum, and thus most of the total genetic variation is observed among populations (N'guni et al., 2011). In Botswana, this practice comes as a result of pressure from the prevailing environmental conditions. Farmers normally select those landraces that can withstand long periods of low moisture conditions and high temperatures, as well as low soil fertility. Quellea bird infestation is also a challenge for farmers; hence

**Table 2.** Microsatellite loci, repeat motif, observed fragment size (OFS), number of alleles (Na), effective number of alleles (Ne), heterozygosity and Shannon's information index (I)

| Microsatellite | Repeat motif | OFS | Na | Ne | *Ho* | *He* | I |
|---|---|---|---|---|---|---|---|
| Dsenhsbm 99 | $(GCA)_6$ | 240-252 | 4 | 1.2 | 0.01 | 0.22 | 0.46 |
| Sb 1-10 | $(AG)_{27}$ | 256-288 | 4 | 1.1 | 0.02 | 0.10 | 0.26 |
| Sb 6-34 | $(AC)/(CG)_{15}$ | 207-208 | 2 | 2.0 | 0.02 | 0.50 | 0.69 |
| Sb AGB 02 | $(AG)_{35}$ | 98-128 | 5 | 1.3 | 0.00 | 0.23 | 0.53 |
| Undhsbm 24 | $(GCA)_{10}$ | 174-183 | 3 | 1.5 | 0.02 | 0.32 | 0.61 |
| Undhsbm 313 | $(TAAA)_{13}$ | 140-146 | 3 | 1.7 | 0.02 | 0.40 | 0.61 |
| Xcup 49 | $(GGAT)_6$ | 152-165 | 4 | 2.7 | 0.48 | 0.64 | 1.12 |
| Xcup 50 | $(ACAGG)_5$ | 146-161 | 4 | 2.8 | 0.03 | 0.65 | 1.12 |
| Xdhsbm 1005 | $(AGATAT)_{16}$ | 177-220 | 9 | 4.0 | 0.20 | 0.71 | 1.62 |
| Xtxp 285 | $(CTT)11CTC(CTT)_{16}$ | 232-270 | 15 | 8.8 | 0.08 | 0.88 | 2.34 |
| Mean | | | 5.3 | 2.7 | 0.09 | 0.46 | 0.93 |
| StDev | | | 3.9 | 2.3 | 0.15 | 0.25 | 0.63 |

StDev = standard deviation

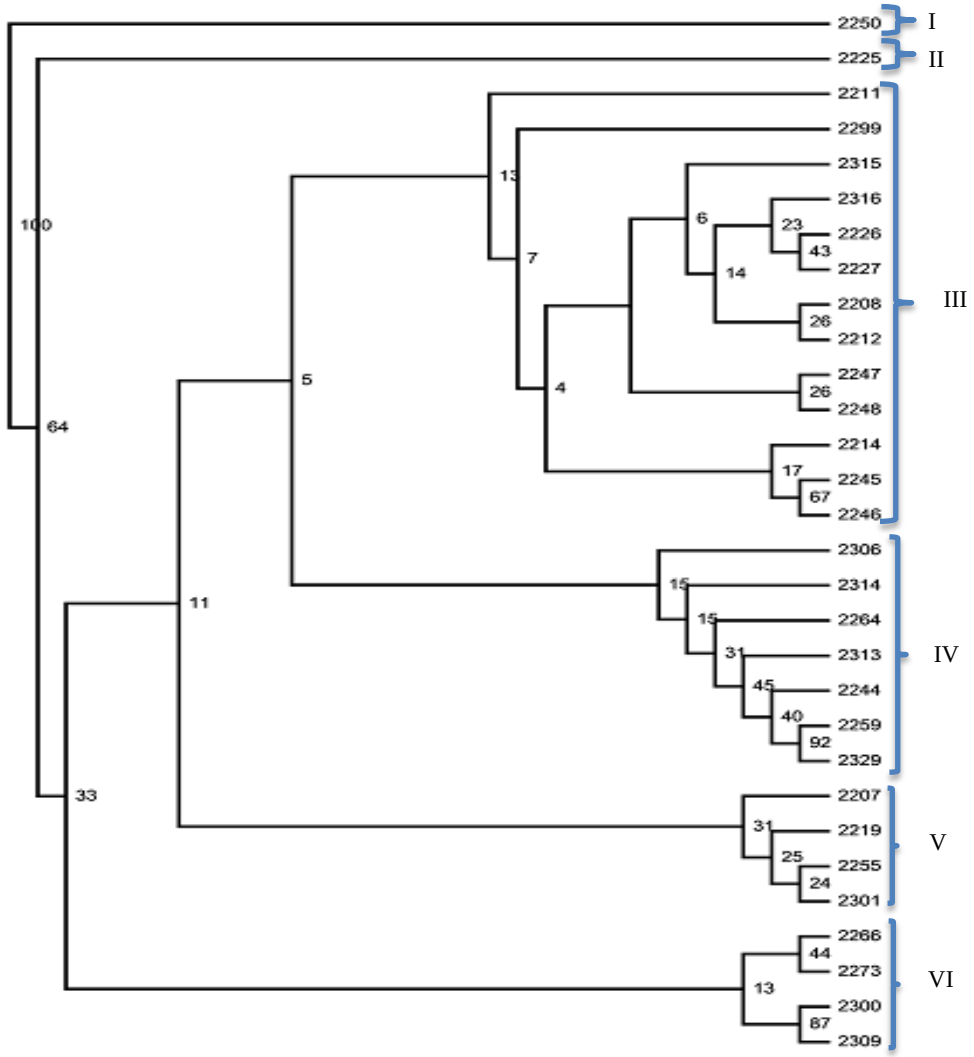


**Fig 2.**. UPGMA dendrogram showing the clustering pattern of 30 sorghum accessions from Botswana based on Nei's standard genetic distance. Bootsrap values generated from 1000 resampling in the FreeTree program are shown between branches.

they tend to select early maturing cultivars, so that their grain will be past the dough stage once the birds migrate into the country. Farmers also use the various landraces for different uses. For instance, the reddish brown grain cultivars, which have high tannin content and hence a bitter taste when served as porridge, are used for making local beer. The white cultivars mostly have desirable taste when cooked; hence they are used for making soft as well as stiff porridge, and they are also cooked as grain and consumed like rice. The tall cultivars with the sweet stalks are used mainly for chewing. Sweet juice is obtained from their stalks just like sugar cane. These cultivars are also used for making silage for livestock feed because of their good fermentation quality, which is due to the high sugar content. They also grow very tall so they yield higher green biomass for silage. Farmers, therefore, need to maintain a range of cultivars for various uses year after year. Variation was not detected among the groups when sorghum accessions were grouped according to

**Table 3.** SSR based analysis of molecular variance for 30 sorghum accessions from Botswana: without grouping the accessions (A), and by grouping the accessions according to agro-ecological zones (B), sorghum races (C) and ethnicity of the local population of the sampling site (D).

| Accessions | Source of variation | d.f. | Variance components | Percentage Variation | Fixation Indices | *P* value |
|---|---|---|---|---|---|---|
| (A) Ungrouped | AA | 29 | Va =1.69 | 69.7 | $F_{ST}$ =0.697 | Va and $F_{ST}$=0.000 |
| | WA | 690 | Vb = 0.73 | 30.3 | | |
| | Total | 719 | 2.40 | | | |
| (B) Ecological zones | AG | 2 | Va = 0.05 | -1.2 | $F_{ST}$ = 0.695 | Vc and $F_{ST}$=0.000 |
| | AAWG | 27 | Vb = 2.03 | 70.7 | $F_{SC}$ = 0.698 | Vb and $F_{SC}$=0.000 |
| | WA | 690 | Vc = 0.83 | 30.5 | $F_{CT}$ = -0.012 | Va and $F_{CT}$=0.634 |
| | Total | 719 | 2.39 | | | |
| (C) Races | AG | 3 | Va =0.02 | 0.97 | $F_{ST}$ = 0.698 | Vc and $F_{ST}$=0.000 |
| | AAWG | 26 | Vb = 1.66 | 68.8 | $F_{SC}$ = 0.695 | Vb and $F_{SC}$=0.000 |
| | WA | 690 | Vc = 0.73 | 30.2 | $F_{CT}$ = 0.100 | Va and $F_{CT}$=0.298 |
| | Total | 719 | 2.41 | | | |
| (D) Ethnicity | AG | 5 | Va =0.06 | 2.53 | $F_{ST}$ = 0.698 | Vc and $F_{ST}$=0.000 |
| | AAWG | 24 | Vb = 1.63 | 67.31 | $F_{SC}$ = 0.690 | Vb and $F_{SC}$=0.000 |
| | WA | 690 | Vc =0.73 | 30.16 | $F_{CT}$ = 0.025 | Va and $F_{CT}$=0.136 |
| | Total | 719 | 2.41 | | | |

*AA = among accessions, WA = within accessions, AG = among groups, AAWG = among accessions within groups

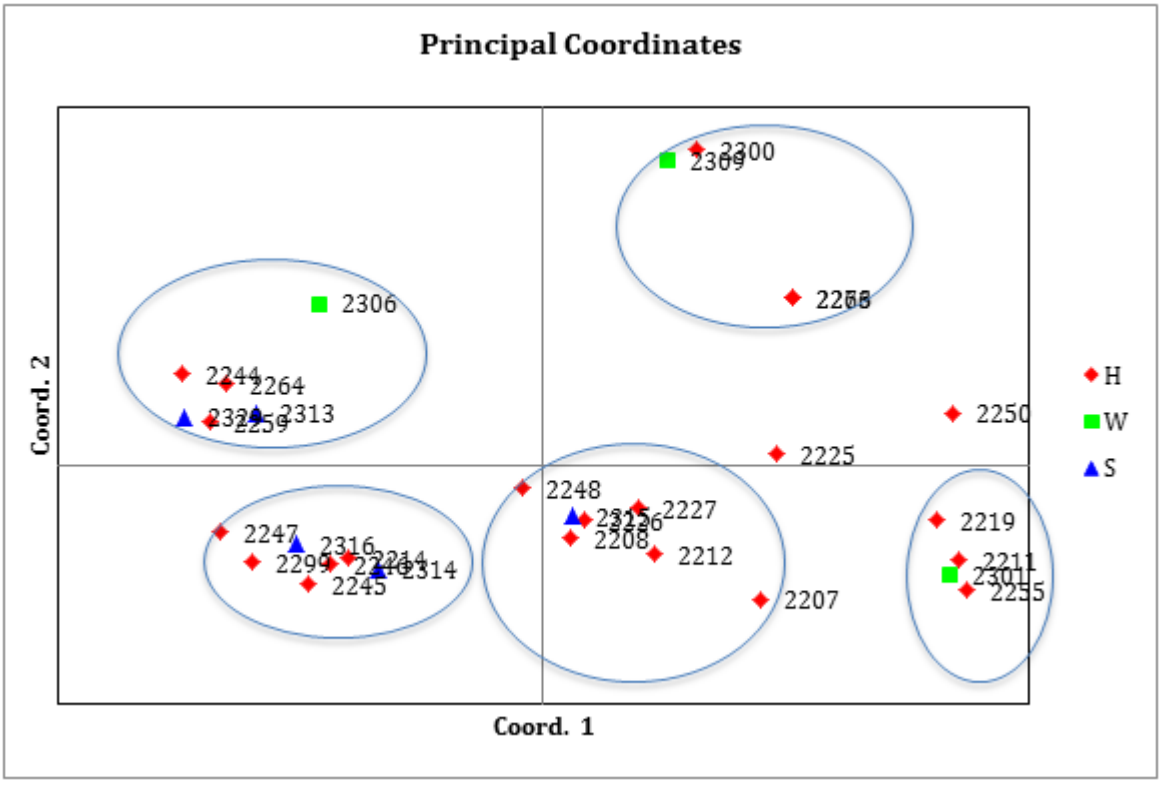


**Fig 3.** Principal co-ordinate analysis (PCoA) using genetic distance matrix of 30 sorghum accessions from Botswana. The 7 clusters were obtained from data using 10 microsatellite markers

ecological zone or ethnicity of farmers. This result indicates that sorghum germplasm in Botswana is freely exchangeable between ethnic groups and across geographical regions.

***Genetic structure in sorghum accessions***

It is interesting to note that cluster analysis did not group the sorghum accessions according to racial classification, ethnicity or agro-ecological regions. A similar observation was made by Dje et al. (2000). They found a scattering of accessions belonging to the same race or geographical region when a matrix plot of individual sorghum accessions based on $R_{ST}$ distance matrix was used. Several other studies have reported weak differentiation among accessions according to geographic region (Ayana et al., 2000; Uptmoor et al., 2003). This observation could be attributed to the practice of seed exchange through traditional and commercial seed systems. Farmers usually exchange seeds in order to access new cultivars with desirable traits. In Botswana, as it is also noted in other locations of sub-Saharan Africa, farmers exchange traditional crops like sorghum following collective social rules that involve relationships between friends, relatives and other social or traditional structures (Deu et al., 2008). N'guni et al. (2011), however, reported that cluster analysis on Zambian sorghum accessions grouped them according to their geographic regions of origin, a finding similar to that reported by Ghebru et al. (2002) for Eritrean sorghum. Among the 30 sorghum accessions studied, the most unique was accession 2250, as it had several rare and unique alleles,

and was grouped separately from all the other accessions. Accession 2273 also had several unique alleles and other accessions bearing at least one unique allele included 2219, 2225, 2255, 2300 and 2316. These accessions need to be further evaluated as they could contain alleles linked to important traits.

## Materials and Methods

### *Plant materials*

Seeds of 30 sorghum accessions were sourced from the National gene bank of Botswana. Passport data (latitude, longitude and racial classification) for the accessions were recorded (Table 1). Before conserving in the gene bank, seed samples were collected from three different sorghum-growing agro-ecological zones in the country (Fig. 1) and consist of four races of sorghum (bicolor, durra, guinea and kafir). During the collecting missions, accessions were systematically sampled to cover different villages in agricultural districts within the agro-ecological zones. Accessions bearing the same latitude and longitude are not duplicates, as they were collected from different fields in the same village.

### *The study area*

Botswana is a semi-arid, landlocked country in southern Africa. It is divided into four agro-ecological zones; hardveld (h), sandveld (s), wet sandveld (w) and the transition sandveld-hardveld (t) (Fig. 1). This zoning is based mainly on types of soil and climate of the areas. The hardveld covers most of the eastern part of the country. The soils in this part are sandy loams, loamy sands and sandy clay loams with low to moderate fertility. The area is characterized by semi-arid conditions, with average rainfall of about 400 to 550 mm, which is very erratic and poorly distributed spatially and temporally. The sandveld is quite arid and is covered by sands with very poor fertility, with an average rainfall of 200 to 400 mm, also quite erratic in nature. The wet sandveld is covered with sandy soils of poor drainage and receives 550 to 600 mm rainfall annually. Arable farming is practiced mostly in the hard veld and the wet sand veld, and sorghum is therefore important in these two zones. Only a few farmers in the sandveld grow crops (including sorghum). Our research gave priority to sorghum accessions collected from villages in the hardveld and the wet sandveld as sorghum is an important food security crop in these two areas.
Six ethnic groups dominantly occupy these villages. On the hardveld are the Barolong (Southern District), the Bakwena (the dominant group in the Kweneng District), the Bakgatla (Kgatleng District), and the Bangwato and Bakalaka (who equally occupy the Central District). On the wet sandveld are found the Batawana in the North West district, in the wet area around the Okavango delta.

### *DNA extraction*

Sorghum seeds were planted in seedling trays and grown in the greenhouse at 25°C for approximately 10 days. Leaves (approximately 6 cm in length) were sampled into 2 ml Eppendorf tubes and frozen in liquid nitrogen, then freeze dried and maintained on silica gel at -80°C until they were milled using a Retsch MM400 shaker (Hann, Germany). DNA was separately extracted from leaves of 12 individual plants per accession using a modified cetyl trimethyl ammonium bromide (CTAB) protocol, as described in Bekele et al. (2007). The quality of the isolated DNA was analyzed using agarose gel electrophoresis and its quantity was determined using Nanodrop ND-1000 spectrophotometer (Saveen Werner, Malmö, Sweden).

### *PCR amplification and microsatellite analysis*

Thirty-five microsatellite primer pairs were initially screened for good amplification and polymorphism. The PCR was performed in 25 µl reaction volumes, containing 10 ng genomic DNA, 1x PCR reaction buffer, 2 mM $MgCl_2$ , 0.2 mM of each dNTPs, 7.5 pmol µl-1 each of the forward and reverse primers, and 0.75 units of Taq DNA polymerase (Saveen Werner AB, Sweden) and made up to 25 µl with Milli Q water. The linkage groups of six of the ten microsatellite loci considered in this study were noted in previous research; Sb 1-10 and Sb 6-34 in chromosome 1 (Brown et al., 1996; Taramino et al., 1997), Xcup 49 and Xcup 50 in chromosome 10 (Schloss et al., 2002), Sb AGB02 in chromosome 7 and Xtxp 285 in chromosome 3 (Menz et al., 2002). The 10 loci were selected from a total of 35 loci that were initially tested, based on specificity of the primer pairs to their target loci, and suitability of the allele size for post-PCR multiplexing. The list of primers used in the study is attached as a supplementary file (Table 4). The forward primers of each pair of selected primers were 5'-labeled with either HEX, 6-FAM, VIC or NED florescent dyes. The reverse primers were PIG-tailed with "GCTTCT" to avoid a non-template addition of a single nucleotide by Taq DNA polymerase to the PCR product, as described in Ballard et al. (2002). Reactions lacking DNA were included as negative controls. PCR reactions were prepared in 96-well thin wall PCR plates and amplifications were run in Gene Amp® PCR System 9700 (Applied Biosystems Inc., USA) at conditions optimized for each pair of primers. The PCR program used involved an initial denaturation step of 3 min at 95° C, followed by 40 cycles of 30 s denaturing at 95° C, 30 s annealing at a temperature optimized for each primer pair, and 45 s primer extension at 72°C, followed by 20 min final extension at 72°C. At the end the samples were held at 4°C until retrieval. For Sb 1-10, Sb 6-34, Xcup 49, Xcup 50 primer pairs, the annealing temperature was 56°C. For Sb AGB 02 and Xtxp 285 primer pairs, the annealing temperature was 55°C, and for Dsenhsbm 99, Undhsbm 24, Undhsbm 313, and Xdhsbm 1005 primer pairs, the annealing temperature was 60°C. Following PCR, a few samples were selected randomly and checked on 1.5 % agarose gels for amplification and product intensity for each locus. The PCR products from the ten loci were multiplexed into two panels. Each panel contained PCR products from five loci. PCR products labeled with the same fluorescent dye but multiplexed in the same panel had a size difference of at least 100 base pairs to avoid overlapping. Multiplexed PCR products were then analyzed using ABI 3730 capillary DNA sequencer (Applied Biosystems) at University of Copenhagen, Denmark. The size standard ROX 58-352 was used as a molecular size marker.

### *Data scoring and analysis*

Genemarker software, version 2.2.0 (SoftGenetics, 2011) was used for peak identification and fragment sizing. Allelic data for a particular locus was recorded as fragment size at a co-dominant locus and the genotype of each individual at each locus was recorded. When a PCR product was not obtained, data for the specific loci and samples were treated as missing values. Genetic diversity parameters for each locus were

estimated using POPGENE version 1.31 (Yeh et al., 1999). Arlequin version 3.0 (Excoffier et al., 2005) was used for the analysis of population genetic structure. For the analysis of molecular variance (AMOVA), sorghum accessions were grouped according to their ecological region, their race and the ethnicity of the local populations inhabiting the sites where the accessions were originally collected. Genetic Analysis in Excel (GenAIEx) version 6.4 (Peakall and Smouse, 2006) was used for principal coordinate analysis (PCoA) (Figure 3). Cluster analysis and bootstrapping were performed with the FreeTree Freeware program (Pavlicek et al., 1999) based on Nei's standard genetic distance (Nei and Li, 1987). TreeView 1.6.6 program (Page, 1996) was used to view the trees.

**Conclusion**

This study provides the first analysis and quantification of genetic diversity present in Botswana's sorghum accessions held by the national gene bank. The data obtained demonstrate that there was substantial genetic diversity within and among the accessions conserved in the gene bank as corroborated by both the AMOVA and by the number of alleles recorded. Sorghum accessions exhibiting high allelic richness and high genetic diversity must be prioritized for core collection and breeding programs. The presence of unique and rare alleles in some accessions indicates the need for further evaluation of Botswana *ex situ* conserved accessions to identify genotypes with desirable traits. The data also showed that geographic region; ethnicity of the local population or racial classification of sorghum did not influence the distribution of the genetic diversity in Botswana's sorghum. Molecular characterization – coupled with morphological characterization – is very important for effective genetic resource conservation and utilization.

**Acknowledgements**

The authors acknowledge the Swedish International Development Agency (Sida) through the SADC Plant Genetic Resources Network Centre (SPGRC) for financial support. Sorghum germplasm used in this study was provided by the government of Botswana, to whom we are indebted. We are grateful to Ms Anna Zborowska of the Swedish University of Agricultural Sciences for technical assistance in the laboratory.

**References**

Ayana A, Bryngelsson T, Bekele E (2000) Genetic variation of Ethiopian and Eritrean sorghum (*Sorghum bicolor* (L.) Moench) germplasm assessed by random amplified polymorphic DNA (RAPD). Genet Resour Crop Evol 47:471–482

Ballard LH, Adams PS, Bao Y (2002) Strategies for genotyping: effectiveness of tailing primers to increase accuracy in short tandem repeat determination. J Biomol Tech 13:20-29

Bekele E, Geleta M, Dagne K (2007) Molecular Phylogeny of genus Guizotia (Asteraceae) using DNA sequences derived from ITS. Genet Resour Crop Evol 54:1419–1427

Bibi A, Sadaqat HA, Akram HM, Mohammed MI (2010) Physiological markers for screening sorghum (*Sorghum bicolor*) germplasm under stress condition. Int J Agric Biol 12:451–455

Brown SM, Hopkins MS, Mitchell SE (1996) Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum (*Sorghum bicolor* (L.) Moench). Theor Appl Genet 111:23–30

Bowditch B M, Albright DG, Williams JGK, Braun MJ (1993) Use of random amplified polymorphic DNA (RAPD) in comparative genome studies. Method Enzymol. 224:294–309

Dean RE, Dahlberg JA, Hopkins MS (1999) Genetic redundancy and diversity among 'orange' accessions in the US national Sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39: 1215-1221.

Deu M, Sagnard F, Chantereau J, Calatayud C, Herault D, Mariac C, Pham JL, Vigouroux Y, Kapran L, Traore PS, Mamadou A, Gerard B, Ndjeunga J, Bezancon G (2008) Niger-wide assessment of *in situ* sorghum genetic diversity with microsatellite markers. Theor Appl Genet 116:903–913

Dillon SL, Shapter FM, Henry RJ, Cordeiro G, Izquierdo L, Slade LL (2007) Domestication to crop improvement: genetic resources for *Sorghum* and *Saccharum* (Andropogoneae). Ann Bot 100:975–989

Dje Y, Forcioli D, Ater M, Lefebvre C, Vekemans X (1999) Assessing population genetic structure of sorghum landraces from north western Morocco using allozyme and microsatellite markers. Theor Appl Genet 99:157–163

Dje Y, Heuertz M, Lefebvre C, Vekemans X (2000) Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatelite markers. Theor Appl Genet 100:918–925

Doggett H (1988) Sorghum, 2nd ed. Longman Scientific and Technical. Harlow, UK

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.0:an intergrated software package for population genetic data analysis. Evol Bioinform Online 1:47–50

FAO (2004) Production Year Book 2002 No. 56. FAO Stat. Ser 176 – Food and Agriculture Organization of the United Nations, Rome, Italy

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

Folketsma RT, Rattunde HFW, Chandra S (2005) The pattern of genetic diversity of Guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. Theor Appl Genet 111:399–409

Ghebru BG, Schmidt RS, Bennetzen JB (2002) Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor Appl Genet 105:229-236

Kumar AA, Reddy BVS, Sharma HC, Hash CT, Rao PS, Ramaia B, Reddy PS (2011) Recent advances in sorghum genetic enhancement research at ICRISAT. Am J Plant Sci 2:589–600

Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC, Klein PE (2002) A high density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP (R), RFLP and SSR markers. Plant Mol Biol 48:483–499

Mutegi E, Signard F, Semagn K, Deu M, Muraya M, Kanyenji B, De Villiers S, Kiambi D, Herselman L, Labuschagne M (2011) Genetic structure and relationship within and between cultivated and wild sorghum (*Sorghum bicolor* (L.) Moench) in Kenya as revealed by microsatellite markers. Theor Appl Genet 122:989-1004

Nei M, Li WH (1987) Molecular evolutionary genetics. Columbia University Press. New York

Ng'uni D, Geleta M, Bryngelsson T (2011) Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) accessions of Zambia as revealed by simple sequence repeats (SSR). Hereditas 148:52-62

Nkongolo KK, Nsapato L (2003) Genetic diversity in *Sorghum bicolor* (L.) Moench from different ecogeographical regions in Malawi assessed with RAPDs. Genet Resour Crop Ev 50:149–156

Page RDM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357-358

Pavlicek A, Hrda S, Flegr J (1999) FreeTree-Freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jacknife analysis of the tree robustness. Application in the RAPD analysis of the genus *Frenkelia*. Folia Biol (Praha) 45:97-99

Peakall R, Smouse P (2006) Genetic Analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288-295

Reddy RN, Madhusudhana R, Mohan SM, Chakravarthi DVN, Seetharama N (2011) Characterization, development and mapping of unigene-derived microsatellite markers in sorghum [*Sorghum bicolor* (L.) Moench]. Mol Breed 10.1007/S11032-011-9571-0

Shehzad T, Okuizumi H, Kawase M (2009) Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. Genet Resour Crop Ev 56:809-827

Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Patterson AH, Kresovich S (2002) Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor Appl Genet 105:912-920

SoftGenetics (2011) Genemarker software version 2.2.0. www.softgenetics.com

Taramino G, Tarchini R, Ferrario S, Lee M, Pe ME (1997) Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor Appl Genet 95:66-72

Thudi M, Fakrudin B (2011) Identification of unique alleles and assessment of genetic diversity of rabi sorghum accessions using simple sequence repeat markers. J Plant Biochem Biot 20:74-83

Uptmoor R Wenzel W, Friedt W, Donaldson G, Ayisi K, Ordon F (2003) Comparative analysis on the genetic relatedness of *Sorghum bicolor* accessions fron southern Africa by RAPDs, AFLPs and SSRs. Theor Appl Genet 106:1316-1325

Varshney RK, Graner A, Sorrels ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48-55

Vierling RA, Xiang Z, Joshi CP, Gilbert ML, Nguyen HT (1994) Genetic diversity among elite sorghum lines revealed by restriction fragment length polymorphisms and random amplified polymorphic DNAs. Theor Appl Genet 87:816-820

Westman AL, Kresovich S (1997) Use of molecular marker techniques for description of plant genetic variation. In: Callous JA, Ford-Llyod BV, Newbury HJ (eds) Plant genetic resources. Conservation and use. CAB Int. Wallingford, United Kingdom. Pp. 9-48

Yeh FC, Yang R, Boyle T (1999) Popgene version 1.31. http://www.ualberta.ca/fyeh/