

Full Length Research Paper

# Random amplified polymorphic DNA (RAPD) based assessment of genetic relationships among some Zimbabwean sorghum landraces with different seed proanthocyanidin levels

Z. Dhlamini<sup>1\*</sup> and I. Sithole-Niang<sup>2</sup>

<sup>1</sup>Department of Applied Biology and Biochemistry, National University of Science and Technology, P O Box AC939, Ascot Bulawayo, Zimbabwe.

<sup>2</sup>Department of Biochemistry, University of Zimbabwe, P O Box MP 167, Mount Pleasant, Harare, Zimbabwe.

Received 27 January, 2014; Accepted 12 March, 2014

Knowledge of genetic distances between genotypes is important for efficient organization and conservation of plant genetic resources for crop improvement programs. In this study genetic distances between genotype pairs (complements of Jaccard's similarity coefficient) were estimated from Random Amplified Polymorphic DNA (RAPD) data collected from 48 Zimbabwean sorghum landraces. These varieties showed variation in their seed proanthocyanidin (PAs) levels with 16 and 29 of them having detectable and non-detectable PA levels respectively. RAPDs revealed considerable genetic variation between the varieties used and 2.7 polymorphisms per primer were obtained. Ninety nine polymorphic RAPD bands were used to calculate genetic distances and the mean genetic distance between the genotypes was 0.494 ( $\pm$  0.113) with a range of 0.051 to 0.761. A multidimensional scaling (MDS) plot of the distance matrix revealed two distinct clusters of cultivated and wild sorghums. No clustering of genotypes according to their seed proanthocyanidin levels was revealed by MDS analysis; also the mean genetic distances of genotypes in the low, medium and high PA categories were not different from each other and none of them was significantly different from the mean genetic distances between all the groups. The RAPD markers used in the present study could not distinguish between sorghums with different PA levels in their seeds; however, the protocol established could be useful in further analysis of this trait in near isogenic lines.

**Key words:** Genetic distances, multidimensional scaling, proanthocyanidins, Random Amplified Polymorphic DNA (RAPD), *Sorghum bicolor*, Zimbabwe

## INTRODUCTION

Sorghum (*Sorghum bicolor* (L.) Moench) is a traditional cereal crop in Zimbabwe and it ranks fourth in production

after maize, wheat, and pearl millet (FAO, 2006). Sorghum utilization is generally influenced by the presence

\*Corresponding author. E-mail: Zephaniah.dhlamini@nust.ac.zw.

of polyphenolic compounds that are produced in large quantities in grain and vegetative tissues of many cultivars (Waniska, 2000). The polyphenolic compounds of sorghum such as the proanthocyanidins (PAs), also known as condensed tannins, have protein binding properties, which tend to reduce the nutritional quality of sorghum based diets in both livestock and humans. However, despite their nutritional side effects, sorghum polyphenols have been implicated in defense against competitors, herbivores and pathogens (Winkel-Shirley, 2001). There is need to develop high yielding sorghum cultivars with desirable levels and types of polyphenolic compounds that will improve the nutritional qualities of the crop without compromising the positive agronomic traits conferred by these compounds.

Future sorghum improvement programmes have increased the utilization of local germplasm resources. There are indications that communal sorghum farmers are selecting some local landraces for traits such as drought and pest tolerance, disease resistance, early maturity, palatability and storability among others (Nagaraj et al., 2013). The availability of high yielding sorghum cultivars with traits desired by farmers, is likely to result in increased sorghum production and thus improved food security in the semi-arid regions.

Zimbabwe is one of the few African countries with a rich and varied gene pool of sorghum landraces that has been selected and built-up over centuries (Chakauya et al., 2006). If this germplasm is to be widely utilized in sorghum breeding programmes, there is need for detailed understanding of its genetic diversity. This genetic diversity can be evaluated by one of several means, most of which enable the estimation of genetic distances between the landraces (Chakauya et al., 2006; Abdel-Fatah et al., 2013; Ng'uni et al., 2011). Information on the degree to which lines or populations are genetically related can help breeders in making plans for genetic crosses, in assigning available breeding materials to specific heterotic groups, and in the identification of individual varieties with reference to plant varietal purity and its maintenance (Mohammadi and Prasanna, 2003). Furthermore, knowing the level of genetic variation in a germplasm collection can facilitate more efficient sampling of genotypes for particular needs and identifying lines that should be kept to preserve maximum genetic diversity in germplasm banks, thereby facilitating efficient handling of germplasm resources (Bretting and Widrechner, 1995).

There are different genetic markers such as morphological traits, protein and DNA markers that can be used in genetic diversity studies. Random amplified DNAs (RAPDs) were used in this study because they are relatively cheap and easy to perform; they require small amounts of DNA material detect relatively small amounts of genetic variation and enable inexpensive generation of data that can be subjected to different statistical manipulations.

RAPD based assessment of genetic similarities in plants usually employs either one of the three commonly used

similarity coefficients (Dudley 1994), which are the simple matching coefficient (Sneath and Sokal, 1973), Jaccard's coefficient (Gower, 1972) and Nei and Li's coefficient (Nei and Li, 1979). All these similarity coefficients are non-negative and have an upper limit of unity. In such cases where the similarity measure is bound by zero and unity there is always a dissimilarity, this dissimilarity is the genetic distance between the two genotypes  $i$  and  $j$ . Just like similarity, the dissimilarity is symmetric and non-negative. Naturally, an organism has maximal similarity to itself, thus  $S_{ii}=1$  and  $GD_{ij}=0$  will mean no genetic difference, while  $GD_{ij}=1$  signifies complete difference between the two genotypes (Everitt, 1993; Nienhuis et al., 1994).

The genetic distances derived from heritable characteristics (genetic markers) define the phenetic patterns of a population (Abbott et al., 1985). Usually discontinuities exist in such patterns, resulting in groupings with different ranges of variation within groups and varying degrees of differences between them. It is important for the investigator to visualise these groupings because a lot of information for decision-making can be deduced from them. Thus, the third step in a genetic distance estimation study is to transform, by statistical methods, the genetic distance matrix into a diagrammatic form (clustering), from which the phenetic groupings can be easily identified. In molecular marker studies, dendrogram construction and ordination techniques are the commonly used clustering techniques. Principal component analysis (PCA), principal co-ordinate analysis and non-metric multi-dimensional scaling (MDS) are the most commonly used ordination techniques. In this study MDS was used.

Our objective in this study was to use the RAPD technique to estimate genetic distances between some Zimbabwean sorghum landraces and to identify RAPD markers that can be used to discriminate between cultivars with different levels of proanthocyanidins in their seeds.

## MATERIALS AND METHODS

### Plant material

Forty-eight sorghum varieties collected from different parts of Zimbabwe were used in this study. These included 37 randomly sampled landraces, cultivated by rural farmers, 5 commercial cultivars, 3 breeder's experimental lines and 3 wild sorghums (*Sorghum arundinaceum* Desv) (Table.1).

### Determination of proanthocyanidin (PAs) levels in sorghum seeds.

Forty five sorghum varieties in this study had their seeds tested for the presence of soluble and insoluble proanthocyanidins (PAs) using the butanol-HCl method described by Bate-Smith (1975). The three wild genotypes were not assayed for tannins because the quantity of seeds available was not adequate for the assay.

### DNA extraction and quantification

Seeds of all the varieties were germinated in the greenhouse. Fresh leaf tissue was harvested from 7-day-old seedlings, per

**Table 1.** Sorghum cultivars used in the study, their common names, areas of origin and seed proanthocyanidin levels.

Genotype	Common Name	Area of Origin	PA Status	
			A550 nm*	Classification
<b>Commercial cultivars</b>				
SV1	SV1	Harare	0.043	Low
SV2	SV2	Harare	0.039	Low
ZWSH1	ZWSH1	Harare	0.005	Low
<b>Experimental lines</b>				
NL700	NL700	Harare	0.004	Low
NL829	NL829	Harare	0.026	Low
DC75m	DC75m	Harare	0.032	Low
DW	Dwarf Wonder	Harare	0.238	Medium
<b>Wild Lines</b>				
W9	Wild sorghum	Harare	U	Unknown
W10	Wild sorghum	Harare	U	Unknown
W15	Wild sorghum	Harare	U	Unknown
<b>Local landraces</b>				
PL01	Isitishi	Matopo	0.004	Low
PL02	Ntelwane	Matopo	0.031	Low
PL04	Katandanzara	Mutoko	0.007	Low
PL05	Ludende	Gwanda	0.002	Low
PL08	Iganu	Matopo	0.014	Low
PL09	Tswetha	Matopo	0.024	Low
PL10	Muchaina	Chipinge	0.005	Low
PL12	Rudende	Mutoko	0.003	Low
PL14	Brown tsweta	Gwanda	0.031	Low
PL15		Harare	0.020	Low
16Ga		Magaisa	0.003	Low
68Db	Sikhothama	Tsholotsho	0.029	Low
102K	Murara	St Joseph	0.041	Low
107D	Chifumbata	St Joseph	0.007	Low
370	Imfe/ipwa	Zaka	0.007	Low
499	Mugonellengo	Gata Camp	0.015	Low
581		Mt Selinda	0.009	Low
646		Mt Selinda	0.021	Low
653		Mt Selinda	0.005	Low
310		Zaka	0.004	Low
671		Zhombe	0.004	Low
686		Zhombe	0.005	Low
877	Imfe/ipwa	Manyoni	0.008	Low
886	Imfe/ipwa	Manyoni	0.557	Medium
4D		Chivi	0.426	Medium
7B	Chibedlani	Gutu	0.473	Medium
34K	Gangara	Malipati	0.138	Medium
44B		Lundi	0.292	Medium
71C	Tsveta	Shurugwi	0.389	Medium
304	Imfe/ipwa	Turgwana	0.417	Medium
304a	Imfe/ipwa	Turgwana	0.417	Medium
21G	Mugore	Kyle	0.639	High
37Ce	Chititshi	Malipati	0.958	High
PL03	Red Swazi	Chipinge	0.722	High
PL06	Chibonda	Mutoko	0.743	High
42K	Gangara	Lundi	0.880	High
We2		Harare	0.622	High
10B3c		Gutu	0.949	High

\*Absorbance at 550 nm.

variety and 0.75 g of the pooled leaf tissue was used for DNA extraction. The DNA extraction buffer was modified from Jhingan (1992), with potassium ethyl xanthogenate (PEX), (Fluka Chemical Corp., USA) replacing sodium ethyl xanthogenate. The weighed leaf tissue was ground to a fine pulp in 50  $\mu$ L of PEX extraction buffer [200 mM Tris buffer (pH 7.5), 1.4 M NaCl, 600 mM PEX, 100 mM EDTA (pH 8.0)] in 1.5 ml micro centrifuge tubes using plastic grinding rods. An additional 450  $\mu$ L PEX extraction buffer was added, the tubes were vortexed briefly and incubated in a water bath at 65°C for 1 h. Thereafter, the samples were centrifuged at 10 000 revolutions per minute (RPM) in a microcentrifuge for 10 min. Supernatants were transferred to clean microcentrifuge tubes containing 1000  $\mu$ L of a 6:1 mixture of absolute ethanol and 7.5 M ammonium acetate to precipitate the nucleic acids at room temperature. After 30 min of precipitation, the nucleic acids were collected by centrifugation at 3000 RPM. The pellets were re-suspended in 300  $\mu$ L of TE buffer (1 mM Tris pH 7.5; 0.1 mM EDTA pH 8.0) containing 100  $\mu$ g/ml RNase A and incubated in a water bath at 37°C for 1 h. Any remaining tissue debris was pelleted from the suspension by centrifugation at 14 000 RPM for 1 min and the supernatants transferred to clean microcentrifuge tubes. To the supernatant, 1200  $\mu$ L of a 10:1 mixture of absolute ethanol and 3 M sodium acetate was added to precipitate the DNA at room temperature for 30 min. The DNA precipitate was pelleted by centrifugation at 3000 RPM for 5 min. The pelleted DNA was washed by gentle vortexing in 70% ethanol followed by centrifugation at 14 000 RPM for 15 min to collect the clean pellet. The washed DNA pellet was air-dried and finally dissolved in 75  $\mu$ L of TE buffer (pH 7.5). The DNA in each sample was quantified by using a DNA fluorometer (Hoeffer Scientific Instruments, USA) and diluted in TE-tartrazine (TE buffer with 200 mM tartrazine) to a working concentration of 4 ng/ $\mu$ L.

#### RAPD Procedures and Primer Screening

All RAPD reactions were done in a total volume of 10  $\mu$ L. Each reaction was carried out in RAPDs buffer [50 mM Tris buffer (pH 8.5), 20 mM KCl, 3.5 mM MgCl<sub>2</sub>, 0.05% (w/v) bovine serum albumin, 0.01% xylene cyanol and 1.25% (w/v) Ficoll 400] on 20 ng of DNA, 1 unit of Taq DNA polymerase enzyme (Promega, USA), 1  $\mu$ M of random decamer (Operon technologies, Alameda, California, USA and University of British Columbia, Canada) and 0.2 mM of each dNTP (Skroch and Nienhuis, 1995). Eight varieties representing four different seed PA groups (low, medium, high and unknown) were evaluated for genetic polymorphisms using 70 randomly selected RAPD primers. The PCR products were electrophoresed on 1.5% agarose gels, stained with Ethidium Bromide and photographed over UV light onto Polaroid 667 film. The gel pictures were then used to identify the primers which produced clear polymorphic bands. On this basis, 40 primers were selected for use in the main RAPD study for estimating genetic distances among the 48 sorghum varieties.

#### Thermal cycling conditions

All RAPD reactions were performed in thin walled 96-well plates in an MJ PC100 thermocycler (MJ Research, Water Town, MA, USA). A total of 39 cycles were performed, in the first cycle the temperature settings were: 91°C for denaturation for 60 s, 42°C for annealing for 15 s followed by elongation at 72°C for 70 s. The subsequent 38 cycles had denaturation time set at 15 s, annealing at 15 s and elongation at 70 s with temperatures similar to first cycle for each of the three PCR steps.

#### RAPD band scoring

Like all dominant molecular marker techniques RAPDs generate

binary data, thus when comparing two genotypes *i* and *j* using this kind of data there are four possible outcomes: [1,1], [0,1], [1,0] or [0,0], (1= presence and 0= absence of a band (genetic marker) in genotype *i* and *j* respectively). Polymorphic bands were scored from the gel photographs. Monomorphic bands, which were the majority of bands seen on the sorghum RAPD gels, were not scored. Two criteria were used in scoring bands: firstly, the band had to stain strongly; secondly, there had to be an unambiguous difference between the allelic states of the band being scored, that is, presence or absence of a band. Each polymorphic band was treated as a unit character, and each variety was scored for the presence or absence of a band, scored 1 or 0, respectively.

#### Statistical analyses

The scored bands data were used to calculate genetic distances using the Jaccard's similarity coefficient ( $J_{ij}$ ) (Gower, 1972):

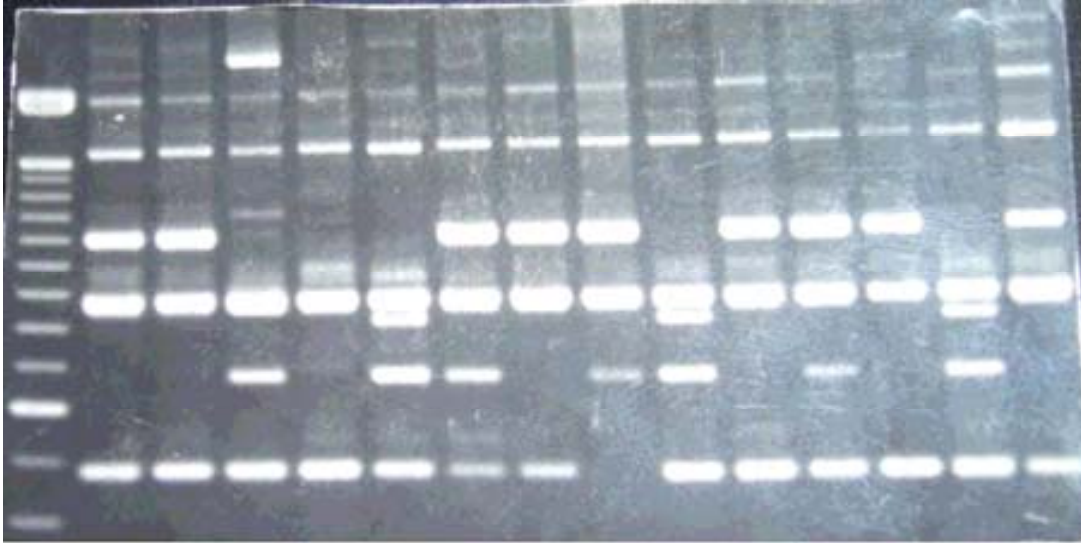
$$J_{ij} = \frac{N_{(1,1)}}{N_{(1,1)} + N_{(1,0)} + N_{(0,1)}} \quad (1)$$

where  $N_{(1,1)}$ ,  $N_{(1,0)}$  and  $N_{(0,1)}$  is the number of times the cultivars *i* and *j* both have a particular band, *i* has a band while *j* does not and *j* has a band while *i* does not, respectively. This similarity coefficient is non-negative and has an upper limit of unity. In such a case where the similarity measure is bound by zero and unity, there is always a dissimilarity, which is the genetic distance ( $GD_{ij}$ ) between two genotypes *i* and *j*. Just like similarity the dissimilarity is symmetric and non-negative.

$$GD_{ij} = 1 - J_{ij} \quad (2)$$

Naturally, an organism has maximal similarity to itself, thus  $S_{ii} = 1$  and  $GD_{ij} = 0$  will mean no genetic difference, while  $GD_{ij} = 1$  signifies complete difference between the two genotypes (Everitt, 1993; Nienhuis et al., 1994). Genetic distances between all the 1128 possible genotype pairs [ $n(n-1)/2$ , where *n* is the number of genotypes in the study] from the 48 accessions were calculated using the correlation procedure of the statistical programme, SYSTAT 5.2 (Wilkinson, 1992). This produced a 48 × 48 genetic distance matrix. For the purposes of visualizing the genetic relationships between the cultivars with different seed proanthocyanidin levels, the cultivars were classified into four groups based on the butanol HCl assay for PAs. The four groups were; high, medium and low (with absorbance greater than 0.5, between 0.140 and 0.5, and less or equal to 0.05, respectively) and unknown in the case of wild lines whose PA levels were not determined (Table 1). The genetic distance matrix was converted to two-dimensional coordinates using the multidimensional scaling (MDS) procedure in SYSTAT 5.2. The objective of MDS is to estimate the coordinates of a set of genotypes in a space of specified dimensionality from data measuring the relationships between pairs of genotypes (SAS Institute, 1990; Schiffman et al., 1981). The coordinates are supposed to represent the information from the genetic matrix so that there is maximum correspondence between the observed proximities and inter-point distances (Everitt, 1993).

Thus the larger the calculated genetic distance between two individuals the further apart should the points representing them on the plot. To determine if a subset of one or more RAPD bands could be selected that would allow classification of the cultivated sorghum genotypes into three proanthocyanidin groups (high, medium and low), an additional analysis was performed. This involved ranking the RAPD bands according to their ability to separate the cultivars into the three groups, thus maximizing variances of band frequencies among groups; this was done by calculating individual



**Figure 1.** RAPD banding patterns with primer OPAS-14 in 15 sorghum cultivars: 10B3c, 304, 304a, 34K, PL14, 4D, 646, 877, DW, 42K 7B, PL9, DC75, 16Ga and PL10 respectively. The first lane is a 100 bp molecular weight marker.

band frequencies in all the groups. In this case the frequency of a band is the proportion of genotypes in a particular group having the band, relative to all evaluated genotypes. The bands were then ranked by variance of band frequencies across groups. The best 15 bands were selected and used to calculate new genetic distances and the genetic distance matrix was used in MDS analysis as described above. This analysis did not give a clear distinction between the three groups of cultivated sorghums. The classification was then changed to include only two PA groups; those with detectable PA levels (high) and those without (low). Band frequencies among these two groups were calculated and 15 bands with the greatest differences in amplification frequencies between the two groups were used to compute genetic distances and MDS analysis was performed on the resultant distance matrix.

## RESULTS

### Variation in seed PA content

The sorghum landraces and commercial cultivars used in this study showed significant variability in their seed proanthocyanidin levels. Of the 45 genotypes assayed for PAs using the butanol-HCl assay, 16 (36%) had detectable PA levels while 29 (64%) did not have detectable PA levels (Table 1).

### Degree of genetic polymorphisms in sorghums as revealed by RAPDs

Of the 70 primers screened for their ability to detect polymorphisms in sorghum, only 5 did not amplify DNA at all from the 8 genotypes used in the primer screening experiments. Among those primers that amplified DNA fragments from the sorghum templates only one primer (OPA-08) produced a single monomorphic band

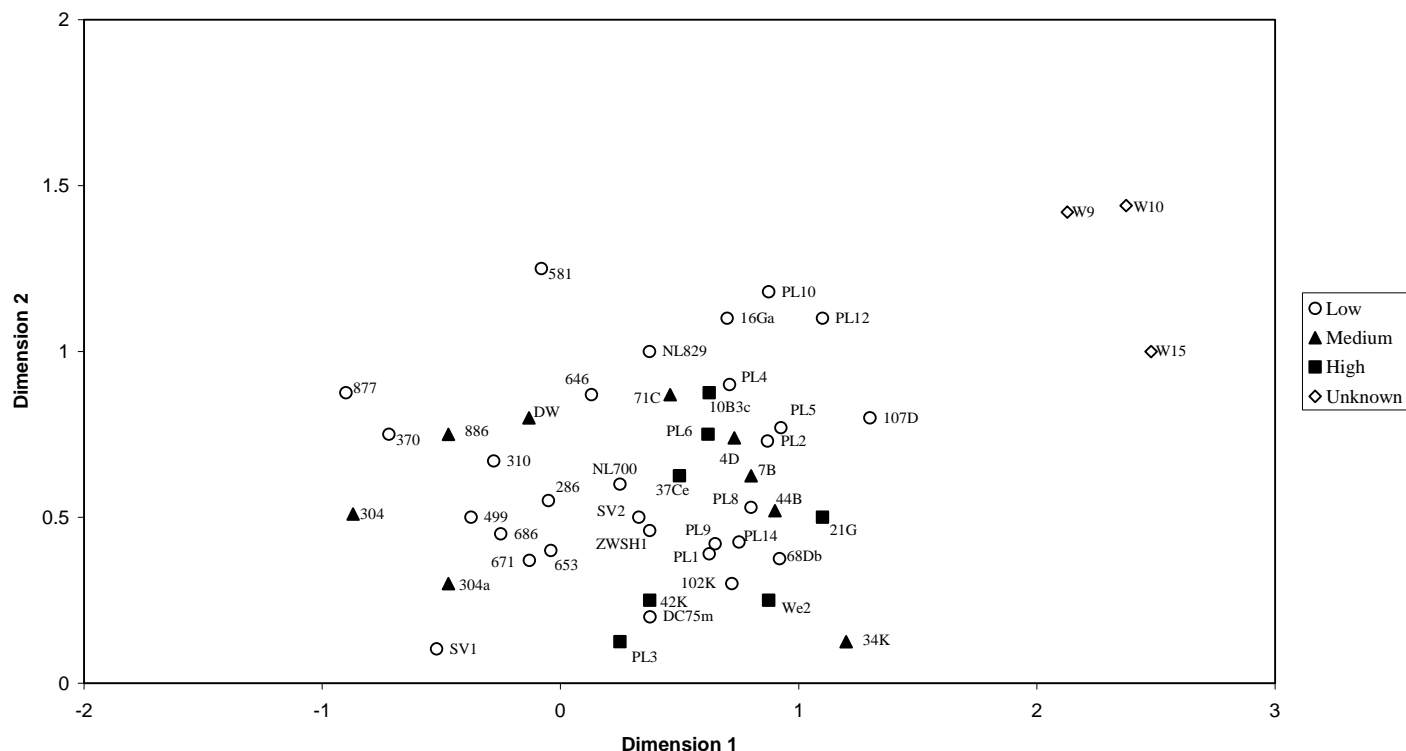
whereas, the rest produced multiple banding profiles (Figure 1). Most of the RAPD bands obtained were monomorphic. The number of polymorphic bands produced by the selected 40 primers among the 48 genotypes ranged from 1 (as in OPK-15 and UBC-72) to 6 (OPG-05), with the average being 2.7 polymorphisms per primer. In total, 99 polymorphic bands were scored and used in the genetic distance studies.

In this study, the mean frequency of amplification of a polymorphic band was  $0.57 \pm 0.05$ . Taking the bands individually, none of them could be used to tag any commercial cultivar. However primer OPA-14 produced a band (~420bp) that was only unique to genotype 42K.

The RAPD bands were used to calculate genetic distances between genotypes and the average genetic distance for the 1128 inter-pair comparisons was  $0.494 \pm 0.113$ , with a range of 0.053 to 0.761. The entire  $48 \times 48$  triangular matrix of genetic distances is too lengthy to be shown here, however part of the matrix is shown in Table 2.

### Relationships revealed by MDS analysis

The MDS plot of the genetic distances derived from the 99 polymorphic RAPD bands is shown in Figure 2. This plot was a good fit to the distance matrix since the stress level (the goodness of fit parameter for MDS) was 0.07. A stress level of 0.05 is described as an excellent fit, 0.10 a good fit, 0.2 a fair fit and 0.4 a poor fit (Kruskal, 1964). The sorghum genotypes fall into 2 clusters. The main cluster is made up of cultivated sorghums and the other cluster is made up of the wild sorghums (W9, W10 and W15). Within the major cluster of cultivated sorghums the



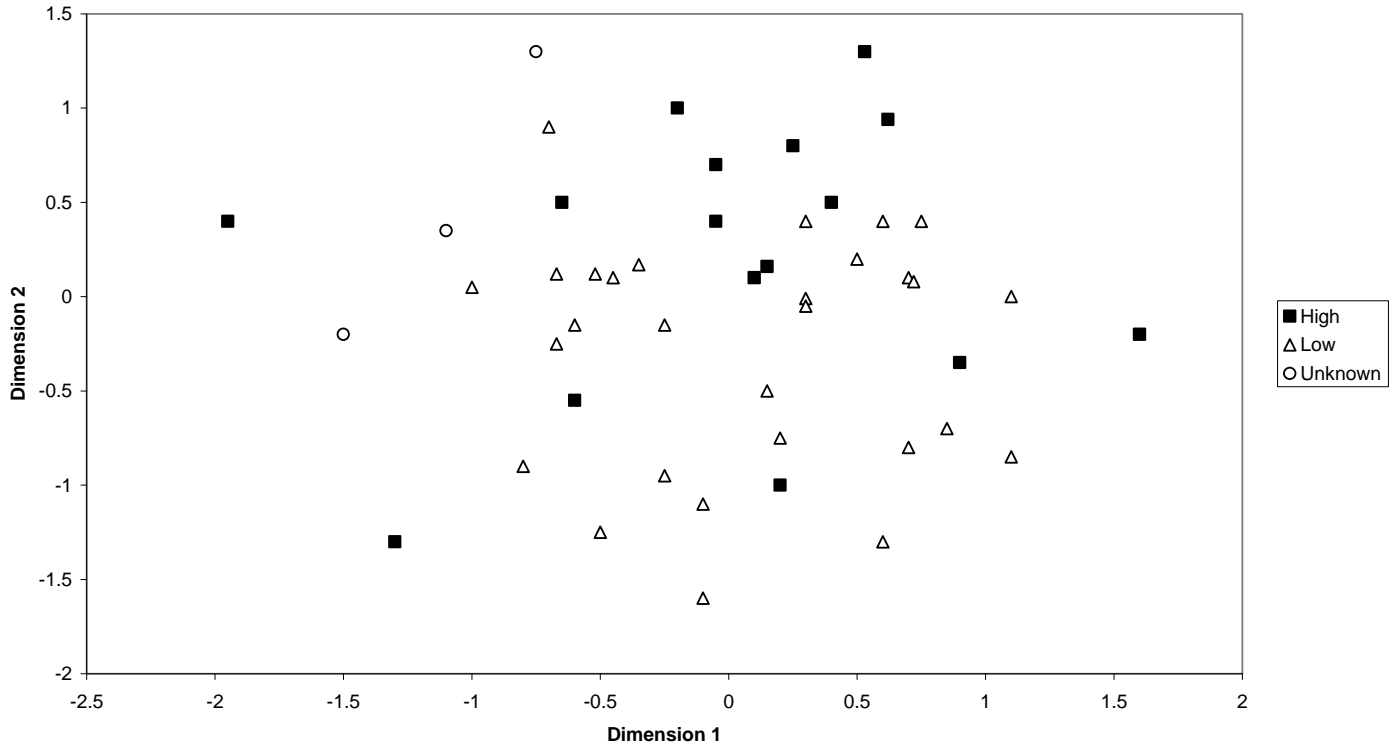
**Figure 2.** MDS plot of genetic distances of 48 Zimbabwean sorghum cultivars calculated from 99 RAPD bands and classified for their seed proanthocyanidin content as high, medium, low or unknown.

**Table 2.** Part of the matrix of genetic distances between some pairs of sorghum cultivars used in the study.

<b>10B3c</b>	0.000															
<b>304</b>	0.557	0.000														
<b>304a</b>	0.535	0.164	0.000													
<b>34K</b>	0.448	0.565	0.549	0.000												
<b>PL14</b>	0.358	0.521	0.457	0.412	0.000											
<b>4D</b>	0.423	0.514	0.493	0.472	0.188	0.000										
<b>646</b>	0.441	0.616	0.600	0.575	0.382	0.423	0.000									
<b>877</b>	0.608	0.484	0.422	0.611	0.587	0.615	0.563	0.000								
<b>DW</b>	0.493	0.522	0.485	0.493	0.573	0.603	0.569	0.460	0.000							
<b>42K</b>	0.414	0.600	0.560	0.500	0.394	0.453	0.406	0.507	0.493	0.000						
<b>7B</b>	0.338	0.560	0.539	0.414	0.200	0.324	0.408	0.620	0.553	0.397	0.000					
<b>PL9</b>	0.343	0.507	0.443	0.420	0.051	0.200	0.391	0.592	0.541	0.380	0.212	0.000				
<b>PL15</b>	0.421	0.425	0.384	0.405	0.390	0.444	0.442	0.494	0.352	0.368	0.412	0.355	0.000			
<b>16Ga</b>	0.364	0.589	0.534	0.486	0.400	0.438	0.541	0.632	0.562	0.486	0.446	0.431	0.475	0.000		
<b>PL10</b>	0.344	0.583	0.587	0.500	0.414	0.452	0.514	0.679	0.575	0.500	0.371	0.423	0.524	0.297	0.000	
	<b>10B3c</b>	<b>304</b>	<b>304a</b>	<b>34K</b>	<b>PL14</b>	<b>4D</b>	<b>646</b>	<b>877</b>	<b>DW</b>	<b>42K</b>	<b>7B</b>	<b>PL9</b>	<b>PL15</b>	<b>16Ga</b>	<b>PL10</b>	

sweet-stem sorghums (877, 886, 370 and 304) group together. The genotypes did not cluster into their PA groups as defined in Table 1. Thus the RAPD markers generated in this study could not distinguish between the low, medium and high PA sorghum cultivars. This was further confirmed by comparing the mean genetic

distances of genotypes in different PA groups. If a particular group of genotypes is genetically well differentiated from the other groups, the mean GD among genotypes within the group should be smaller than that between the other groups (Menkir et al., 1997). In this study the mean GDs among the genotypes with high, medium and low



**Figure 3.** MDS plot of genetic relationships among sorghum genotypes having high and low seed PA levels. Genetic distances were derived from 10 RAPD markers showing the greatest differences between high and low groups in terms of the frequency of amplification within each group.

PA levels which were  $0.4802 (\pm 0.0788)$ ;  $0.4876 (\pm 0.0533)$  and  $0.4680 (\pm 0.1166)$ , respectively are not significantly different hence we cannot distinguish these groups using the 99 RAPD bands. After maximizing band frequency variances between groups no single RAPD band was found that had an absolute ability to distinguish the groups. Band UBC-180(2000) had the highest separation ability having a frequency variance of 0.220. The majority of the bands had much lower values ranging from 0.0 to 0.128. Fifteen bands with the highest frequency variances between the groups were selected and used to compute genetic distances. These genetic distances still failed to separate the three groups according to their differences in seed PA content (MDS plot not shown). Even after classifying the genotypes as either high or low in PAs and band frequencies between the groups calculated, there still wasn't any band capable of separating the two groups. However, the MDS plot (Figure 3) produced from distances derived from the best 15 bands separated the two groups in a much better way compared to when they were classified into three groups.

## DISCUSSION

The genotypes used in this study were collected from different parts of Zimbabwe and the fact that about 35% of these sorghums were high in PAs may be an indication that this random sample was almost representative of the

national sorghum germplasm. Obilana (1991), after evaluating a considerable number of Zimbabwean sorghum landraces found that about 33% of the sorghums are high in PAs by looking at seed colour.

The RAPD technique used in this study seems to have been well optimized for sorghum since the average number of polymorphisms detected (2.7 per primer) is within the range of what other groups (Zhan et al., 2012; Agama and Tuinstra, 2003) working on sorghum usually get. The mean genetic distance of about 0.5 between all these genotypes demonstrates adequate coverage of the genome and that there is a substantial level of variation within Zimbabwean sorghums. Furthermore it can be said that the genotypes were discriminated efficiently since the RAPD markers used showed genetic independence, with frequency of amplification for any polymorphic marker being 0.567 (this value must be  $>0.5$  if the RAPD markers are to distinguish genotypes efficiently (Noormohammadi et al., 2012).

Other sorghum studies (Menkir et al., 1997) reveal that there is a considerable variation in the world collection of sorghum and that genotypes from Southern Africa are generally less diverse than those from East and Central Africa. This observation is further substantiated by the absence of clear genetic clusters within the cultivated Zimbabwean sorghum landraces used in this study.

This lack of major genetic subdivisions in this sorghum collection may be an indication that these landraces have

not been significantly isolated in space and thus the introgression of genetic material between them has been occurring over time. Moreover, in the communal farming sector different sorghum landraces are grown in close proximity. The high levels of polymorphisms between some individual genotypes identified in this and other studies may be due to artificial selection for different traits and creation of new genetic combinations by breeders.

This study did not reveal any significant genetic relationships between sorghum cultivars based on their seed PA levels. Since no marker for PAs was found in this population, which was seemingly adequately covered with the RAPD markers used, one can conclude that the trait is controlled by a relatively small portion of the genome otherwise polymorphisms for it could have been identified. Since the MDS plot (Figure 2) does not reveal any PA level group clusters it is concluded that the presence or absence of this trait did not in any way influence the evolution or selection of sorghums over the years, thus sorghums with high PA levels did evolve together with low PA cultivars and have many other traits in common. Furthermore, the observation that MDS plots derived from RAPD markers with maximized variances between the high and low PA groups (Figure 3) revealed a better separation of the genotypes, maybe an indication that the presence or absence of PAs in sorghum seeds is not a polygenic trait. This is in agreement with the results from classical genetics experiments carried out by Woodruff et al., (1982). The RAPD markers for PAs can be obtained if another approach such as use of the near isogenic-lines (NILs) segregating the trait is adopted. The RAPDs protocol established in this work, together with the genetic distance information and PA analysis data can be used as a quick and cost effective method to pre-screen NIL production protocols.

It is possible that the different cultivars are mutants at different structural and regulatory loci controlling flavonoid biosynthesis (Wu et al., 2012). Studying these loci using high throughput genotyping methods such as TILLING (Targeting Induced Local Lesions in Genomes) can aid the development of useful genetic markers for PAs in sorghum (Blomstedt et al., 2011).

## REFERENCES

- Abbott LA, Bisby FA, Rogers DJ (1985). *Taxonomic Analysis in Biology: Computers, Models, and Data Bases*. Columbia University Press, New York.
- Abdel-Fatah BE, Ali EA, Tag AA, Hessein EM (2013). Genetic diversity among sorghum (*Sorghum bicolor* L. Moench) Landraces in Agromorphological Traits and Molecular Markers. *Asian J. Crop Sci.* 5(2): 106-124.
- Agama HA, Tuinstra MR (2003). Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. *Afr. J. Biotechnol.* 2(10): 334-340.
- Blomstedt CK, Gleadow RM, O'Donnell N, Naur P, Jensen K, Laursen T, Olsen CE, Stuart P, Hamill JD, Møller BL (2012). A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. Moench resulting in acyanogenic forage production. *Plant Biotechnol. J.* 10(1): 54-66.
- Bretting PK, Widrechner MP (1995). Genetic markers and horticultural germplasm management. *HortScience.* 30(7):1349-1356.
- Chakauya E, Tongoona P, Matibiri EA, Grum M (2006). Genetic diversity assessment of sorghum landraces in Zimbabwe using microsatellites and indigenous Local Names. *Int. J. Bot.* 2(1):29-35.
- Dudley JW (1994). Comparison of genetic distance estimators using molecular marker data. In: *Analysis of Molecular Marker Data*, Am. Soc. Hortic. Sci., Corvallis, Oregon USA.
- Everitt BS (1993). *Cluster Analysis* 3<sup>rd</sup> ed. Edward Arnold, London.
- FAO (2001). *FAO production year-book, 2001*, Rome Italy.
- Gower JC (1985). Measures of similarity, dissimilarity and distance; In *Encyclopedia of Statistical Sciences* 5, eds. Kotz. S and Johnson, pp.397 - 405 N. L. Wiley, New York.
- Jhingan AK (1992). A novel technology for DNA isolation. *Methods Mol. Cell. Biol.* 3(15): 15-22.
- Kruskal JB (1964). Multidimensional Scaling by optimizing goodness of fit to a non metric hypothesis. *Psychometrika* 29:1-27.
- Menkir A, Goldsbrough P, Ejeta G (1997). RAPD based assessment of genetic diversity in cultivated races of sorghum. *Crop Sci.* 37: 564-569.
- Mohammadi SA, Prasanna BM (2003). Analysis of Genetic Diversity in Crop Plants: Salient Statistical Tools and Considerations. *Crop Sci.* 43:1235-1248.
- Nagaraj N, Sachin M, Surajit H, Ban MCS, Chandrakanth, MG (2013). Baseline Scenario of Post-rainy Season Sorghum Economy in Marathwada Region of Maharashtra. Working Paper Series No. 38. Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 28 pp.
- Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* 76: 5269-5273.
- 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.
- Nienhuis J, Tivang J, Skroch P (1994). Analysis of genetic relationships among genotypes based on molecular marker data. In: *Analysis of Molecular Marker Data*, American Society for Horticultural Science, Corvallis, Oregon USA.
- Nienhuis J, Tivang J, Skroch P, dos Santos JBN (1995). Genetic relationships among cultivars and landraces of lima bean (*Phaseolus lunatus* L.) as measured by RAPD markers. *J. Am. Soc. Hortic Sci.* 120(2): 300-306.
- Noormohammadi Z, Fasihee A, Homae-Rashidpoor S, Sheidai M, Baraki SG, Mazooji A, Tabatabaee-Ardakani AZ (2012). Genetic variation among Iranian pomegranates (*Punica granatum* L.) using RAPD, ISSR and SSR markers. *Aust. J. Crop Sci.* 6(2): 268-275.
- Obilana AT (1991). Description and evaluation of sorghum germplasm from Zimbabwe. In: *Proceedings of the 21<sup>st</sup> Anniversary: Crop production Congress*, Harare, Zimbabwe.
- SAS Institute 1990 *SAS/STAT User's Guide*, 6<sup>th</sup> Edition, SAS Institute, Cary, NC, USA.
- Schiffman SS, Reynolds ML, Young FW (1981). *Introduction to Multidimensional Scaling*. Academic Press, New York.
- Skroch P, Nienhuis J (1995). Qualitative and quantitative characterization of RAPD variation among snap bean (*Phaseolus vulgaris*) genotypes. *Theor. Appl. Genet.* 91:1078-1085.
- Waniska RD (2000). Structure, phenolic compounds, and antifungal proteins of sorghum caryopses. Pages 72-106 in *Technical and institutional options for sorghum grain mold management: proceedings of an international consultation*, 18-19 May 2000, ICRISAT, Patancheru, India (Chandrashekar A, Bandyopadhyay R, Hall AJ, eds.). Patancheru 502 324, Andhra Pradesh, India: Int. Crops Res. Inst. Semi-Arid Tropics.
- Wilkinson L (1992). *SYSTAT 5.2: The system for statistics*. SYSTAT, Evanston, Illinois, USA
- Winkel-Shirley B (2001). Flavonoid Biosynthesis: A colourful model for genetics, biochemistry, cell biology and plant biotechnology. *Plant Physiol.* 1126: 485-493.
- Woodruff BJ, Cantrell RP, Axtell JD, Butler LG (1982). Inheritance of tannin quantity in sorghum. *J. Hered.* 73:214-218.
- Wu Y, Li X, Xiang W, Zhu C, Lin Z, Wu Y, Li J, Pandravada S, Ridder DD, Bai G, Wang ML, Trick HN, Scott R, Bean SR, Tuinstra MR,



Tesfaye T, Tesso TT, and Yu J (2012). Presence of tannins in sorghum grains is conditioned by different natural alleles of Tannin1. PNAS 109(26):10281-10286

Zhan Q, Zhou L, Bi N, Wu H, Li J, Lu J, Lin P (2012). A Comprehensive Analysis of Genetic Diversity and Relationships of 13 Sweet Sorghum Varieties as Energy Sources. J. Sust. Bioenergy Syst. 2(4):86-91.