Biodiversity and Conservation (2006) 15:3251–3265
DOI 10.1007/s10531-005-0313-7


# Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers

NEMERA GELETA[1,2], MARYKE T. LABUSCHAGNE[2,*] and CHRIS D. VILJOEN[2]

*[1]Department of Plant Sciences, College of Agriculture, Alemaya University, P.O. Box 138, Dire Dawa, Ethiopia; [2]Department of Plant Sciences, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa; *Author for correspondence (e-mail:labuscm.sci@mail.uovs.ac.za; phone: +27-51-4012715; fax: +27-51-4305692)*

Received 27 April 2004; accepted in revised form 8 June 2005

**Key words:** AFLP, Genetic diversity, Germplasm, Morphology, Sorghum, SSR/microsatellites

**Abstract.** A comparison of the different methods of the estimation of genetic diversity is important to evaluate their utility as a tool in germplasm conservation and plant breeding. Amplified fragment length polymorphism (AFLP), microsatellites or SSR and morphological traits markers were used to evaluate 45 sorghum germplasm for genetic diversity assessment and discrimination power. The mean polymorphism information content (PIC) values were 0.65 (AFLPs) and 0.46 (SSRs). The average pairwise genetic distance estimates were 0.57 (morphological traits), 0.62 (AFLPs) and 0.60 (SSRs) markers data sets. The Shannon diversity index was higher for morphological traits (0.678) than AFLP (0.487) and SSR (0.539). The correlation coefficients obtained by the Mantel matrix correspondence test, which was used to compare the cophenetic matrices for the different markers, showed that estimated values of genetic relationship given for AFLP and SSR markers, as well as for morphological and SSR markers were significantly related ($p < 0.001$). However, morphological and AFLP data showed non-significant correlation ($p > 0.05$). Both data sets from AFLP and SSR allowed all accessions to be uniquely identified; two accessions could not be distinguished by the morphological data. In summary, AFLP and SSR markers proved to be efficient tools in assessing the genetic variability among sorghum genotypes. The patterns of variation appeared to be consistent for the three marker systems, and they can be used for designing breeding programmes, conservation of germplasm and management of sorghum genetic resources.

## Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is one of the most important cereals of the semi-arid tropics. It is the third most important cereal crop after *tef* [*Eragrostis tef* (Zucc.) Trotter] and maize, and first in the eastern regions of Ethiopia, in terms of cultivation area and production (CSA 2000). Current models of sorghum race and variety distribution differentiate the main *S. bicolor* races as Bicolor, Caudatum, Durra, Guinea, and Kafir (Harlan and de Wet 1972). All of these (except Kafir) are found in Ethiopia (Stemler et al. 1977; Teshome et al. 1997), and have a broad agro-ecological variation, which has resulted in the accumulation of genetic diversity in this crop species.

Estimation of genetic diversity to identify groups with similar genotypes is important for conserving, evaluating and utilising genetic resources, for studying the diversity of different germplasm as possible sources of genes that can improve the performance of cultivars, and for determining the uniqueness and distinctness of the phenotypic and genetic constitution of genotypes with the purpose of protecting the breeder's intellectual property rights (Subudhi et al. 2002). In the past, plant breeders made selections of breeding material on the basis of morphological characteristics that were readily observable and that were co-inherited with the desired trait. Although these methods remain effective, morphological comparisons have limitations, including the influence of environment or management practices (Gepts 1993). Subjectivity in the character evaluation is also linked to developmental stage (Morell et al. 1995).

Different techniques are used to generate DNA based markers that result in different estimates of genetic similarity depending on the number of markers generated and the genome coverage. The SSR marker technique has been used to characterize genetic diversity represented by elite inbred genotypes and cultivated races of sorghum (Brown et al. 1996; Dean et al. 1999; Djé et al. 2000; Smith et al. 2000).

Although DNA markers have been compared in the assessment of sorghum genetic diversity (de Oliveira et al. 1996; Yang et al. 1996; Smith et al. 2000), both AFLPs and SSRs are more recent techniques, and have not been evaluated for use in discriminating between different sorghum accessions. The objective of the study was to compare the use of AFLP's, SSR's and morpho-agronomical traits markers to assess genetic diversity among the 45 accessions of sorghum from the eastern highlands of Ethiopia.

## Materials and methods

Forty-five sorghum accessions including 34 landraces grown by the farmers, six elite breeding entries and five improved cultivars acquired from the Alemaya University (AU), Sorghum Improvement Programme, were used (Table 1).

### *Morphological traits*

The accessions were grown at the Department of Plant Sciences Experimental field, AU, Alemaya during the 2001 growing season. For data collection the Sorghum Descriptors (IBPGR/ICRISAT 1993) was employed. Ten qualitative (plant colour, stalk juiciness, leaf midrib colour, inflorescence exsertion, panicle compactness and shape, awns, glume colour, grain covering, grain colour and endosperm texture) and 16 quantitative (days 50% flowering, leaf number, leaf length, leaf width, leaf area, internode length, leaf sheath length, plant height, panicle length, panicle width, number of primary branches panicle$^{-1}$, head weight, grain yield panicle$^{-1}$, 1000-seed weight, threshing percent and grain

*Table 1.* List of the sorghum germplasm used in the study and their collection information.

| No. | Local/cultivar name | Collection site information | | Sample status[b] |
|---|---|---|---|---|
| | | Aanaa[a]/Source | Altitude (m) | |
| 1 | Wagare 1 | Chinakssen | 1950 | LR |
| 2 | Wagare 2 | Chinakssen | 1970 | LR |
| 3 | Wagare 3 | Haro Maya | 2120 | LR |
| 4 | Muyra adi | Haro Maya | 2120 | LR |
| 5 | Muyra 1 | Kurfa Challe | 2220 | LR |
| 6 | Fandisha duudaa | Kurfa Challe | 2310 | LR |
| 7 | Fandisha faca'a | Kurfa Challe | 2050 | LR |
| 8 | Wagare 4 | Meta | 2050 | LR |
| 9 | Hamedaya | Meta | 2180 | LR |
| 10 | Muyra 2 | Meta | 2180 | LR |
| 11 | Abedelota | Deder | 2080 | LR |
| 12 | Ambajeette | Deder | 2080 | LR |
| 13 | Muyra 3 | Tulo | 2000 | LR |
| 14 | Dassile 1 | Tulo | 1900 | LR |
| 15 | Hanchiro | Tulo | 2100 | LR |
| 16 | Wagare 5 | Tulo | 1940 | LR |
| 17 | Key Fendisha | Tulo | 2140 | LR |
| 18 | Sharif | Tulo | 1960 | LR |
| 19 | Alaa guuraacha | Tulo | 1940 | LR |
| 20 | Suuta naqaaphu | Tulo | 1900 | LR |
| 21 | Qirendaye | Tulo | 2200 | LR |
| 22 | Alegid | Tulo | 1900 | LR |
| 23 | Fandisha | Doba | 1900 | LR |
| 24 | Fandisha gababa | Doba | 2230 | LR |
| 25 | Bulo | Doba | 2010 | LR |
| 26 | Janga | Doba | 2000 | LR |
| 27 | Harka basi | Doba | 2000 | LR |
| 28 | Shafare 1 | Doba | 2000 | LR |
| 29 | Shafare 2 | Chiro | 2270 | LR |
| 30 | Gababe | Chiro | 2240 | LR |
| 31 | Warabi | Chiro | 1900 | LR |
| 32 | Zangada 1 | Habro | 1940 | LR |
| 33 | Zangada 2 | Habro | 1940 | LR |
| 34 | Dassile 2 | Habro | 1900 | LR |
| 35 | ETS 721 | Alemaya University | – | BE |
| 36 | ETS 993 | Alemaya University | – | BE |
| 37 | ETS 789 | Alemaya University | – | BE |
| 38 | ETS 804 | Alemaya University | – | BE |
| 39 | Wotet begunche | Alemaya University | – | BE |
| 40 | AL–70 | Alemaya University | – | IC |
| 41 | ETS 2752 | Alemaya University | – | IC |
| 42 | Chiro | Alemaya University | – | IC |
| 43 | ETS 1005 | Alemaya University | – | IC |
| 44 | ETS 576 | Alemaya University | – | IC |
| 45 | Long muyra | Alemaya University | – | LC |

[a] Administrative unit.
[b] LR = Landrace, BE = Breeding entry, IC = Improved cultivar, LC = Local check.

number panicle$^{-1}$) traits. The heritability values were high for qualitative traits and moderately high (>60%) for most of the quantitative traits (Geleta 1997). The morphological traits data were transformed to binary data in order to compare them with AFLP and SSR markers data as described by Sneath and Sokal (1973).

### *DNA markers*

A total of 45 accessions (Table 1) were used in this study. Three to four plants were grown in 8 l size pots, containing soil, under standard glasshouse conditions at the University of the Free State, Bloemfontein, South Africa, during March through August 2001. The growth temperature was set at 14 ± 2 °C night and 28 ± 2 °C day.

Leaf material was taken from 10 plants (4–6 week old) of each accession. Single-plant samples were ground to a powder in liquid nitrogen using a mortar and pestle. A modified monocot extraction procedure (Edwards et al. 1991) was followed to isolate the DNA. Extraction buffer (10 ml) (1M Tris–HCl pH 8: 0.25 M EDTA, and 1.25% (w/v) SDS) and 1 ml (10% w/v) Cetyl triethyl ammonium bromide (CTAB) was added. The homogenate was vortexed and incubated at 65 °C for 60 min, with periodic shaking. Chloroform extraction was performed to remove cellular debris and proteins by the addition of 10 ml chloroform-isoamyl alcohol (24:1v/v) followed by centrifugation for 15 min at 10,000 rpm. Thereafter, the DNA was precipitated by the addition of two volumes of cold absolute ethanol. The precipitate was spooled using a sterile Pasteur pipette and washed twice in 70% ethanol. The DNA was dissolved in 250 $\mu$l sterile distilled water and stored at −20 °C.

The DNA was diluted to a working concentration of 100 ng/$\mu$l in sterile distilled water. Equal quantities (100 ng) of genomic DNA from 10 plants for each accession were bulked and used in AFLP and SSR analyses. Although the bulking of plants discards the intra-accession diversity, it makes it possible to assess the full extent of genetic diversity between lines, taking into account the potential genetic diversity that may exist within a line.

### *AFLP*

Genomic DNA (250 ng of the bulked DNA) was double digested with five units each of *EcoR*I and *Mse*I endonuclease, at 37 °C for 2 h. The digested DNA fragments were ligated to *EcoR*I and *Mse*I adaptors (Table 2) with T4 DNA ligase for 2 h at 20 ± 2 °C. The ligated DNA was diluted to 1:10 in TE buffer (10 mM Tris–HCl (pH 8.0), 0.1 mM EDTA) and stored at −20 °C. PCR was performed in two consecutive reactions: a pre-selective and selective PCR, following the protocol supplied by the manufacturer (GIBCO BRL). In the pre-selective reaction, genomic DNA was amplified using an AFLP primer

*Table 2.* Adaptors and primers used for pre-selective and selective AFLP amplification reactions.

| Primer/adaptor code | Sequence |
|---|---|
| Adaptors | |
| *Eco*RI adaptor | 5′-CTCGTAGACTGCGTACC-3′ |
| | 3′-CATCTGACGCATGGTTAA-5′ |
| *Mse*I adaptor | 5′-GACGATGAGTCCTGAG-3′ |
| | 3′-TACTCAGGACTCAT-5′ |
| Primer | |
| *Eco*RI primer | |
| E-AAC | 5′-GATCTGCCTACCAATTCAAC-3′ (NED) |
| E-ACA | 5′-GATCTGCGTACCAATTCACA-3′ (FAM) |
| *Mse*I primer | |
| M-CAA | 5′-GATGAGTCCTGAGTAACAA-3′ |
| M-CAT | 5′-GATGAGTCCTGAGTAACAT-3′ |
| M-CTA | 5′-GATGAGTCCTGAGTAACTA-3′ |
| M-CAG | 5′-GATGAGTCCTGAGTAACAG-3′ |

pair, each having one selective nucleotide (Table 2). Accordingly, a 5 $\mu$l diluted ligation product, 40 $\mu$l pre-amplification primer mix, 5 $\mu$l 10× PCR buffer mixed with $MgCl_2$ for AFLP and 1 unit/0.2 $\mu$l of *Taq* polymerase were mixed for the pre-selective reaction. The pre-selective reactions were performed as follows: 20 cycles of 30 s at 94 °C, 60 s at 56 °C and 2 min at 72 °C. Pre-selective PCR amplification was confirmed by gel electrophoresis and the amplified product diluted to 1:50 in TE buffer (10 mM Tris–HCl [pH 8.0] and 0.1 mM EDTA) and used as template for the selective amplification using AFLP primers, each containing three selective nucleotides.

Selective PCR amplification was performed in 20 $\mu$l reactions consisting of, 5 $\mu$l pre-selective template DNA (1:50 dilution), 4.5 $\mu$l (6.7 ng/$\mu$l, dNTPs) *Mse* primer with selected nucleotide extensions, 1 $\mu$l (1 $\mu$M) Eco primer with selected nucleotide extensions (Eco-ACA and Eco-AAC labelled with FAM and NED, respectively) (PE Biosystems), 2 $\mu$l 10× PCR buffer (200 mM Tris–HCl [pH 8.4], 15 mM $MgCl_2$, 500 mM KCl) and 0.1 $\mu$l *Taq* polymerase (5 units/$\mu$l). Selective PCR amplification reactions were performed for 35 cycles, with 30 s at 94 °C, and 30 s at 65 °C, followed by 2 min at 72 °C. The annealing temperature was lowered 0.7 °C in each subsequent cycle during the first 12 cycles down to 56 °C. All amplification reactions were performed in a PCR System 2700 (Applied Biosystems). Following selective amplification, 5 $\mu$l of amplification product was mixed with 24 $\mu$l formamide (deionised) and 1 $\mu$l GeneScan$^{TM}$ 500 ROX$^{TM}$ size standard marker (PE Biosystems), denatured at 94 °C for 10 min and quickly cooled in ice slurry and resolved according to size on a Perkin–Elmer ABI310 Automated Capillary Sequencer (PE Biosystems).

AFLP analysis was performed using GeneScan® software. Only clear and unambiguous bands were included in the analysis. AFLP fragments larger than or equal to 60 bp with a peak height above or equal to 45 reflective fluorescent units (RFUs) were scored. A visual comparison was used to correlate the binary output of electropherograms.

*Microsatellites (SSR)*

Fifteen SSR sorghum primer pairs (Brown et al. 1996) were used in this study. Primers were excluded from the study if they failed to amplify consistently in all 45 accessions. The 10 SSR primer pairs used in the final analysis were presented in Table 3.

A standard PCR method was used to amplify microsatellites. PCR conditions were optimised for each primer pair by adapting the annealing temperature (Tm). The PCR reaction was performed by taking 0.5 $\mu$l of bulked DNA, 2.5 $\mu$l 10× PCR buffer (200 mM Tris–HCl (pH 8.4), 500 mM KCl), 0.75 $\mu$l $MgCl_2$(50 mM), 0.5 $\mu$l dNTP's (40 mM), 0.2 $\mu$l *Taq* polymerase (5 units/$\mu$l) and 18.25 $\mu$l sterile distilled water in a total reaction volume of 25 $\mu$l. PCR cycling conditions were: 2 min initial denaturation at 95 °C; followed by 30 cycles of 30 s at 94 °C, 45 s at either 45 °C (Sb4-32, sb5-85 and sb6-84), 50 °C (sb4-22) or 55 °C (sb1-10, sb4-15, sb5-236, sb6-36, sb6-57 and sb6-342) and 1 min elongation at 72 °C, followed by a final elongation of 10 min at 72 °C. All PCR reactions were performed on a PCR System 2700 (Applied Biosystems).

The result of the PCR amplification was analyzed by electrophoresis on MS agarose (Roche, 2% gels) designed to separate high resolution PCR products such as SSRs for resolution of fragments ranging from 50 to 1000 bp in TAE

*Table 3*. Summary of the SSR-primer pairs used in the study.

| SSR Locus | Primer sequence | Repeat motif | Linkage group | Size range (bp)[1] |
|---|---|---|---|---|
| Sb1-10 | F:GTGCCGCTTTGCTCGCA<br>R:TGCTATGTTGTTTGCTTCTCCCTTCTC | $(AG)_{27}$ | D | 242–488 |
| Sb4-15 | F:GCTGCTAAGCCGTGCTGA<br>R:TTATTTGGGTGAAGTAGAGGTGAACA | $(AG)_{16}$ | E | 120–134 |
| Sb4-22 | F:TGAGCCGAAAACCGTGAG<br>R:CCCAAAACCAAGAGGGAAGG | $(ACGAC)_4(AG)_6$ | na | 270–300 |
| Sb4-32 | F:CTCGGCGGTTAGCACAGTCAC<br>R:GCCCATAGACAGACAGCAAAGCC | $(AG)_{15}$ | E | 160–216 |
| Sb5-85 | F:AGACGCTTTTCTCTCTCTCTCTCTCTCT<br>R:TAGCCCTGCCGCATACTGAATG | $(AG)_{12}$ | na | 200–225 |
| Sb5-236 | F:GCCAAGAGAAACACAAACAA<br>R:AGCAATGTATTTAGGCAACACA | $(AG)_{20}$ | G | 162–222 |
| Sb6-36 | F:GCATAATGACGGCGTGCTC<br>R:CTTCCAAGTGAAAGAAACCATCA | $(AG)_{19}$ | C | 155–199 |
| Sb6-57 | F:ACAGGGCTTTAGGGAAATCG<br>R:CCATCACCGTCGGCATCT | $(AG)_{18}$ | C | 283–313 |
| Sb6-84 | F:CGCTCTCGGGATGAATGA<br>R:TAACGGACCACTAACAAATGATT | $(AG)_{14}$ | F | 170–212 |
| Sb6-342 | F:TGCTTGTGAGAGTGCCTCCCT<br>R:GTGAACCTGCTGCTTTAGTCGATG | $(AC)_{25}$ | A | 250–320 |

[1] Data from Brown et al. (1996), Dean et al. (1999), and Ghebru et al. (2002).
na = not available.

buffer (40 mM Tris–acetate, 1 mM EDTA, pH 8.0) run at 80 V for 2.5 h. Amplified fragments were visualised and sized using the Gel Doc 1000$^{TM}$ image analysis system (Biorad) after ethidium bromide (0.5 $\mu$g/ml) staining. The presence or absence of each fragment was coded as 1 or 0, respectively, and scored in a binary data matrix. The PIC was determined according to the formula described in Smith et al. (2000):

$$\mathrm{PIC} = 1 - \sum_{i=1}^{n} (\mathrm{f}_i)^2$$

where $\mathrm{f}_i$ is the frequency of the $i$th allele carried by the population, calculated for each SSR locus. PIC values ranged from 0 (monomorphic) to 1 (highly discriminative). The PIC values provide an estimate of the discriminatory power of a marker by taking into account not only the number of alleles at a locus, but also the relative frequencies of those alleles in the population under study.

### *Statistical analyses*

PIC, Shannon index, genetic distance, cluster analyses, and the Mantel test procedures were performed on the binary data. Shannon index was estimated using POPGENE (version 1.2, Yeh and Boyle 1997). NCSS statistical analysis software (NCSS 2000) was used to estimate genetic distance (Euclidean method) for all possible pairwise comparisons between accessions. Matrices of Euclidean dissimilarity coefficients based on morphological, AFLP and SSR data sets were tested for correlation (1000 permutation) using the Mantel test (Liedloff 1999). Cluster analysis (Unweighted Pair Group Method using Arithmetic Averages) on the similarity indices (Jaccard's coefficient) and principal coordinate analysis (PCA) was performed to identify genetic variation patterns among the sorghum genotypes using NTSYSpc (version 2.11s; Rohlf 2000–2003).

## Results and discussion

The levels of polymorphism detected by the different marker approaches showed large differences (Table 4). The eight AFLP primer pairs used generated 649 polymorphic bands (average 81 pair$^{-1}$), and the 10 microsatellite loci used produced 43 polymorphic bands (4.4 pair$^{-1}$) across 45 sorghum accessions. The percentage of polymorphic bands was slightly higher in SSR (90%) than AFLP (85%) data (data not shown). However, the AFLP technique was 14 times more efficient in detecting polymorphism per assay, as it has a higher multiplex ratio. This probably resulted from the use of agarose gels for allele detection in SSRs, since sequences differing by 2 bp could not be resolved using

*Table 4*. Number of polymorphic markers, average and range of pairwise genetic distance, and Shannon index estimates among 45 sorghum accessions.

| Marker system | Number of polymorphic markers | Genetic distance | | Shannon index[a] |
|---|---|---|---|---|
| | | Average | Range | |
| Morphological | 96 | 0.566 | 0.354–0.707 | 0.678 |
| AFLP | 552 | 0.615 | 0.413–0.745 | 0.487 |
| SSR | 43 | 0.604 | 0.152–0.762 | 0.539 |
| Combined | 691 | 0.609 | 0.426–0.723 | 0.495 |

[a] Conducted using the function of Nei (1972) (Yeh and Boyle 1997).

agarose gels (Senior et al. 1998). Acrylamide gels can resolve nucleotide differences of one base pair.

The average PIC values among the 45 accessions for AFLP and SSR markers were 0.464 and 0.645, respectively (data not shown). Polymorphism detection efficiency among sorghum accessions by AFLP's and SSR's compared favourably with other available marker systems. Dje et al. (1999) reported a higher degree of polymorphism in microsatellite markers than allozymes. In addition, Yang et al. (1996) detected 55, 25, 44% polymorphic bands for RFLP, RAPD, and ISSR techniques, respectively, in a selection of 34 Chinese sorghums. For sorghum accessions from Ethiopia and Eritrea, it was reported that morphological traits have shown a higher level of variation than those obtained for allozymes and RAPD markers (Ayana 2001).

The genetic distances (GDs) were estimated using 649 polymorphic AFLP fragments, 43 SSR polymorphic alleles, and 26 morpho-agronomical traits with 96 variants. The GD indexes calculated from morphological traits (0.57), AFLPs (0.62), SSRs (0.60) and combined data (0.61) showed that both AFLPs and SSRs were slightly more efficient than morphological traits in detecting variability (Table 4). Range-wise, AFLP data produced lower (0.413–0.745) GD estimates compared to SSR (0.152–0.762), but the average GD's were very close in both. GD estimates can be affected by several factors such as, the distribution of markers in the genome (genome coverage) and the nature of evolutionary mechanisms underlying the variation measured (Powell et al. 1996). AFLPs are believed to detect mainly point mutations while SSRs are specific to hypervariable loci (Giancola et al. 2002). Another important factor is the influence of individual loci used for the analysis. While the SSR loci were based on availability, AFLP loci were randomly distributed whereas morphological traits were selected.

The PIC values for SSR loci ranged from 0.52 for sb4-32 and sb6-342 loci to 0.79 for sb4-22, while one primer pair (sb4-15) was found to be monomorphic (Table 5). The mean PIC value was 0.645. Smith et al. (2000) reported similar observations of PIC in sorghum detected by SSR markers. In this study, however, no significant correlation was detected between the repeat number and the allele number ($r = -0.31$) or polymorphic information content

*Table 5*. Number of alleles, size range (in base pairs), and PIC value for SSR loci found in 45 accessions of sorghum.

| SSR locus | Repeat number[a] | No. of alleles | Size range | PIC value |
|---|---|---|---|---|
| Sb1-10 | $(AG)_{27}$ | 4 | 110–510 | 0.717 |
| Sb4-15 | $(AG)_{16}$ | 1 | 135 | 0.000 |
| Sb4–22 | $(ACGAC)_4/(AG)_6$ | 6 | 254–492 | 0.793 |
| Sb4-32 | $(AG)_{15}$ | 9 | 112–533 | 0.521 |
| Sb5-85 | $(AG)_{12}$ | 7 | 109–510 | 0.594 |
| Sb5-236 | $(AG)_{20}$ | 4 | 110–265 | 0.578 |
| Sb6-36 | $(AG)_{19}$ | 5 | 165–279 | 0.773 |
| Sb6-57 | $(AG)_{18}$ | 4 | 292–343 | 0.683 |
| Sb6-84 | $(AG)_{14}$ | 3 | 173–295 | 0.630 |
| Sb6-342 | $(AC)_{25}$ | 5 | 254–634 | 0.520 |
| Average | | 4.8 | | 0.645 |

[a] From Brown et al. (1996).

($r = -0.07$). The 43 polymorphic SSR alleles collectively yielded unique genotypes for each of the 45 accessions. A wide range of fragment sizes (differences between the shortest and longest alleles ranged from 51 to 421 bp) was obtained, with most within the ranges previously reported in studies with different sorghum germplasm (Dean et al. 1999; Djè et al. 2000; Ghebru et al. 2002).

The Shannon diversity index estimates obtained were 0.678 (morphological), 0.487 (AFLP) and 0.539 (SSR) data sets. The result of a Mantel test to compare similarity of matrices for different marker types showed that estimated values of genetic relationship given for AFLP and SSR markers, and SSR and morphological markers were significantly related (Table 6). However, the AFLP markers showed non-significant correlation with morphological traits. In a similar study, Tatineni et al. (1996) reported a high correlation between RAPD and morphological characters. However, a lower correlation between AFLP and SSR genetic distance estimates has been reported (Giancola et al. 2002). Powell et al. (1996) also reported that SSR similarity estimates were not significantly correlated to RFLP's, RAPD's, or AFLP's in soybean.

Dendrograms resulting from UPGMA cluster analyses of morphological, AFLP, SSR, and combined data are shown in Figure 1a–d, respectively. All the dendrograms, except morphological traits that failed to distinguish between two accessions, clearly discriminated the 45 sorghum accessions. Though the

*Table 6*. Mantel test values estimated for the three marker techniques, with sample size of 990.

| | AFLPs | SSRs |
|---|---|---|
| SSRs | $0.28^{**}$ | |
| Morphological traits | $0.08^{NS}$ | $0.19^{**}$ |

** Significant at $p < 0.001$ and NS non-significant at $p > 0.05$.

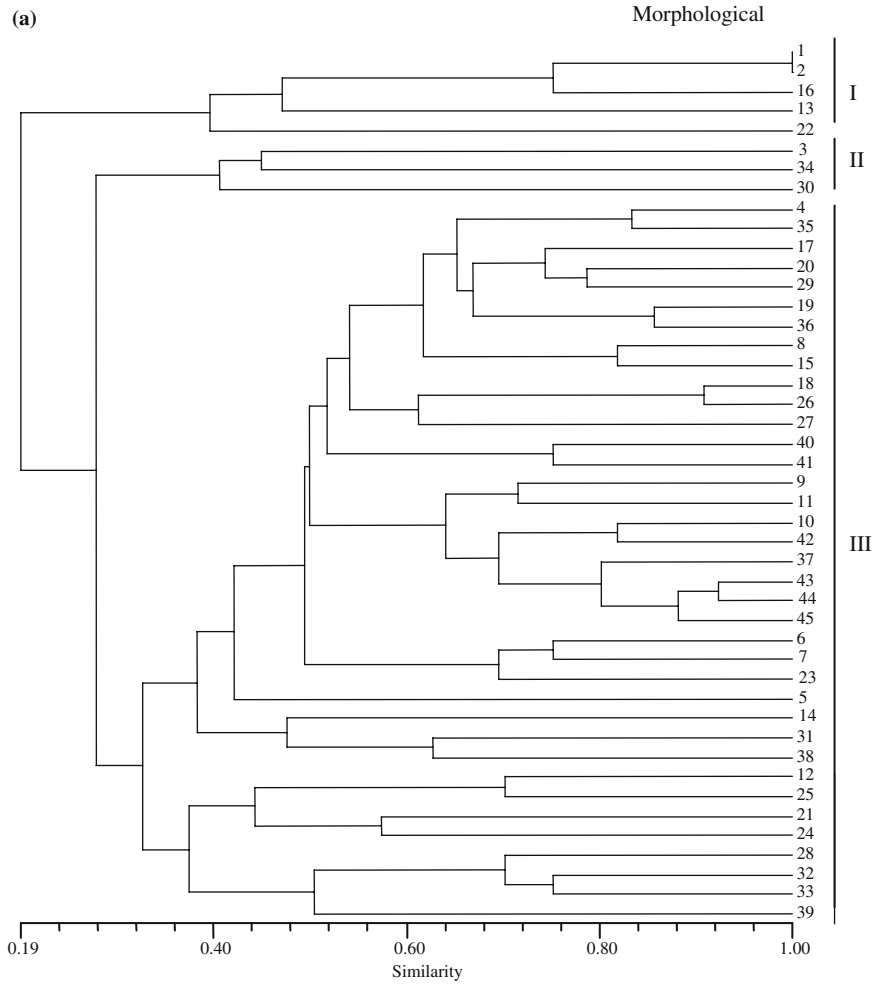


*Figure 1.* Dendrograms of 45 sorghum accessions showing the genetic similarity based on morphological traits (a), AFLP (b), SSR (c) and combined (d) data sets using Jaccard's coefficient and UPMGA cluster analysis. Accession numbers as shown in Table 1.

output of each tree was rather unique, there are sufficient similarities among the three dendrograms, for instance in all dendrograms accessions no. 4, 5, 6 and 7 were consistently grouped together. Clustering based on morphological data analysis produced three major clusters (Figure 1a). Accessions 1 and 2 were not resolved into their individual branches, which shows the inadequacy of the morphological traits. Cluster III contained the largest number of accessions from different localities.

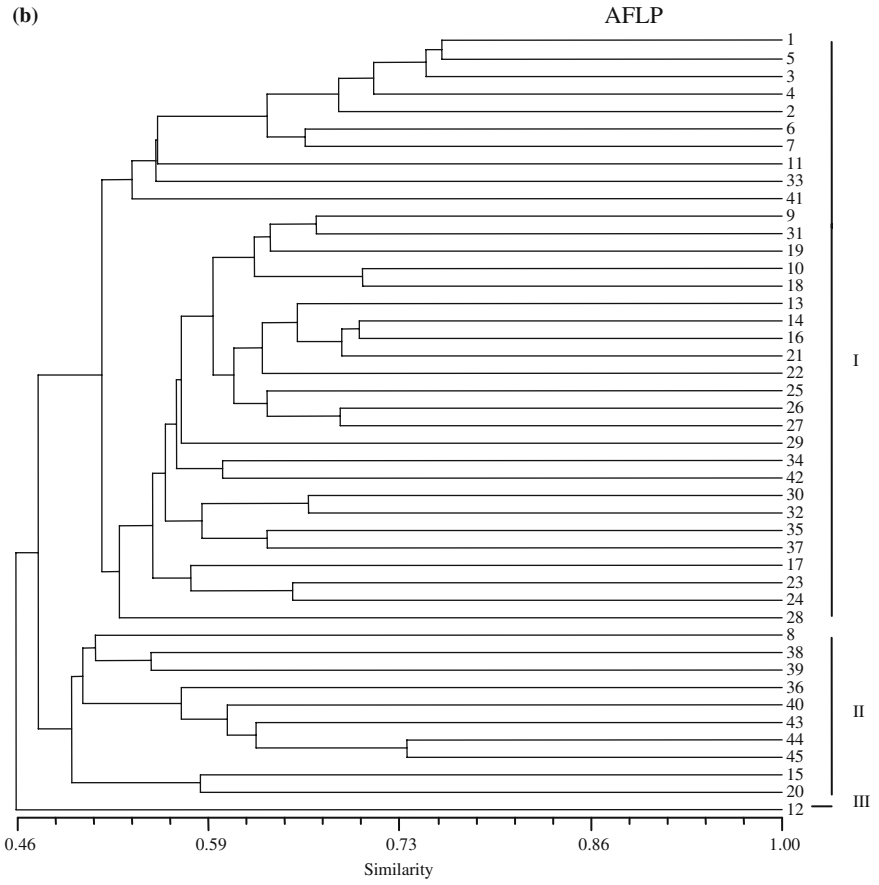


*Figure 1.* Continued.

In the AFLP based dendrogram (Figure 1b) three main clusters were also formed, where the first cluster constituted the largest number of accessions coming from various localities. These three clusters were identified at the 0.48 similarity level. In the first cluster, two subgroups were formed, where accession 41 was the most distinct in the first subgroup and accession 28 in the second. In cluster II, two subclusters were identified. Seven of the 10 accessions from AU constituted the first subcluster. Cluster III consisted of the single accession no. 12. In the dendrogram constructed from SSR markers (Figure 1c), four main clusters were identified at 0.33 similarity level. The highest association was found within cluster IV between accessions no. 36 and 38 at 0.97 similarity level. It was observed that there was more grouping based on the site of collection than observed in the morphological and AFLP based clustering. For instance, accessions from Meta and Kurfa Challe formed close relationships within cluster II and cluster III,

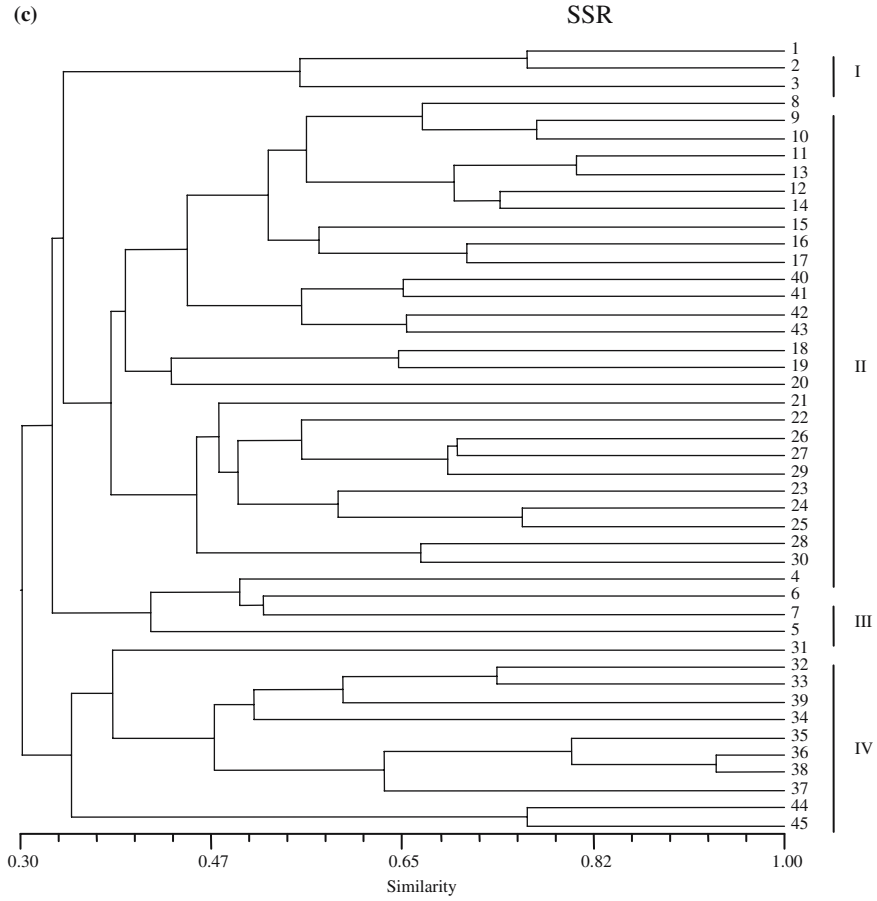


*Figure 1.* Continued.

respectively. The AU accessions predominantly constituted cluster IV while the other were positioned close to each other in cluster II. A clustering performed on the combined data is shown in Figure 1d. The first cluster contained a large group of accessions from various localities. Accessions 41 and 12 were found to be different in their respective cluster. In this study, with a few exceptions, no clear tendency of clustering was observed based on the accession names or altitude of collection.

The PCA plots (Figure not shown) obtained with the three data sets also showed clustering similar to that of the dendrogram, although this was more clear for the SSR data set. The consistent clustering of most breeding entries/improved cultivars from AU close to each other in the present study apparently substantiates that the marker systems used have a high potential in quantifying the level of similarity and relationships among sorghum

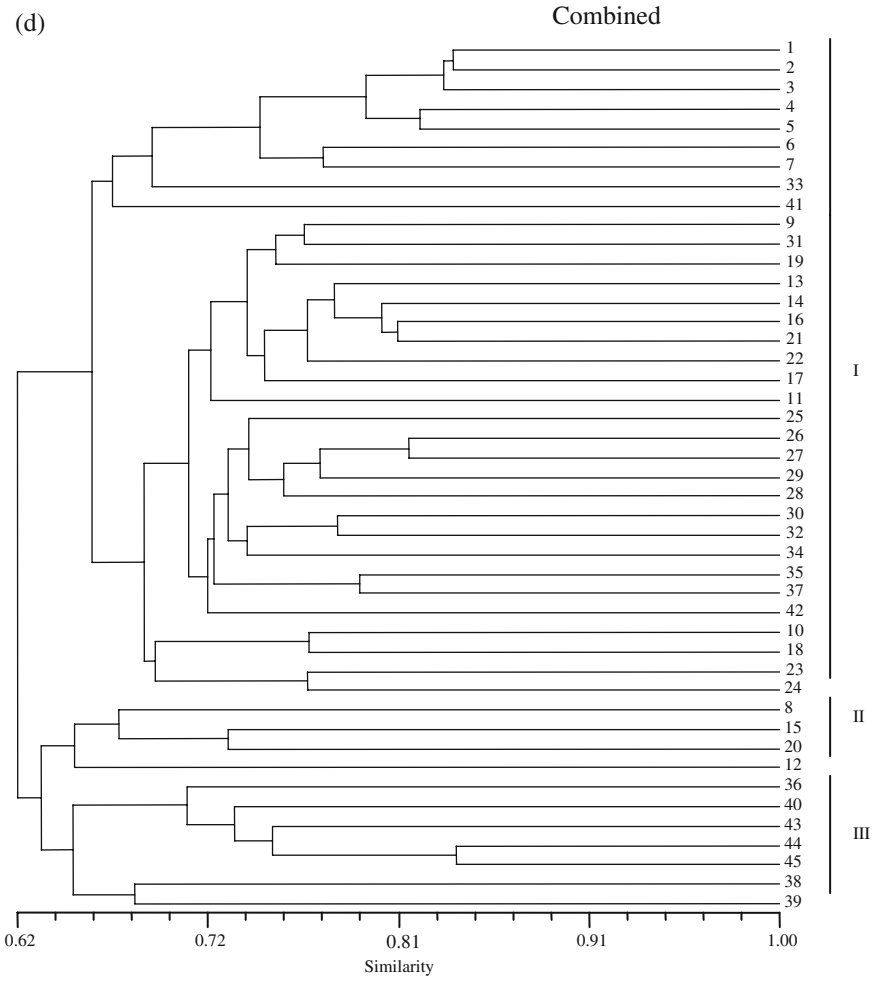


*Figure 1.* Continued.

germplasm. Furthermore, the results showed that by using the AFLP or SSR technique, a large set of informative data could be generated in less time than with morphological traits. Also when simultaneously using DNA markers and morphological traits to classify genotypes, it is possible to obtain a relevant minimum subset of marker-fragments that can be used in conjunction with available morphologic data to better classify genotypes compared to using only the quantitative or only the qualitative traits. The high genetic diversity value among the sorghum accessions (landraces, breeding entries and improved cultivars) indicates that the level of genetic diversity was not influenced by breeding activities.

## Conclusions

Detailed studies on genetic diversity in germplasm can be performed by studying morphological traits or by employing marker systems such as allozymes, RFLPs, RAPDs, AFLPs or SSRs. Examples are known for many species, including cereals. The AFLP and SSR markers used in this study were polymorphic, and can be used to distinguish accessions of sorghum. In addition, this study indicated that although morphological characterisation is influenced by the environment and is time consuming in general, it can still be an important and practical means of making progress in sorghum germplasm evaluation. Morphological traits based data has shown significant correlation with SSR, but no correlation with AFLP based data sets. The low similarity value among the majority of sorghum accessions could indicate that there is a high level of genetic diversity among the test materials for these markers systems. Both AFLP and SSR markers can be utilized as a method of choice for revealing genetic variation and identifying slightly different genotypes in a sorghum breeding program.

## Acknowledgements

This research was supported by a Grant from the Agricultural Research and Training Project (ARTP) through the Alemaya University, Ethiopia. The Institute of Biodiversity Conservation and Research (IBCR)/Ethiopia is gratefully acknowledged for giving permission to take the sorghum germplasm used for the study in South Africa.

## References

Ayana A. 2001. Genetic diversity in sorghum (*Sorghum bicolor* (L.) Moench) germplasm from Ethiopia and Eritrea. Ph.D. dissertation. Department of Biology, Science Faculty, Addis Ababa University, Addis Ababa, Ethiopia.

Brown S.M., Hopkins M.S., Mitchell S.E., Senior M.L, Wang T.Y., Duncan R.R., Gonzalez-Candelas F. and Kresovich S. 1996. Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Genet. 93: 190–198.

CSA 2000. Central Statistical Authority, the Federal Democratic Republic of Ethiopia: agricultural sample survey report on area and production for major crops. Statistical Bulletin No 227, Addis Ababa, Ethiopia, pp. 16–68.

De Olivera A.C., Richter T. and Bennetzen J.L. 1996. Regional and racial specificities in sorghum germplasm assessed with DNA markers. Genome 39: 579–587.

Dean R.E., Dahlberg J.A., Hopkins M.S., Mitchell S.E. and Kresovich S. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39: 1215–1221.

Djè Y., Forcioli D., Ater M., Lefebvre C. and Vekemans X. 1999. Assessing population genetic structure of sorghum landraces from north-western Morocco using allozyme and microsatellite markers. Theor. Appl. Genet. 99: 57–163.

Djè Y., Heuertz M., Lefébvre C. and Vekemans X. 2000. Assessment of genetic diversity within and among germplasm accessions in cultivated sorghum using microsatellite markers. Theor. Appl. Genet. 100: 918–925.

Edwards K., Johnstone C. and Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res. 19: 1349–1458.

Geleta N. 1997. Variability and association of morpho-agronomic charatcters with reference to highland sorghum [*Sorghum bicolor* (L.) Moench] landraces of Haraghe, eastern Ethiopia. M.Sc. Thesis. Alemaya University, Alemaya, Ethiopia.

Gepts P. 1993. The use of molecular and biochemical markers in crop evaluation studies. In: Hecht MK (ed), Evolutionary Biology Vol. 27. Plenum Press, New York, pp. 51–94.

Ghebru B., Schmidt R.J. and Bennetzen J.L. 2002. Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor. Appl. Genet. 105: 229–236.

Giancola S., Marcucci Poltri S., Lacaze P. and Hopp H.E. 2002. Feasibility of integration of molecular markers and morphological descriptors in a real case study of a plant variety protection system for soybean. Euphytica 127: 95–113.

Harlan J.R. and de Wet J.M.J. 1972. A simplified classification of cultivated sorghum. Crop Sci. 12: 172–176.

IBPGR/ICRISAT. 1993. Descriptors for Sorghum [*Sorghum bicolor*(L) Moench]. International Board of Plant Genetic Resources. Rome, Italy/International Crop Research Institute for Semi-Arid Tropics, Patancheru, India.

Liedloff A. 1999. Mantel Version 2.0, Mantel-nonparametric test calculator. http://www.Sci.qut.edu.au/NRS/mantel.htm.

Morell M.K., Peakall R., Appels R., Preston L.R. and Lloyd H.L. 1995. DNA profiling techniques for plant variety identification. Aust. J. Exp. Agric. 35: 807–819.

NCSS 2000. Number Cruncher Statistical Systems. Dr. Jerry L. Hintze, 329 North 1000 East, Kaysville, Utah 84037, Canada.

Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Dingey C. and Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2: 225–238.

Rohlf F.J. 2000–2003. Numerical Taxonomy and Multivariate Analysis System, Version 2.11s. Applied Biostatistics, New York.

Senior M.L., Murphy J.P., Goodman M.M. and Stuber C.W. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci. 38: 1088–1098.

Smith J.S.C., Kresovich S., Hopkins M.S., Mitchell S.E., Dean R.E., Woodman W.L., Lee M. and Porter K. 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci. 40: 226–232.

Sneath P.H.A. and Sokal R.R. 1973. Numerical Taxonomy. The principles and practice of Numerical Classification, XV. Freeman, San Francisco, CA, pp. 573.

Stemler A.B.L., Harlan J.R. and de Wet J.M.J. 1977. The sorghums of Ethiopia. Econ. Bot. 31: 446–460.

Subudhi P.K., Nguyen H.T., Gilbert M.L. and Rosenow D.T. 2002. Sorghum improvement: past achievements and future prospects. In: Kang M.S. (ed), Crop improvement: Challenges in the Twenty-First Century. The Haworth Press, Binghamton, New York, pp. 109–159.

Tatineni V., Cantrell R.G. and Davis D.D. 1996. Genetic diversity in elite cotton germplasm determined by morphological characters and RAPDs. Crop Sci. 36: 186–192.

Teshome A., Baum B.R., Fahrig L., Torrance J.K., Arnason T.J. and Lambert J.D. 1997. Sorghum [*Sorghum bicolor* (L.) Moench] landrace variation and classification in North Shewa and South Welo, Ethiopia. Euphytica 97: 255–263.

Yang W., de Oliveira A.C., Godwin I., Schertz K. and Bennetzen J.L. 1996. Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. Crop Sci. 36: 1669–1676.

Yeh F.C. and Boyle T.J.B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits, Version 1.2. Belgian J. Bot. 129: 157.