J. Plant Biochem. Biotechnol. (Jan–June 2011) 20(1):74–83
DOI 10.1007/s13562-010-0028-z

ORIGINAL ARTICLE

Identification of unique alleles and assessment of genetic diversity of *rabi* sorghum accessions using simple sequence repeat markers

Mahendar Thudi · Bashasab Fakrudin

Received: 23 February 2010 / Accepted: 5 October 2010 / Published online: 29 January 2011

Abstract Genetic diversity among 42 sorghum accessions representing landraces (19), advanced breeding lines (16), local cultivars (2) and release varieties (5) with 30 simple sequence repeat (SSR) markers revealed 7.6 mean number of alleles per locus showing 93.3% polymorphism and an average polymorphism information content of 0.78 which range from 0.22 (*Xtxp*12) and 0.91(*Xtxp*321). The average heterozygosity and effective number of alleles per locus were 0.8 and 6.65 respectively. Cluster analysis based on microsatellite allelic diversity clearly demarcated the accessions into ten clusters. A total of 24 unique alleles were obtained from seven SSR loci in 23 accessions in a size range of 110–380 bp; these unique alleles may serve as diagnostic tools for particular region of the genome of respective genotypes. Selected SSR markers from different linkage groups provided an accurate way of determining genetic diversity at the molecular level.

Keywords Polymorphism information content · Simple Sequence Repeats (SSR) · Fingerprint · Heterozygosity · Effective number of alleles · Landrace

B. Fakrudin (✉)
Institute of Agri-Biotechnology,
University of Agricultural Sciences, Dharwad,
Krishinagar,
Dharwad 580 005 Karnataka, India
e-mail: bfakrudin@gmail.com

Present Address:
M. Thudi
International Crops Research Institute for the Semi-Arid Tropics,
Patancheru,
502324, Hyderabad, India

Abbreviations

MDS	multidimensional scaling
PIC	polymorphism information content
CTAB	cetyl trimethyl ammonium bromide
UPGMA	unweighted pair-group method with arithmetic average
SSR	simple sequence repeats

Introduction

Sorghum [*Sorghum bicolor* (L.) Moench] is fifth most important C4 cereal species next only to wheat, rice, maize and barley. It is predominantly a self-pollinating member of gramineae, with a basic chromosome number ten ($2n=2x=20$) and with a nuclear DNA content of 1.6 pg and genome size of 735 Mbp (Dillon et al. 2007). It is cultivated on 44 mha in 99 countries and is it a staple food crop for over 500 million people in Africa, Asia, Oceania, and the Americas. About 21% (9.2 mha) of world's sorghum area is in India. It is cultivated on 5.0 mha during *kharif* (rainy) season and 4.2 mha during *rabi* (winter) season (FAOSTAT 2006). The *rabi* belt in India extends through the states of Karnataka, Maharashtra, and Andhra Pradesh. Generally *rabi* sorghum in these regions is grown on typical deep black soils and the crop experience terminal drought stress due to receding soil moisture. Most of these regions experience either long or short term moisture stress or high temperature stress. Further, there is a drastic decline in the cultivated area and production of sorghum from 6.8 mha (4.0 mt) during 1964–65 to 5.0 mha (2.9 mt) in 2002–03 (http://icrtest:8080/sorghum/sorghum.htm). Such a drastic decline in area and production can be attributed to low soil fertility, rain-fed farming characterized by erratic and

inadequate rainfall, negligible external inputs, continued use of mostly unimproved cultivars with a low harvest index (<20%), and prevalence of diseases and insect pests. Therefore, improvement of *rabi* sorghum productivity has a great impact on socio-economic status of the sorghum growing regions in India.

Effective plant breeding and crop improvement program depend on the availability of diverse genetic resources. Evaluation of germplasm diversity could help to identify genotypes with great novelty and thus be useful in crop improvement program in achieving both short and long-term goals. There have been several reports of assessing patterns of genetic variation mainly based on morphology (Appa-Rao et al. 1996; Dje et al. 1998), pedigree (Jordan et al. 1998) or combination of agronomic and morphological traits in sorghum (Harlan and Dewet 1972). The utility of isozymes (Aldrich et al. 1992), restriction fragment length polymorphism (RFLP) (Aldrich and Doebley 1992), random amplified polymorphic DNA (RAPD) (Ayana et al. 2000) and simple sequence repeat (SSR) (Abu-Assar et al. 2005) markers have been attempted in *kharif* types of sorghum germplasm. However, application of such approaches in *rabi* type of germplasm are scanty to our knowledge. SSR markers have been found to be more reliable and useful over others owing to their inherent merits and abundance (Varshney et al. 2005). In sorghum and maize, compared to RAPD (Vierling et al. 1994) and RFLP (Tao et al. 1993), SSRs markers have been shown to be highly useful to study genetic diversity (Uptmoor et al. 2003; Dillon et al. 2005).

A comprehensive genetic diversity analysis of local cultivars and landraces of *rabi* types is highly desired in India which would in turn help in deploying the genetic variation in plant breeding programmes. The current investigation was aimed at unraveling the extent of diversity among a set of selected 42 *rabi* germplasm lines comprising of local cultivars, advanced breeding lines and landraces of Karnataka and Maharashtra regions of Southern India using SSR markers and to integrate the results in developing mapping populations and selection of parents in deriving segregating material for *rabi* sorghum improvement.

Materials and methods

Plant material

A set of 42 *rabi* sorghum germplasm that represents five released varieties, 16 advanced breeding lines, two local cultivars and 19 selected landraces were examined for the extent of existing genetic diversity. The pedigree information and salient features of these accessions are presented in the Table 1.

DNA extraction and normalization

The total genomic DNA was isolated from bulk of five individual plants from each entry. Three to four week old seedlings were taken from field, lyophilized and tissues were stored at −80°C till use. Genomic DNA was extracted following CTAB mini preparation (Mace et al. 2003).

PCR amplification and microsatellite assay

A set of 30 microsatellite loci that span all the ten linkage groups (Table 2) of sorghum were chosen for genotyping 42 sorghum germplasm accessions that include 41 *rabi* sorghum genotypes and one parental line (C-43) used in developing sorghum hybrid (CSH 16) which is predominantly grown in *kharif* season. The sequence information of these markers was obtained from Bhattramakki et al. (2000). The PCR amplification reactions for all 30 microsatellite loci on the selected genotypes set was performed in 25 µl reaction volume: 2 µl (25 ng/µl) DNA template, 1 µl of 2.5 mM dNTP mix (Bangalore Genei Pvt Ltd), 2.5 µl 10X PCR assay buffer (Bangalore Genei Pvt Ltd), 1 µl each of forward and reverse primer (4 pM/µl Bangalore Genei Pvt Ltd), 0.1U of *Taq* polymerase (Bangalore Genei Pvt Ltd) and made up to 25 µl with MilliQ water. Amplifications were performed in 96 well thin wall polycarbonate microtitre plates (Corning Inc.) in a Master Cycler gradient (5331-eppendorf versions 2.30. 31-09, Germany) with initial denaturation at 94°C for 5 min followed 34 cycles of denaturation at 94°C for 1 min, primer annealing at 50–60°C depending on the primer (Table 2) for 1 min and primer extension at for 2 min. A final extension at 72°C for 10 min was given at the end of the cycles and samples were held at 4°C until retrieval.

Separation and visualization of PCR amplicons

Initially the PCR products of all markers were tested for amplification on 2% agarose gel. Further, the amplified products of all 30 markers were resolved on 6% polyacrylamide gels using Sequi-Gen® GT nucleic acid electrophoresis cell (Bio-rad Pvt. Ltd, India). The amplicons were visualized by silver staining procedure (Tegelstrom 1992).

Data analysis

The clear and unambiguous alleles produced by all the micosatellite loci were scored for the presence as ‘1’ and absence as ‘0’, for each corresponding allele among the

Table 1 Information on pedigree, source and distinguishing characters of the *rabi* sorghum germplasm

Genotypes	Pedigree or source	Distinguishing character
Hybrids		
SPV-570	5-4-1 × SB40, Parbhani	Hybrid variety has good fodder quality, creamy and lustrous seed, and high yielding lines with 120 days maturity. Promising restorer line on Milo cytoplasm
SSV-74	Released from UAS, Dharwad	Hybrid variety, has good fodder quality, semi-compact elliptic panicle
SPV-1359	Land race selection made from *rabi* genotype Dhulia Released from Rahuri Agri'l University, Maharashtra	High yielding hybrid variety, susceptible to drought
SPV-86	R24 × R16; released from National Research Center For Sorghum	High yielding hybrid variety, charcoal rot susceptible
DSV-4	SPV 86 × E 36-1, released from UAS, Dharwad	Charcoal rot tolerant and high yielding genotype
Mugthi	M 35-1 × Viramgaon local	High yielding released variety
A and B lines		
A-1	M 35-1 × GS-56-1-1, Annigeri	Male sterile line
M31-2A	Raichur mutant from M35-1	Male sterile line
SB401-A	SB 1066 × Pulgar White, UAS, Dharwad	Male sterile line
104-A	296B × Swati, ARS, Mohol	Cytoplasm male sterile line, shoot fly resistant, high seed protein
104-B	Milo	Longer panicle and shoot fly tolerant
SB 401-B	Milo cytoplasm	Promising B line on Milo cytoplasm
M31-2B	Maldandi cytoplasm	High yielding, bold seed matures at 130 days
Landraces and local collections		
Doodmogra	CRS, Solapur	Landrace
Barsizoot	Barshi local, Maharashtra	Landrace, high iron content in the grain
Yangir local	Gadag, Karnataka	Landrace
Harnidagad	CRS, Solapur	Landrace
Basavanapada	Dharwad, Karnataka	Extra early, landrace, high number of trichomes
Pop sorghum	NA	Landrace, small, pearly, completely covered grain
Sakkari jola	Bijapur, Karnataka	Landrace, very small lustrous grain
Kadabin jola	Bijapur, Karnataka	Landrace, yellow colored grain
Kagimothi	Bijapur, Karnataka	Landrace, red colored grain
Kodamuthy	NA	Landrace, susceptible to head smut
Belagrinda	Bagalkot, Karnataka	Landrace, shoot fly resistant, compact oval panicle
Y-4	Annegari, Karnataka	Landrace, tall growing (255 cm)
Chittapur local	Gulbarga, Karnataka	Landrace, high grain yield
Bidar local	Bidar, Karnataka	Landrace compact elliptical ear head
Hegari-1	Bellari district, Hegari, Karnataka	Landrace, resistant to shoot fly
Kandakur	NA	Landrace, susceptible to head smut
Muddihalijola	Bijapur, Karnataka	Landrace, shoot fly resistant with semi-compact elliptic panicle
Tandur local	Tandur, Karnataka	Local collection, compact elliptic inflorescence
Dagadisolapur	Dharwad, Karnataka	Landrace grown in Maharashtra region of India, susceptible to grain smut
Dagadi local	Dharwad, Karnataka	Landrace, grown in some parts of Deccan plateau
Drought tolerant/susceptible		
M35-1	Selection from Maldandi bulk	A long standing variety, good quality grain, tolerant to drought
RSLG-262	Rahuri, landrace selection, Mahatma Phule Agri'l University	Exhibits terminal drought tolerance
E36-1	Ethiopan landrace	Resistance to drought and charcoal rot, striga susceptible and staygreen

Table 1 (continued)

Genotypes	Pedigree or source	Distinguishing character
RS-29	NA	Stay-green, promising restorer donor for increased grain number
IS 22380	NA	Drought susceptible
Biotic stress resistant/susceptible		
C-43		It can grow both in *kharif* and *rabi* seasons, source of grain mould resistance
DSV-5	Selection from Natte maldandi of Gulbarga district	Dual purpose, charcoal rot resistant *rabi* sorghum variety
JP1-5	Selection from Chittapur local, Gulbarga	Bold, pearly white grain, susceptible to grain smut
Giddamaladandi	Dharwad, Karnataka	Dwarf genotype, resemble M35-1 in morphology, compact elliptic inflorescence, susceptible to smut, grown in some places of Karnataka and Maharashtra states of India

NA information not available

genotypes. Estimates of inter individual genetic similarity were obtained according to (Nei and Li 1979) as $S_{ij} = 2a_{ij}/(2a_{ij} + b_i + c_j)$, where S_{ij} is the similarity between two individuals i and j, a_{ij} is the number of bands present in both individuals i and j, b_i is the number of bands present in individual "i" but absent in individual "j", and c_j is the number of bands present in individual "j" but absent in individual "i". The resulting 42×42 similarity matrix was subjected to multi-dimensional scaling (MDS) (Kruskal and Wish 1978) to assess whether the observed molecular variation indicated any evidence of clustering among accessions. The UPGMA-based dendrogram was constructed using the NTSYS 2.1 software, version 2.1. The unweighted pair-group method with arithmetic average (UPGMA) was used to independently confirm the clustering indicated by the two-dimensional MDS plot. MDS scatter plot was prepared using the pair-wise Dice similarity coefficient values among the 42 genotypes (Rohlf 2002). Bootstrapping over loci with 100 replications was carried out to assess the strength of evidence for the branching patterns in the resulting UPGMA dendrograms using PAUP 4.0b10 (http://www.lms.si.edu/PAUP; developed by Swofford 2003).

Expected heterozygosity was computed according to Belaj et al. (2003) as $He = 1 - \sum Pi^2$

Where Pi is the frequency of the i^{th} allele

For each SSR marker, polymorphism information content (PIC) values were calculated according to (Botstein et al. 1980), which provide an estimation of the discriminatory power of a locus as follows:

$$PIC = 1 - \left[\sum_{i=1}^{n} Pi^2\right] - \left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2Pi^2Pj^2\right]$$

Effective number of alleles per locus (n_e) according to (Morgante and Oliveri 1993): $n_e = 1/\sum Pi^2$; Where Pi is the frequency of the i^{th} allele

Total number of effective alleles Ne as defined by (Pejic et al. 1998): $N_e = \sum n_e$

Results

Allelic diversity

All 30 microsatellite loci successfully produced at least one allele per accession that was used to characterize and evaluate genetic diversity. A total of 228 alleles were generated by 30 SSR loci in the germplasm analyzed The allele sizes across all microsatellite loci were in the range of 100–750 bp. The total number of alleles ranged from 2 (*Xtxp*21) to 16 (*Xtxp*318) with an average of 7.6 alleles per locus (Table 2). A wide range of heterozygosity (0.25 to 0.91) among the accessions was detected for the SSR loci examined. The average heterozygosity was 0.80. The PIC values, derived from allelic diversity and frequency among the genotypes, were not uniform for all of the SSR loci tested. The PIC values for the markers ranged from 0.22 to 0.91 with an average of 0.78. The number of effective alleles per marker ranged from 1.33 to 11.73 with an average of 6.65. The total number of effective alleles produced by the 30 SSR loci was 199.72. Among amplicon, a few primer pairs amplified orthologous loci among the tested sorghum accessions.

Genetic relationships among the selected *rabi* genotypes

The average genetic similarity among the accessions studied was 0.56 with a range of 0.41 to 0.74. The accession C-43 recorded least genetic similarity with all other accessions studied (0.41). Similarly the landraces have shown lower genetic similarity with advanced breeding lines, local cultivars and release varieties that were examined in the current study. The genetic similarity between the landraces

Table 2 Microsatellite loci, number of alleles, heterozygosity, polymorphism information content and effective number of alleles estimated based on analysis of 42 sorghum germplasm accessions

Locus ID	Repeat motif	Annealing temp. (°C)	Amplicon size (bp)	No. of alleles	Heterozygosity	PIC Value	n_e^*	LGΔ
*Xtxp*326	$(GT)_{12}$	55	440–450	3	0.66	0.59	2.94	–
*Xtxp*63	$(GA)_{24}$	55	100–115	9	0.90	0.89	9.64	B
*Xtxp*205	$(AG)_{12}$	55	200–215	4	0.71	0.66	3.44	C
*Xtxp*343	$(AGT)_{21}$	55	250–320	8	0.88	0.86	8.05	D
*Xtxp*235	$(TC)_{19}$	55	220–245	7	0.87	0.85	7.59	–
*Xtxp*230	$(GA)_{28}$	55	170–205	10	0.89	0.88	9.44	F
*Xtxp*141	$(GA)_{23}$	55	200–208	4	0.75	0.70	3.98	G
*Xtxp*321	$(GT)_4$ $(AT)_6$ $(CT)_{21}$	55	100–130	5	0.75	0.71	3.93	H
*Xtxp*265	$(GAA)_{19}$	55	150–380	5	0.76	0.71	4.12	I
*Xtxp*9	$(TG)_{12}TT(TG)_{14}(AG)_{13}$	55	162–285	8	0.82	0.79	5.62	C
*Xtxp*86	$(AG)_{13}(GA)_{16}$	55	265–320	4	0.68	0.61	3.14	B
*Xtxp*88	$(AG)_{31}$	53	145–750	13	0.90	0.89	10.10	A
*Xtxp*318	$(AGA)_{12}$	55	400–600	16	0.90	0.89	10.15	–
*Xtxp*8	$(TG)_{31}$	60	120–145	6	0.79	0.76	4.65	B
*Xtxp*211	$(CT)_{23}$	55	150–200	7	0.86	0.84	7.02	B
*Xtxp*283	$(TTC)_{12}$	55	110–170	10	0.89	0.88	8.87	B
*Xtxp*286	$(CTT)_{11}$ CTC $(CTT)_{16}$	55	100–125	7	0.88	0.87	8.40	B
*Xtxp*12	$(CT)_{22}$	55	105–190	13	0.90	0.90	10.31	D
*Xtxp*21	$(AG)_{16}$	60	100–120	2	0.25	0.22	1.33	D
*Xtxp*158	$(AC)_{10}$	55	110–130	3	0.56	0.49	2.26	–
*Xtxp*278	$(TTG)_{12}$	50	180–210	4	0.70	0.65	3.34	E
*Xtxp*67	$(GA)_{28}$	55	140–253	6	0.81	0.78	5.22	F
*Xtxp*258	$(AAC)_{19}$	55	555–700	8	0.87	0.86	7.68	F
*Xtxp*295	$(CTT)_{16}$ $(AGG)_6$	55	180–190	6	0.76	0.73	4.21	E
*Xtxp*276	$(GAA)_{12}$ $(GAAA)_6$ $(GAAA)_{21}$ $(GTA)_3$ $(GTA)_3$	55	215–225	9	0.87	0.86	8.00	–
*Xtxp*354	$(GA)_{21}$ $(AAG)_3$	55	200–205	9	0.88	0.87	8.13	H
*Xtxp*6	$(CT)_{33}$	50	100–130	15	0.91	0.91	11.73	F
*Xtxp*274	$(TTC)_{19}$	55	235–250	11	0.89	0.88	8.93	I
*Xtxp*65	$(ACC)_4$ $(CCA)_3$ $CG(CT)_6$	55	120–150	8	0.88	0.87	8.58	J
*Xtxp*229	(CT) $(CA)_6$ CCC $(CA)_6$	55	200–208	8	0.89	0.88	8.92	A
Average				7.6	0.80	0.78	6.65	

Δ Linkage information obtained from Bhattramakki et al. (2000)

and advanced breeding lines ranged between 0.41 and 0.47; for instance, it was 0.44 between kodamuthy and SSV-74; 0.43 between SB 401-A and popsorghum.

The dendrogram constructed using Dice similarity coefficient and UPGMA clustering grouped all the 42 sorghum germplasm accessions into ten clusters (Fig. 1). Cluster III (Cl III) was largest cluster with 10 genotypes followed by Cl I and Cl II with 7 accessions; Cl VI and Cl VII with three accessions each. However, Cl VIII and Cl X comprised of only one accession SB401-A and C-43 respectively while Cl IX comprised of two accessions (SB401-B and Harinidagad). All the examined landraces were distributed among all the clusters except Cl VIII and Cl X. However, among the 19 examined landraces, a total of 9 (47%) landraces were grouped into Cl III. The accession C-43, cultivated both during *kharief* and *rabi* seems to be distinct in its genotypic composition is being grouped separately (Cl X)

Unique alleles for accessions

Unique alleles were detected that can be used to distinguish genotypes among themselves; several primer pairs produced specific (at least one unique allele per accession) to distinguish any two accessions at a time. Unique alleles for different genotypes, sizes of the fingerprint are provided in the Table 3. The markers *Xtxp*354 amplified eight different

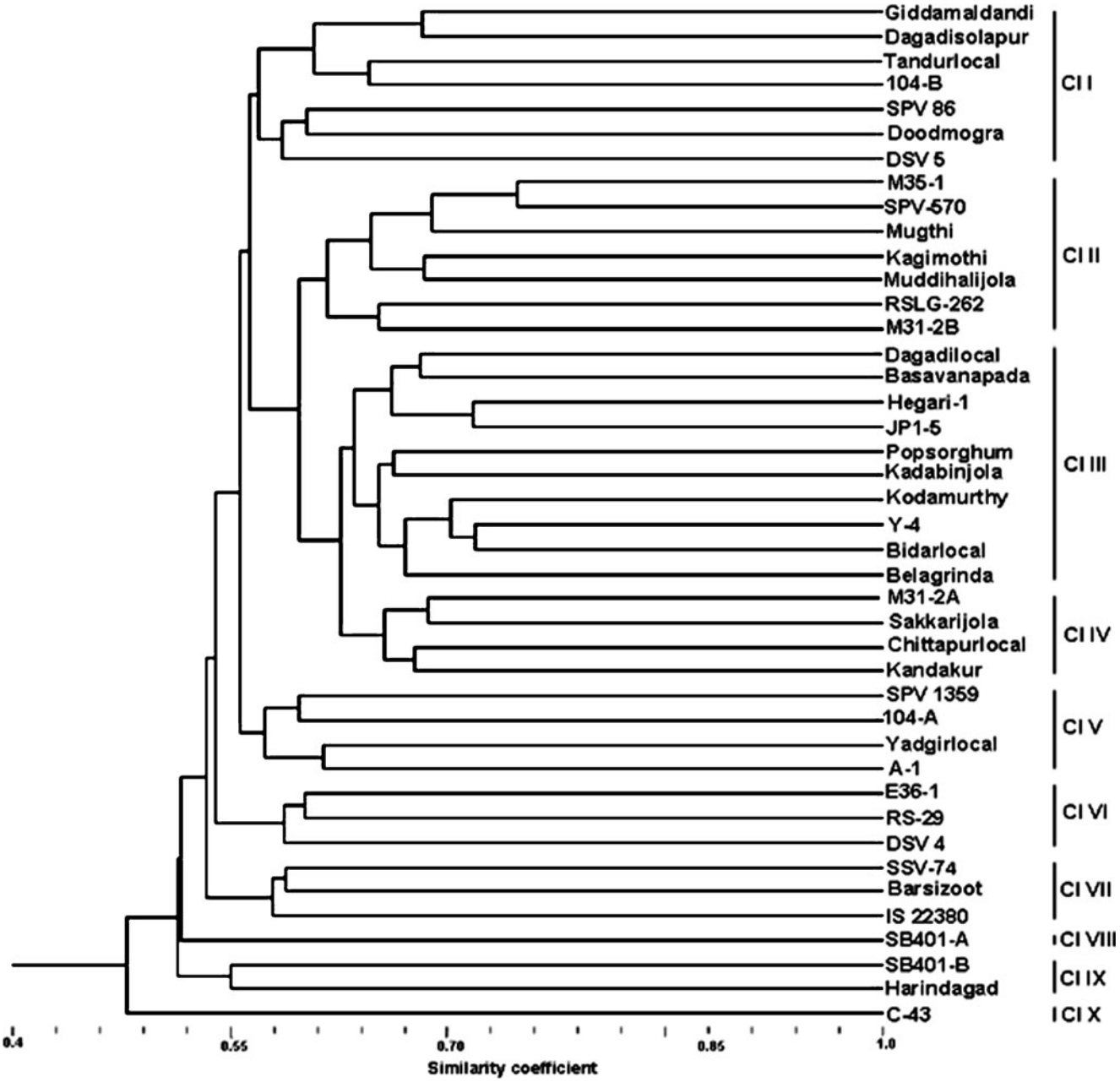

Fig. 1 UPGMA dendrogram constructed using allelic data obtained for 30 SSR markers depicts genetic relationships among 42 sorghum accessions

amplicons of varying size (290 to 380 bases) in IS 22380, E36-1, Yangir landrace, Kadabinjola, Hagari, Dagadi landrace, SPV-86 and SSV-74 genotypes. The markers *Xtxp*343 and *Xtxp276* amplified 4 specific amplicons each, in eight different genotypes.

Discussion

The characterization and quantification of genetic diversity within closely related crop germplasm has long been a major goal, as it is essential for a rational use of genetic resources. Furthermore, the analysis of genetic variation among breeding materials is of fundamental interest to plant breeders, as it contributes immensely to selection, monitoring of germplasm, and also to prediction of potential genetic gains (Chakravarthy and Rambabu 2006). Limitations on morphological characterization, including difficulties concerning the definition and validation of neutral traits, experimental costs, and evaluation time and genotype × environment interaction are widely discussed in germplasm characterization studies (Marita et al. 2000; Chandra et al. 2002). Considering these, DNA-based molecular markers have proven to be powerful tools in the assessment of

Table 3 Unique alleles for selected accessions of *rabi* sorghum

Sl. No.	Locus ID	Genotype	Fingerprint (bp)
1	*Xtxp*326	Kandakur	185
2	*Xtxp*343	E36-1	227
3	*Xtxp*343	C-43	250
4	*Xtxp*343	Doodmogra	245
5	*Xtxp*343	Barsizoot	240
6	*Xtxp*321	DSV-4	260
7	*Xtxp*321	RS-29	125
8	*Xtxp*21	A-1	120
9	*Xtxp*354	IS 22380	110
10	*Xtxp*354	SSV-74	290
11	*Xtxp*354	E36-1	300
12	*Xtxp*354	Dagadi landrace	320
13	*Xtxp*354	SPV-86	350
14	*Xtxp*354	Yangir landrace	355
15	*Xtxp*354	Kadabinjola	380
16	*Xtxp*354	Hagari	110
17	*Xtxp*63	SSV-74	125
18	*Xtxp*63	Dagadi landrace	135
19	*Xtxp*63	Harinigad	130
20	*Xtxp*276	Sakkarajola	135
21	*Xtxp*276	Dagadi sholapur	128
22	*Xtxp*276	104-A	100
23	*Xtxp*276	Yangir landrace	130
24	*Xtxp*276	SB401-B	132

genetic variation and in the elucidation of genetic relationships within and among species, characterized by abundance and untouched by environmental influence. Nevertheless, microsatellite or SSR markers have become the preferred molecular markers for studying genetic diversity in many crops owing to their co-dominance, multi-allelic nature, and ease of use and repeatability of assays.

In the present investigation, 30 SSR marker loci distributed across the ten linkage groups of sorghum genome were used to assess the genetic diversity among 42 sorghum genotypes. The advantage of using markers with known map positions instead of a random sample is that there is control over the coverage of the genome. Thus, it is possible to avoid over representation of certain regions of the genetic map that can produce inaccurate estimates of genetic similarities among individuals (Menz et al. 2002). In general, the level of allelic diversity among the germplasm included in this study was high. The 30 SSR markers evaluated in this study consisted of 14 di-nucleotide, seven tri-nucleotide, six-compound type repeat units. The allele sizes across all microsatellite loci were in the range of 100–750 bp. Maximum number of alleles were observed in case of SSRs with trinucleotide repeat motifs, similar results were report by Jayashree et al. (2006). Mutations in SSR markers resulting in allele size differences are often caused by deletions or insertions of single or multiple repeat units due to unequal crossing over followed by concerted evolution. Possibly, insertion of DNA segments or unusual high rate of multiplication of certain repeats be responsible for observed orthology of amplicons in some accessions. However, homology of primer pairs to orthologous regions within genome might also yield amplicons of too high or low in molecular weight (Li and Gill 2002; Kumar et al. 2009). The average number of alleles per locus was 7.6, indicating a greater magnitude of diversity among the plant materials included in this investigation. Though the marker, *Xtxp*318 revealed highest number of alleles, the linkage group of this marker is not known. Similarly, the marker *Xtxp*6 on chromosome 6 (LG F) generated a maximum of 15 alleles, while marker *Xtxp*21 in chromosome 4 (LG D) had a minimum of 2 alleles (Table 2). Many studies have also reported significant differences in allelic diversity among various microsatellite loci (Uptmoor et al. 2003). The alleles revealed by markers showed a high degree of polymorphism, with as many as 28 markers producing 100% of bands polymorphic. This amply suggested that the genotypes selected for this study harbor enough genetic divergence. Nevertheless, the average expected heterozygosity was very high (0.8), expected heterozygosity is a more accurate measure of polymorphism, as it further measures the distribution of those alleles across the germplasm being examined. Although the number of alleles and expected heterozygosity are dependent on the specific markers selected and the diversity of germplasm used, values were similar to those found in previous studies. The average number of SSR alleles reported has ranged from 4.4 to 6.8 (Uptmoor et al. 2003; Dillon et al. 2005; Menz et al. 2002; Agrama and Tunistra 2003) and expected heterozygosity values have ranged from 0.58 to 0.81 (Uptmoor et al. 2003; Dillon et al. 2005; Chandra et al. 2002). The higher heterozygosity observed in the present study might be due to the higher number of landraces being sampled for the experiment. Markers with PIC values of 0.5 or higher are highly informative for genetic studies and are extremely useful in distinguishing the polymorphism rate of a marker at a specific locus (DeWoody et al. 1995).The SSR markers used in this study were highly informative, because PIC values higher than 0.5 indicate high polymorphism. The observed number of alleles across the loci was more than the effective number of alleles (1.699 to 7.251) as expected. Similarity coefficients ranged from 0.41 to 0.75 for all accessions; the minimum genetic relatedness was 41 per cent between restorer line C-43, and with all the other accessions under study. C-43, a restorer line, included in the study, placed itself in a separate cluster with least indicative similarity with others, whereas highest similarity (0.75) was observed between SPV-570 and M35-1 and 0.72 was observed between M35-1 and Mugthi. SSR markers detect finer levels of

variation among closely related lines. Nevertheless, 30 microsatellite markers were able to discriminate between the landraces/cultivars/release varieties/advanced breeding lines and demonstrated a maximum genetic similarity value of 0.66 between the landraces Basavanapada, Belagrinda and Y-4. These landraces were originated from Karnataka and had similar morphological features, such as a taller and erect plant type, short bold and awnless grains.

The genotypes studied in this work, represent germplasm with important agronomic characteristics for potential commercial exploitation. The genus sorghum possesses enormous wealth of genetic diversity. Several factors could contribute to high level of genetic variation present in cultivated sorghum (DeWoody et al. 1995). Multiple origins of domesticated sorghums, cross-pollination between selected races and out crossing between domestic cultivars and highly variable wild species are considered to be major factors contributing to the extensive genetic diversity observed in sorghum (Doggett 1988). In addition, part of diversity is attributed to domestication of sorghum from several wild species; the high rate of naturally occurring hybridization between landraces and their wild relatives can lead to highly polymorphic genotypes. Overtime, several genetic changes must have happened in ecotypes for their adaptation to a specific situations and quality, which made them to prevail even today. A number of unwanted nucleotide sequences might have been eroded during this process of adaptation in the agriculture system. Sorghum SSR markers revealed higher levels of genetic polymorphism in present study. The high level of polymorphism associated with SSR markers may be a function of the unique replication slippage mechanism, loss or gain of specific nucleotide/s during evolution responsible for generating SSR allelic diversity (Morgante et al. 2002). Based on the simple matching coefficients a genetic similarity matrix was constructed using the SSR data to assess the genetic relatedness among the 42 selected accessions. All the selected accessions were grouped into ten clusters. Cluster III is the largest cluster with ten accessions; followed by followed by Cl I and Cl II with seven accessions; Cl VI and Cl VII with three accessions each. However, Cl VIII and Cl X comprised of only one accession SB401-A and C-43 respectively while Cl IX comprised of two accessions (SB401-B, Harinidagad). Similarity matrices constructed based on shared allele frequencies revealed that the average genetic similarity between genotypes was lowest (0.41). The polyallelic nature of SSR markers has the advantage of discriminating the individuals more precisely. Nine out of nineteen landraces were grouped in cluster III, depicting considerable similarity between them. Cl I comprise of four landraces, one maintainer line 104-B and two advanced breeding lines (SPV-86 and DSV-5). Male sterile line, SB401-A and male fertile line, SB401-B were grouped into separate clusters indicating the divergence between them. Similarly, male sterile line 104-A and corresponding male fertile lines, 104-B were grouped into separate clusters, indicating high degree of diversity among the male sterile and male fertile lines. This justifies, perhaps,

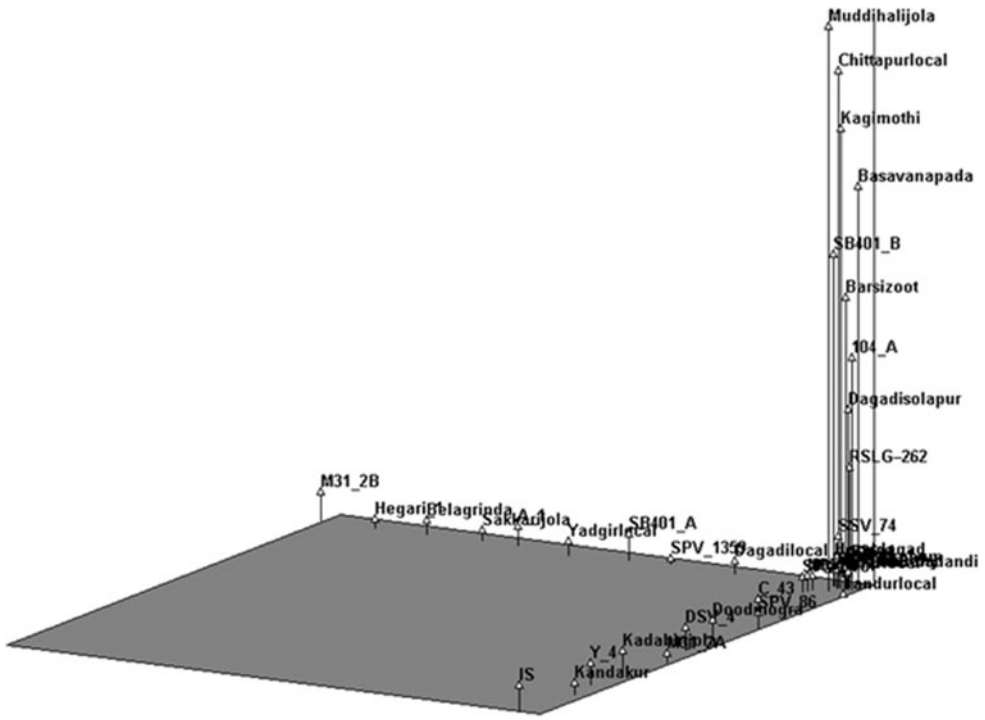

Fig. 2 Multidimensional scatter plot showing genetic relations among 42 sorghum accessions

the reason why they are included as parental lines in breeding programmes. DSV-5 and DSV-4, developed at Main Research Station, Dharwad got grouped into Cl I and Cl VI respectively, with a similarity coefficient of 0.52. The accessions M35-1, Mugthi, SPV-570 with Milo cytoplasm were grouped into cluster II as expected. However, SB401-B, a promising B line on Milo cytoplasm grouped along with a landrace Harinidagad (Cl IX). The very low bootstrap value (35%) observed in the clad that comprised cultivars and landraces indicate that the relative position of each accession may vary within the cluster if the dendrogram is rebuilt (Fig. 1). Although examined set of germplasm was grouped into four clusters through multi-dimensional scatter plot (MDS) the grouping pattern were in accordance with the dendrogram constructed (Fig. 2).

A total of 24 unique alleles were obtained on screening the 42 selected accessions with 30 SSR markers. Seven SSR markers amplified 24 unique alleles in 23 genotypes. The size of specific amplicons ranged from 80 to 355 bases. E36-1 recorded two specific amplicons of size 227 bases and 300 bases with *Xtxp*354 and *Xtxp*343 markers respectively. The *Xtxp*354 marker amplified eight unique alleles of varying size (290 to 380 bases) in IS 22380, E36-1, Yangir landrace, Kadabinjola, Hagari, Dagadi landrace, SPV-86 and SSV-74 genotypes, this indicates *Xtxp*354 marker is a highly informative SSR marker that differentiate these genotypes every effectively. The primers *Xtxp*343 and *Xtxp276* amplified 4 unique alleles each, in eight different genotypes. From this it can be said that the primer *Xtxp*354 has higher capacity to distinguish between genotypes. Such unique alleles are important in diagnostic for particular regions of the genome specific to a particular type of sorghum and their utility in registration and breeding.

Thus, the molecular characterization based on 30 microsatellite markers allowed for an in-depth look at the genetic information and organization of the germplasm collection evaluated. The results of this study indicated that there was substantial genetic variation and polymorphism across the loci studied. Furthermore, the study provided a first detailed analysis and quantification of genetic diversity present in the selected accessions of *rabi* core collection. The data also reaffirms the power of SSR markers to distinctly group closely related landraces. Several earlier studies have indicated that SSR technology is highly cost-effective (Smith et al. 2000) and could be easily employed in resource poor countries. Our data demonstrated accessions studied contain a great deal of genetic diversity as indicated by the observed number of alleles, far beyond that observed in any other sorghum germplasm source of comparable sample number. The utility of PCR-based markers such as SSRs for measuring diversity, for assigning genotypes to heterotic groups and for genetic fingerprinting should prove valuable for sorghum breeding programs.

Acknowledgements Financial assistance was from Ministry of Information Technology and Biotechnology, Government of Karnataka, and Department of Biotechnology, Government of India. We are thankful Dr M.Y. Kamatar, and Dr M. G. Palakshappa for providing sorghum accessions and Dr. N. Seetharama, Director, National Research Center for Sorghum, Hyderabad, India, for encouragement. We thank anonymous reviewers for their critical comments which helped improving the scientific merit of the manuscript.

References

Abu-Assar AH, Uptmoor R, Abdelmula AA, Salih M, Ordon F, Friedt W (2005) Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Sci 45:1636–1644

Agrama HA, Tunistra MR (2003) Phylogenetic diversity and relationships among sorghum accessions using SSR and RAPDs. Afr J Biotechnol 2:334–340

Aldrich PR, Doebley J (1992) Restriction fragment variation in nuclear and chloroplast genomes of cultivated and wild sortghums (*Sorghum boicolar* (L.) Moench). Theor Appl Genet 85:293–302

Aldrich PR, Doebley J, Schertz KF, Stec A (1992) Patterns of allozyme variation in cultivated sorghum and wild sorghum (*Sorghum bicolar* L Moench). Theor Appl Genet 85:451–460

Appa-Rao S, Prasada-Rao KE, Mendesha MH, Gopal-Reddy V (1996) Morphological diversity in sorghum germplasm from India. Genet Resour Crop Evol 43:559–567

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

Belaj A, Satovic Z, Cipriani G, Baldoni L, Testolin R, Rallo L, Trujillo I (2003) Comparative study of the discriminating capacity of RAPD, AFLP and SSR markers and of their effectiveness in establish genetic relationships in olive. Theor Appl Genet 107:736–744

Bhattramakki D, Dong J, Chhabra AK, Hart GE (2000) An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. Genome 43:988–1002

Botstein D, White RL, Skolnick M, Davis RV (1980) Construction of a genetic map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Chakravarthy KB, Rambabu N (2006) SSR marker based DNA fingerprinting and diversity study in rice (*Oryza sativa* L.). Afr J Biotechnol 5:684–688

Chandra S, Huaman Z, Hari Krishna S, Ortiz R (2002) Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data—a simulation study. Theor Appl Genet 104:1325–1334

DeWoody JA, Honeycutt RL, Skow LC (1995) Microsatellite markers in white-tailed deer. J Hered 86:317–319

Dillon SL, Lawrence PK, Henry RJ (2005) The new use of *Sorghum bicolor*-derived SSR markers to evaluate genetic diversity in 17 Australian Sorghum species. Plant Genet Resour 3:19–28

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, Ater M, Lefebvre C, Vekemans X (1998) Patterns of morphological and allozyme variation in sorghum landraces of Northwestern Morocco. Genet Resour Crop Evol 45:541–548

Doggett H (1988) Sorghum, 2nd edn. Wiley, New York

FAOSTAT: agricultural data. http://apps.fao.org, last accessed: October 2006

Harlan JR, Dewet JMJ (1972) A simplified classification of sorghum. Crop Sci 12:172–173

Jayashree B, Ramu P, Prasad P, Bantte K, Hash CT, Chandra S, Hoisington DA, Varshney RK (2006) A database of simple sequence repeats from cereal and legume expressed sequence tags mined *in silico*: survey and evaluation. In Silico Biol 6:0054

Jordan DR, Tao YZ, Godwin ID, Henzel RG, Cooper M, Mcintyre CL (1998) Loss of genetic diversity associated with selection for resistance to sorghum midge in Australia sorghum. Euphytica 102:1–7

Kruskal JB, Wish M (1978) Multidimensional scaling. Sage, Newbury Park

Kumar S, Mohan A, Balyan HS, Gupta PK (2009) Orthology between genomes of *Brachypodium*, wheat and rice. BMC Research Notes 2:93

Li W, Gill BS (2002) The colinearity of the *Sh2/A1* orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the triticeae. Genetics 160:1153–1162

Mace ES, Buhariwalla HK, Crouch JH (2003) A high throughput DNA extraction protocol for tropical molecular breeding programs. Plant Mol Biol Rep 21:459a–459h

Marita JM, Rodriguez JM, Nienhuis J (2000) Development of an algorithm identifying maximally diverse core collections. Genet Resour Crop Evol 47:515–526

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®, RFLP and SSR markers. Plant Mol Biol 48:483–499

Morgante M, Oliveri AM (1993) PCR amplified microsatillite as marker in plant genetics. Plant J 3:175–182

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194–200

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci (USA) 76:5369–5373

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet 97:1248–1255

Rohlf FJ (2002) NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 2.10. Exeter Software, New York

Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Den RE, Woodman WL, Lee M, Porter K (2000) Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci 40:226–232

Swofford DL (2003) PAUP* V4. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland

Tao Y, Manners JM, Ludlow M, Henzel RJ (1993) DNA polymorphisms in grain sorghum (*Sorghum bicolor* L.). Theor Appl Genet 86:679–688

Tegelstrom H (1992) Detection of mitochondrial DNA fragments. In: Hoelzel AR (ed) Molecular genetic analysis of populations: a practical approach. IRL, Oxford, pp 89–114

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

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotech 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