African Journal of Biotechnology Vol. 11(90), pp. 15626-15635, 8 November, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB12.321
ISSN 1684–5315 ©2012 Academic Journals

Full Length Research Paper

Molecular characterization and assessment of genetic diversity of sorghum inbred lines

R. Madhusudhana*, D. Balakrishna, P. Rajendrakumar, N. Seetharama and J. V. Patil

Directorate of Sorghum Research, Rajendranagar, Hyderabad 500 030, India.

Accepted 16 April, 2012

Selecting parents of diverse genetic base with contrasting phenotype is an important step in developing mapping populations for quantitative trait loci (QTL) detection and marker-assisted selection. We studied genetic diversity in 31 sorghum parents using 413 sorghum simple sequence repeats (SSR) markers. The polymorphism information content (PIC), a measure of gene diversity, varied from 0 to 0.92 with an average of 0.53 and was significantly correlated with number of alleles. The primers IS10215, IS10270 and IS10333 could differentiate all the 31 lines conclusively. Clustering analysis based on the genetic dissimilarity grouped the 31 parents into eight clusters and grouping was in good agreement with pedigree, race and geographic origin. Diverse pairs of sorghum parents were identified with contrast phenotype for various biotic and abiotic stresses with higher genetic diversity for developing recombinant inbred line (RIL) mapping populations to identify QTLs/genes for important traits in sorghum. One of the mapping populations resulted in the identification of QTLs for resistance to sorghum shoot fly and these QTL results were validated in a second mapping population.

Key words: Simple sequence repeats (SSR) markers, genetic diversity, sorghum, mapping parents.

INTRODUCTION

Sorghum (*Sorghum bicolor* L. Moench) is the fifth most important cereal crop in the world after wheat, rice, maize and barley. It is a C_4 crop adapted to semi-arid tropical environments that are too dry for cultivation of other cereals like maize (Dogget, 1988; Rooney, 2004). Sorghum is a major staple food and fodder crop and is considered as a pillar of food security in these regions around the world (Bhosale et al., 2011).

In recent past, interest in exploiting sorghum as a biofuel crop is also growing owing to its rich stalk sugar content (Wang et al., 2009). Sorghum breeding across the world is working on the development of high-yielding varieties and hybrids with better quality, disease resistance, drought tolerance and agronomic traits (Klein et al. 2008). Significant breakthrough has been made in developing and releasing sorghum hybrids and varieties for commercial cultivation both in India and elsewhere (Kumar et al., 2011).

However, several constraints of biotic and abiotic nature were limiting sorghum production at farmers' field. Major constraints to sorghum production have been identified and genetic enhancement programmes have been undertaken to breed for increased resistance (Kumar et al., 2011). Genetic diversity assessment is one key step in any plant breeding programme and knowledge of the genetic relationships among different accessions is essential for developing appropriate strategies for breeding and germplasm management. Effective genetic enhancement of a trait depends basically on the level of genetic diversity available in that crop species. Diverse landraces, breeding stocks and wild relatives are useful in broadening genetic base for greater genetic gains. Such studies are also helpful in choosing parents to create and maintain genetic diversity (Becelaere et al., 2005), develop mapping populations for detecting quantitative trait loci (QTL)s/genes (Varshney, 2011) and to categorize lines into heterotic groups for hybrid crop breeding (Menz et al.,2004). Earlier estimates of diversity

*Corresponding author. E-mail: madhu@sorghum.res.in. Tel: +91 40 24015349 226. Fax: +91 40 24016378.

Abbreviations: QTL, Quantitative trait loci; **SSR,** simple sequence repeats; **PIC,** polymorphism information content; **RIL,** recombinant inbred line.

using pedigree, agronomic and morphological traits were found to be generally inflated and found unrealistic (Fufa et al*.,* 2005; Almanza-Pinzon et al*.,* 2003; Van Beuningen and Busch, 1997; Cox et al*.,* 1986; Souza and Sorrells, 1989), and may not reflect the true genetic differences of the studied genotypes. On the other hand, deoxyribonucleic acid (DNA) based molecular markers, which are being increasingly used in genetic diversity studies represent more coverage, are not influenced by environment, reflect more accurate genetic similarity (and differences) and do not require previous pedigree information (Bohn et al*.,* 1999).

Many molecular marker technologies have been developed and applied for studying patterns of genetic diversity in sorghum (Ghebru et al., 2002; Folkertsma et al., 2005; Ali et al., 2011; Bhosale et al., 2011; Muraya et al., 2011; Ng' Uni et al., 2011; Sagnard et al., 2011; Thudi and Fakrudin, 2011). Simple sequence repeat (SSR) markers were markers of choice as they are polymerase chain reaction (PCR)-based, easy to use, co-dominant, locus-specific, highly reproducible, hypervariable and informative (Powell et al*.,* 1996). Molecular markers and QTL analysis offer a more efficient selection strategy for plant breeders to manipulate more complex quantitative traits. Identification of QTLs for a trait of interest mainly depends on the level of genetic contrast that exists for a target trait between parents, accurate phenotypic and genotypic data and the stringency of statistical methods employed. Mapping population is a key class of genetic resource for use in linkage mapping and marker-assisted breeding (Varshney, 2011). Identification of appropriate genotypes that are genetically diverse in addition to contrast phenotype is vital for using them as parents in the development of mapping populations. It is important to note that the genotypes with contrasting trait pheno-types may be good for trait mapping but may not be suitable for linkage map development owing to their genetic closeness. It is therefore essential to estimate the levels of genetic diversity available at molecular level among a set of phenotypically different genotypes to identify them as parents of mapping population. Earlier studies have employed relatively a limited number of SSR markers for diversity analysis in sorghum. The present study was therefore undertaken to assay genetic diversity among a set of sorghum genotypes using a large number of SSR markers to select phenotypically contrasting genotypes to be useful as parents for developing trait-based mapping populations and linkage maps for QTL detection for various biotic and abiotic stresses in sorghum.

MATERIALS AND METHODS

Thirty-one sorghum genotypes included in the study were selected on the basis of their contrast phenotype for important All India Coordinated Sorghum Improvement Project (AICSIP) abiotic and biotic stresses. Nineteen of them have been originated in India, nine in Africa, two in USA and one in Portugal. These lines have been evaluated in stress nurseries under the All India Coordinated Sorghum Improvement Project (AICSIP) over the past few years, and their reactions (susceptibility/resistance) to various biotic and abiotic stresses are known. Eighteen of these lines have been publicly released either as varieties or as a seed/pollinator parent of commercial hybrids. The details of the genotypes are given in Table 1.

DNA extraction

Fifty seeds of each genotype were grown in small plastic pots and were watered till the length of the seedlings was around 10 to 15 cm. Genomic DNA from each of the genotypes was extracted from a bulk of 15 plants using cetyltrimethylammonium bromide (CTAB) procedure (Saghai-Maroof et al*.,* 1984). The DNA quantity of each sample was estimated and DNA concentrations were normalized at 2.5 ng/µl. The DNA quality of each sample was evaluated by running 1 µl of DNA on a 1% agarose gel.

PCR and SSR assay

DNA prepared from bulk of 15 seedlings of each sample was used for PCR reactions. The parameters set for PCR amplification conditions were followed as described by Folkertsma et al*.* (2005) and the PCR and SSR assay was carried out at the Center of Excellence for Genomics, ICRISAT, and Hyderabad. Four hundred and thir-teen sorghum SSR primer pairs with single-copy SSR loci distributed across the 10 sorghum chromosomes were used for genotyping the 31 sorghum genotypes. PCR conditions were optimized for each of the 413 SSR markers and PCR reactions were set up in 5 µl volumes in 384-well PCR plates. Each PCR reaction contained 2 to 4 pmol of primer, 1 to 4 mM $MgCl_2$, 0.1 to 0.2 mM dNTP, 0.1 to 0.125 U Amplitaq Gold Polymerase (PE-Applied Biosystems) and 1X PCR buffer (PE-Applied Biosystems).

Temperature cycling was carried out using the Gene- Amp PCR System 9600 (PE-Applied Biosystems) and touch-down PCR amplification: one 15 min denaturation cycle, followed first by ten cycles of 94°C for 10 sec, 61°C for 20 sec (ramp of 1°C per cycle) and 72°C for 30 sec, then by 31 cycles of 94°C for 10 sec, 54°C for 20 sec and 72°C for 30 sec. After completion of the 31 cycles, a final extension of 20 min at 72°C was included to minimize the +A overhang PCR products were pooled post-PCR, where 1 µl of the FAM-labelled product, 1 µl of the HEX-labelled product and 1.5 µl of the NED-labelled product were mixed with 7 µl of formamide (PE-Applied Biosystems), 0.3 µl of the ROX-labelled 400 HD size standard (PE- Applied Biosystems) and 4.2 µl of distilled water. DNA fragments were denatured and size-fractioned using capillary electrophoresis on an ABI 3700 automatic DNA sequencer (PE-Applied Biosystems). The GENESCAN 3.1 software (PE-Applied Biosystems) was applied to size peak patterns, using the internal ROX 400 HD size standard and GENOTYPER 3.1 (PE-Applied Biosystems) for allele calling.

Cluster analyses

Dissimilarity indices were estimated using allelic data by simple allele matching and cluster analysis based on unweighted neighbor-joining (Gascuel, 1997) was carried using DARwin 5.0 dissimilarity analysis software (Perrier et al*.,* 2003). To ascertain the statistical strength of genetic relationships identified through this analysis, bootstrapping of the data (10,000 permutations) was performed. The total number of alleles detected, the number of common alleles with allelic frequencies of at least 5%, the observed size range (in base pairs; bp), the allele size differences (in bp) and the polymorphism information content (PIC) values (Smith et al., 2000), and frequencies of unique alleles were calculated for each SSR marker

Table 1. Details of sorghum genotypes used in the study.

Genotype	Status	Primary trait	Race	Origin	Pedigree
2219B	Seed parent	M, SF, SB, GM	*Caudatum*	DSR, India	Selection from Kharif Shallu
296B	Seed parent	SF, SB, GM	*Kafir- durra*	DSR, India	IS3922 x Karad local
27B	Seed parent	SF, SB, GM	*Durra-Caudatum*	DSR, India	(CS3687 x CS3922) x (2219B x CS3922)
CS3541	Pollinator parent	FD	*Kafir-durra-Zera-zera*	DSR, India	IS3675 x IS3541
RS29	Pollinator parent	GY	*Caudatum-Zera-zera*	DSR, India	SC108 x SPV126
SPV1616	Variety	GY	*Caudatum*	DSR, India	[(SPV475 x SPV462) x kh89-246]
R16	Pollinator parent	D	*Durra*	DSR, India	IS2950 x M35-1
CSV8R	Variety	D	*Durra*	Parbhani, India	R24 x R16
M35-1	Variety	D, PBM, GY	*Durra*	Rahuri, India	Selection from Maldandi Bulk
CSV216R	Variety	D, PBM, GY	*Durra*	Rahuri, India	Pure line selection from RSLG112-1-6
DJ6514	Genetic stock	SF, MD	*Caudatum*	Dharwad, India	
CSV17	Variety		*Caudatum*	Udaipur, India	[(SPV475 x SPV462)] x {[IC419-1 (2077A x CS3541) x Merta local] x [SC108-3 x CS3541)-3-1]}
IMS9B	Seed parent	SF, SB, GM	*Kafir*	Indore, India	[2077A x (MA9B x Vidhisha 60-1)-11-4-2-5-5A]
N13	Variety	D	*Durra*	Nandyal, India	Nandyal selection
HC308	Variety	FD	*Durra*	Hissar, India	SPV8 (=IS508 (*durra-caudatum*) x Karad local) x IS4776 (*durra*)
IS2205	Genetic stock	SF, SB	*Durra-bicolor*	India	Landrace
B58586	Genetic stock	GM	*Guinea*	India	Landrace
ICSB38	Seed parent	SS	*Durra-Caudatum*	ICRISAT, India	[(BTx623 x MR862) B line Bulk]-5-1-3-5
PB15881-1	Genetic stock	SB	*Kafir-Caudatum*	ICRISAT, India	(296B x IS18579)-2-3-4-1

Table 1 Contd.

IS18551	Genetic stock	SF, SB	*Durra*	Ethiopia	Landrace
IS2122	Genetic stock	SF, SB	*Durra*	Ethiopia	Landrace
IS8525	Genetic stock	E, GM	*Caudatum*	Ethiopia	Landrace
E36-1	Variety	D, M	*Guinea-Caudatum*	Ethiopia	Landrace
IS2312	Genetic stock	SB	*Durra*	Sudan	Landrace
IS22380	Genetic stock	D	*Caudatum*	Sudan	Landrace
IS9830	Genetic stock	D	*Caudatum*	Sudan	Landrace
IS10284	Genetic stock	FD	*Guinea-Caudatum*	Sudan	Landrace
IS26866	Genetic stock	FD	*Guinea-Caudatum*	Nigeria	Landrace
IS14131	Genetic stock		*Bicolor*	Portugal	Landrace
BTx623	Seed parent	SF, SB, GM	*Kafir x Zera-zera*	USA	BTx3197 x SC170-6-4
B35	Seed parent	D	*Durra*	USA	Derivative of IS12555, a landrace from Ethiopia

Maturity (M), shoot fly (SF), grain mold (GM), drought (D), ergot (E), charcoal rot (C), foliar diseases (FD), plant height and biomass (PBM), grain yield (GY), Midge (MD), stalk sugar (SS).

using PowerMarker V3.25 (Liu and Muse, 2005).

RESULTS

Polymorphism of SSR markers

For the 413 SSRs genotyped over 31 genotypes, 2533 alleles were observed of which 995 (40%) were rare. The number of alleles per locus ranged from 1 to 19 with an average of 6.13. The chromosomes, LG B (7.3) showed highest number of alleles followed by LG H (6.55) and LG A (6.46). Twenty-six (6%) SSRs were mono-morphic (one allele per marker) and the rest 387 (94%) SSRs were polymorphic. The mean polymorphism information content (PIC) was 0.53, and it ranged from 0 (monomorphic) to 0.92. Among the SSR series, Xtxp, XIS, Xgap, msbCIR and gpsb had more average allele number compared to Xisep, Xcup and S series of SSRs. The frequency of the major alleles at different SSR loci ranged from 0.13 (IS10264, IS10270, Xtxp329) to 0.98 (ISEP0612, Xcup41, Xcup09 and Xtxp231). The SSR marker, IS10215 showed highest polymor-phism with 19 alleles.

Number of repeats in the SSR motif had strong correlation with allele number ($r = 0.77$. $P = 0.000$) and PIC ($r = 0.64$. $P = 0.000$). However, repeat motif type had significant negative correlation with repeat number ($r = -0.44$, $P = 0.000$), allele number ($r = -0.39$, $P = 0.000$) and PIC values ($r = -0.38$, $P = 0.000$). Of the types of repeat motifs, di-nucleotide repeat motifs showed higher number of alleles per SSR locus (7.05) and PIC (0.6) compared to tri- (4.87, 0.46), tetra- (3.22, 0.34) and penta-(3.06, 0.32) nucleotide repeat motifs. A total of 59 compound SSRs were available and exhibited an average of 8.10 alleles per locus and PIC value of 0.63. Both di- and tri-nucleotide motif SSRs showed repeat units from 4 to 42 while tetra- and penta-motifs had repeats up to 6 numbers. Both di-and tri- nucleotide with $n>30$ repeat units had significantly higher number of alleles and PIC values as compared to classes with $n \leq 30$.

Cluster analysis

Cluster analysis based on the simple allele matching and unweighted neighbor-joining was performed on the 413 SSR markers for 31

genotypes. The cluster analysis discriminated the genotypes and has a cophenetic value 0.97. Eight main clusters were recognized which corresponded well with the pedigree information, race and geographic origin of the genotypes. The cluster I had 11 genotypes which were mostly of *durra* race. The cluster I has five sub-clusters (Ia-Ie). The sub-cluster Ia contained four lines all of which were of African origin. Sub-cluster Ib consists of a land race from Portugal. Sub-cluster Ic contained four genotypes of Indian origin belonging to *durra* race, while subclusters Id consisted of single fodder genotype. Subcluster Ie consists of single genotype grown in Andhra Pradesh, India. In cluster II, seed parent 296B and its cross derivatives of Indian origin have been grouped, while in cluster III, the staygreen genotype B35 of American origin is included. Two landraces of Sudanese origin belonging to *caudatum* and *guinea-caudatum* have been grouped in cluster IV. Cluster V consisted of eight genotypes, both cultivated and land races from African and Indian origin belonging mostly to *caudatum* and its intermediate race with *guinea*. *BT* × *623* and its derivative, ICSB38 grouped separately in cluster VI. The cluster VII contained two seed parents of Indian origin belonging to *caudatum* and *durra-caudatum* races. Finally, the cluster VIII had two genotypes of *caudatum* race originated from Ethiopia and India.

DISCUSSION

Molecular markers such as restriction fragement length polymorphism, random amplified polymorphic DNA, simple sequence repeats, diversity array technology (RFLP, RAPD, SSRs, DArT etc) have been used in many crops including sorghum for characterizing and quantifying genetic diversity for use in genetic enhancement programmes. The analysis of genetic variation among genetic materials is of fundamental importance to plant breeders, as it contributes immensely to selection, monitoring of germplasm and prediction of potential genetic gains (Chakravarthy and Rambabu, 2006). Limitations on use of morphological traits, their experimental costs, and evaluation time and genotype × environment interactions are widely discussed (Fufa et al., 2005; Marita et al., 2000; Chandra et al., 2002). In this context, DNA-based markers have become powerful application tools for characterizing and quantifying genetic differences within and among species. Microsatellite or SSRs have become markers of choice for studying genetic diversity in many crops owing to their co-dominance, multi-allelic nature, ease of use and repeatability. With the availability of complete sorghum genomic sequence (Paterson et al., 2009), thousands of SSRs have been published in sorghum and have been employed in genetic diversity and other studies.

In the present study, 413 SSR marker loci comprising both genomic and genic SSRs were used to assess the genetic diversity among 31 sorghum genotypes. Compared to earlier studies, use of large number of markers in this study provided reasonably good coverage of the sorghum genome. On an average, each chromosome was represented by 40 SSR markers. Chromosome- wise LG A was represented by highest SSRs (63) followed by B and C (56), D (40), E and H (29), F (25), G (32), I (24) and J (36). The map positions of the rest of the markers (21) are not known. The level of allelic diversity observed in the present study was high with SSRs generating 2533 alleles. The average number of alleles revealed per SSR locus detected in the present study was 6.13 and was higher to that detected by Schloss et al. (2002), Agrama and Tuinstra (2003), Smith et al. (2000) and Ali et al. (2008) and lower from other studies (Wang et al., 2006; Thudi and Fakrudin, 2011). This could be due to levels of polymorphism of SSR markers, the diversity of genotypes and the sensitivity of DNA fragment separation systems. The differences between allele numbers among the SSRs could be due to the mechanisms that generate variability for example, punctual mutation, insertion / deletion, polymerase slippage and/or unequal crossing-over for microsatellite markers (Grenier et al., 2000). Mutations in SSR locus resulting in allele size differences are often caused by deletions or insertions of single or multiple repeat units due to unequal crossing-over (Thudi and Fakrudin, 2011).

Though the average allele number per SSR locus differed numerically between the 10 chromosomes, this was not statistically significant. The allele number ranged from 5.38 (LG G) to 7.3 (LG B). LG B (7.3) showed highest number of alleles followed by LG H (6.55) and LG A (6.46). The correlation between the SSR number per chromosome and average allele number per chromosome was non-significant ($r = 0.61$, $P = 0.06$). Most (94%) markers were polymorphic and gave multiple alleles in the range from 2 to 19. Marker series Xtxp, XIS, Xgap, msbCIR and gpsb, which were of genomic origin were highly polymorphic compared to gene-based marker series Xisep, Xcup and S. Genic SSRs have been reported to be less polymorphic compared with genomic SSRs in crop plants because of greater DNA sequence conservation in transcribed regions (Schloss et al., 2002; Varshney et al., 2005). The SSRs used in the study represented di-, tri-, tetra- and penta- nucleotide repeat units. The allele sizes among the genotypes for all the 413 microsatellites varied from 87 to 462 bp. Number of repeats in the SSR motif had strong correlation with allele number and their polymorphism information content. SSRs with di-nucleotide repeats are the most polymorphic marker class followed by tri-, tetra- and penta-repeat units. A direct relationship exists between marker information content and the number of repeat units (Weber 1990; Innan et al., 1997; Schloss et al., 2002). However, repeat motif type had significant negative correlation with repeat number, allele number and PIC values.

The PIC varied from 0.0 (monomorphic) to 0.92 with an

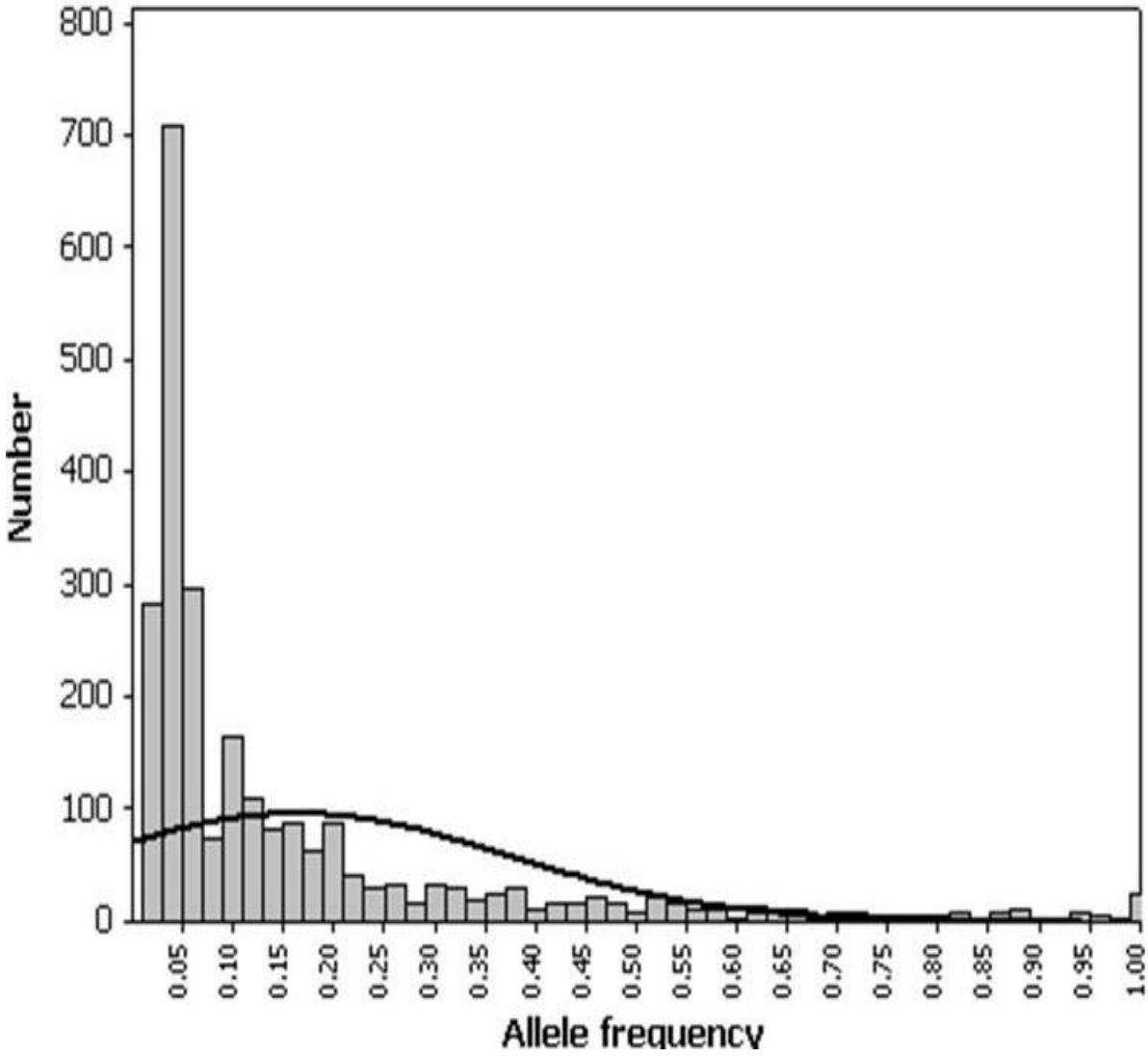

Figure 1. SSR allele frequency observed among the 31 sorghum parents.

average of 0.53 and were significantly correlated with allele number (r = 0.85, $p < 0.001$). PIC provides an estimate of the discriminatory power of a locus or loci by the number of alleles expressed and the relative frequencies of those alleles. According to PIC values, 68 markers were slightly informative (PIC<0.25 with mean = 0.09), 107 were reasonably informative (0.25<PIC<0.5, mean PIC = 0.39), 129 were highly informative (0.5<PIC>0.75, mean PIC = 0.63) and 104 were extremely informative (PIC>0.75, mean PIC = 0.84). The PIC values for the 10 sorghum chromosomes did not differ statistically and varied from 0.48 (LG 7) to 0.60 (LG 2). As reported in earlier study (Casa et al., 2005), around 40% alleles were found to be rare with an allelic frequency <0.05 (Figure 1). These rare alleles could be of particular interest as they are uniquely linked to some particular genotypes and may be diagnostic for particular genotypes or for particular regions of the genome specific to a particular type of sorghum (Agrama and Tuinstra, 2003; Thudi and Fakrudin, 2011). The SSR Xcup63 has been identified as rare in an earlier study (Ali et al., 2008).

The gene diversity observed in this study (Mean PIC = 0.53) is similar to the diversity value (0.40, 0.46, 0.62, 0.58) reported by Ali et al. (2008), Schloss et al. (2002), Agrama and Tuinstra (2003) and Smith et al. (2000), respectively. The SSR loci Xtxp32, Xtxp88, ISEP1012, Xtxp316, IS10333, IS10362, IS10215, Xtxp8, IS10270, IS10270 and Xtxp329 were rich in allelic diversity exhibiting (17-19) alleles with highest PIC of 0.92. These primers IS10215 and Xtxp329 could discriminate 29 parents decisively except for the two genotypes, IS2205 and IS18551 (Figure 2) which could be differentiated with a third primer IS10333 decisively. These 3 primers are extremely informative and discriminative with unique PCR products that could distinguish all the 31 genotypes successfully. Thus, these primers could be of great use in DNA fingerprinting to characterize sorghum genetic stocks in view of the emerging needs for DUS charac-

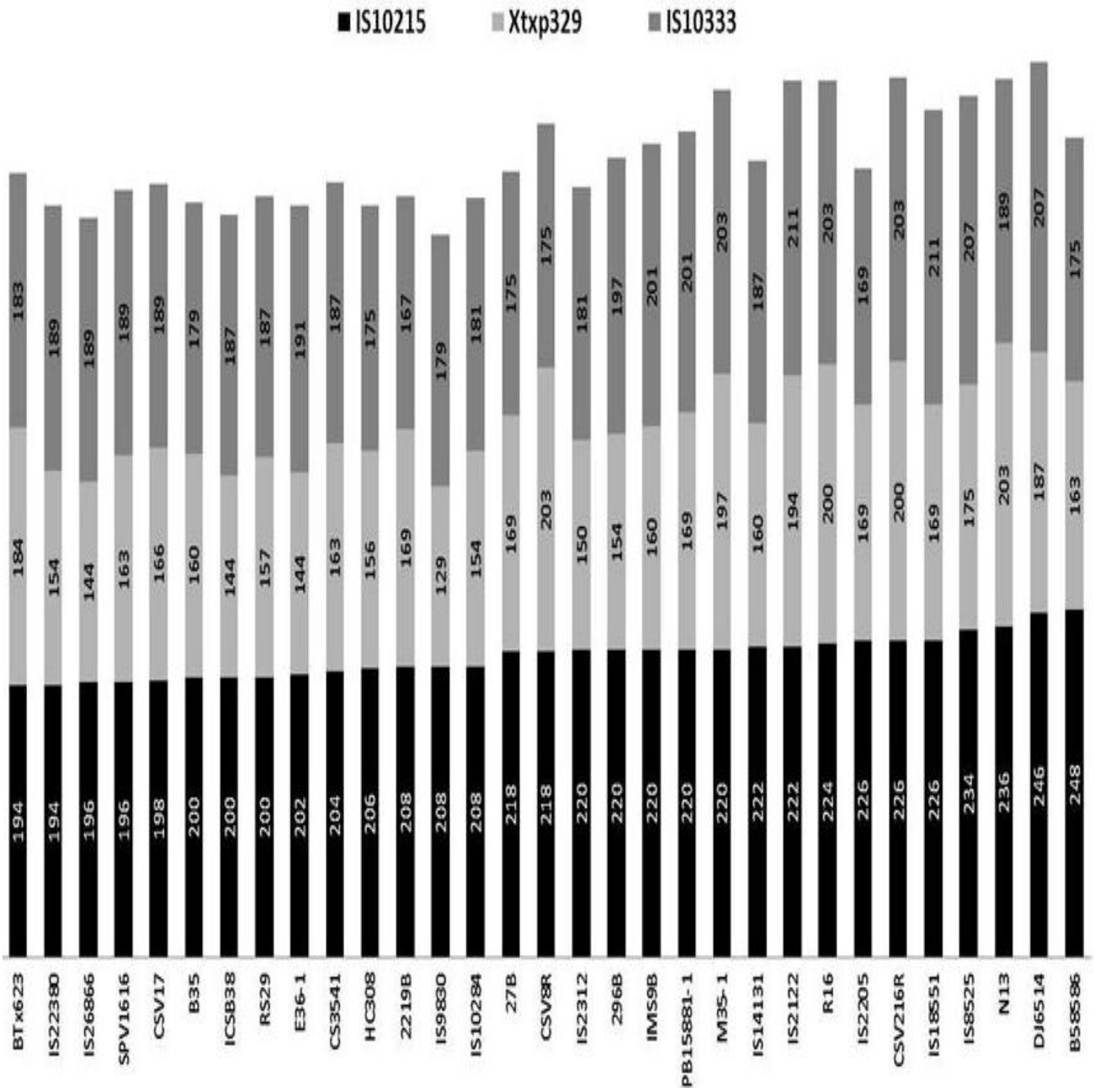

Figure 2. Amplicon sizes (bp) of three SSRs differentiate the 31 sorghum parents.

terization and plant varietal registration.

Cluster analysis

The pair-wise dissimilarity indices among the sorghum inbred lines were estimated using allelic data by simple allele matching followed by cluster analysis using unweighted neighbor-joining algorithm. All the 31 genotypes fell into eight (I to VIII) clusters which broadly correspond with their geographic origin, race and pedigree. Harlan and de Wet (1972) recognized five basic races (bicolor, guinea, caudatum, kafir and durra) and ten intermediate races (*guinea- bicolor, guinea-caudatum, guinea-kafir, guinea-durra, caudatum- bicolor, kafir-bicolor, durra-bicolor, kafir-caudatum, kafir-durra and durra-caudatum*) that originated as a result of natural intercrossing among basic races. All are recognizable on the basis of spikelet / panicle morphology alone (Smith and Frederiksen, 2000). Race *bicolor* is widely distributed in Africa and Asia, *guinea* predominant in West Africa, *caudatum*, throughout Central Africa, *kafir* in south of the equator in Africa and the *durras* in Ethiopia and India (Harlan and de Wet, 1972; Upadhyaya et al., 2009). Cluster I is the largest with 11 genotypes followed by cluster V with 8 acces-

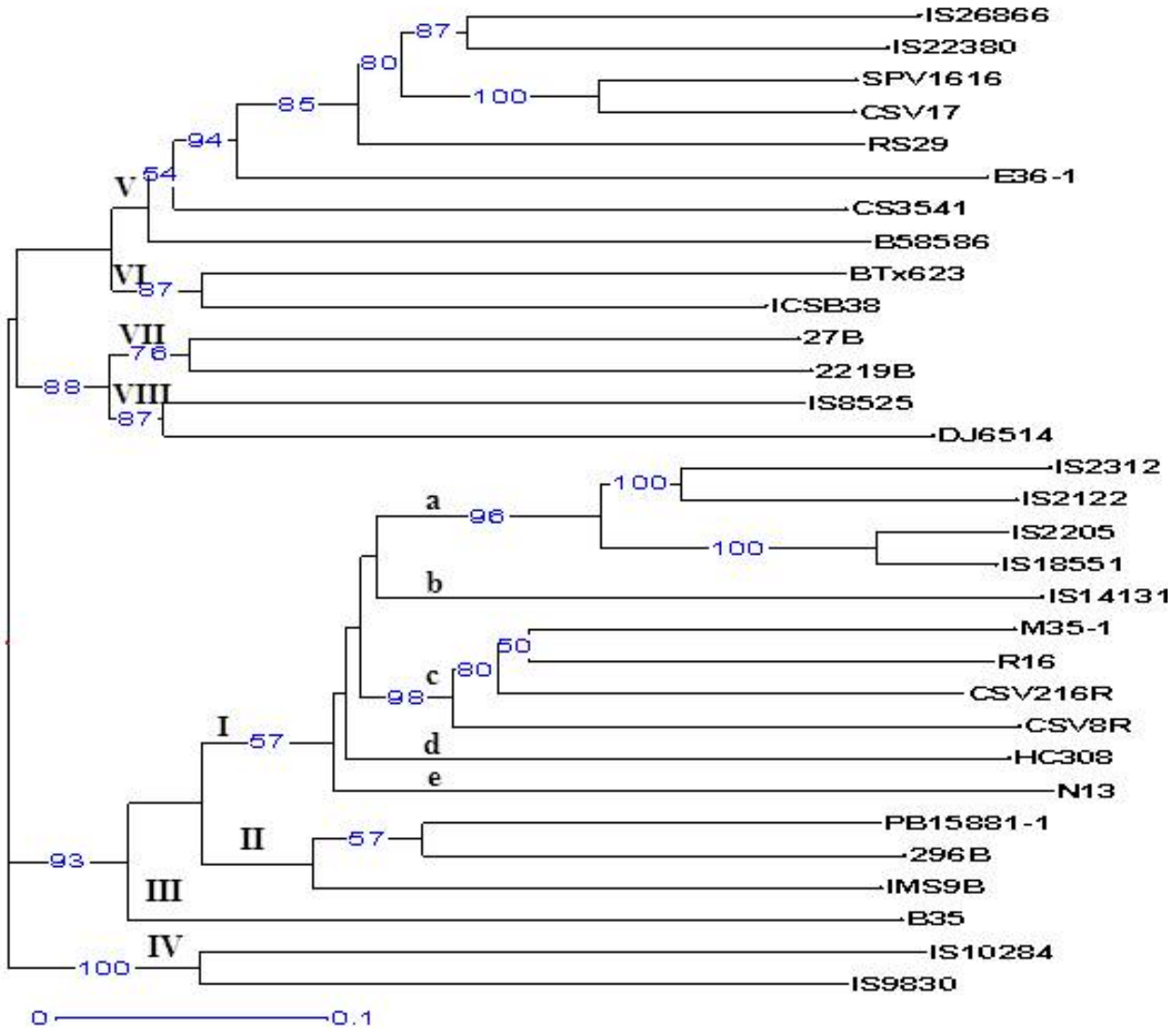

Figure 3. Neighbour joining phenogram depicting genetic relationships among 31 sorghum genotypes. Number along branches denotes bootstrap support (shown only for values greater than 50).

sions. Cluster II had 3 while clusters, IV, VI, VII and VIII had 2 genotypes each. Cluster III had only one accession.

The phenogram (Figure 3) discriminated all the sorghum genotypes in this study and has a cophenetic correlation value of 0.97, indicating a good fit of the dissimilarity matrix data to the phenogram topology. The cluster I, which comprises most of *durra* race genotypes, has five sub-clusters (Ia to Ie). In sub-cluster Ia, all the African *durra* landraces that are known for their shoot fly and stem borer resistance are clustered together. Similarly, group Ic contain Indian *durra* rabi cultivars, M35-1, its cross derivatives R16, CSV8R and a popular rabi cultivar CSV216R. Besides these, the genotypes HC308, N13 (both *durra* race) and IS14131 (a Portuguese *bicolor* landrace) clustered individually. This analysis indicates genotypic relatedness at DNA level and geographic origin. In cluster II, the seed parent 296B (*kafir-durra*), its cross derivative PB15881-1 (*kafir-caudatum*) and another seed parent IMS9B (*kafir* race) are clustered indicating their genetic closeness. It is important to note that the staygreen genotype B35 (*durra*), a BC_1F_2 derivative of Ethiopian *durra* landrace IS12555 separated from the main *durra* sub-groups of group I indicating its high genetic unrelatedness. This line has been predominantly used as a source of staygreen globally to enhance terminal drought tolerance in sorghum (Crasta et al., 1999). Group IV contains two Sudanese landraces of *guinea-caudatum* and *caudatum* race. Group V consists of genotypes belonging mainly to *caudatum* and its cross derivatives with *zera-zera*, *kafir* and *guinea* races. BTx623, a *kafir-zera-zera* line of the USA and its cross

derivative, ICSB38 (*durra-caudatum*) grouped separately in group VI. The group VII and VIII consists of genotypes of *caudatum* race, which are diverse from the Group V genotypes.

Utility of diversity analysis

Development of mapping population for trait mapping is an important requirement in marker-assisted breeding. Identification of genotypes which are genetically diverse and phenotypically contrasting is a challenging step. In the present study, genotypes included are valuable in mapping QTL and gene tagging. The study was useful in identifying genotypes which are genetically diverse with contrasting phenotypes for various biotic and abiotic stresses. For shoot fly, an important seedling insect pest of sorghum, the resistant donors (IS18551, IS2122) and susceptible genotypes (296B, 27B, IMS9B) are genetically diverse and are suitable as parents in development of mapping populations to identify shoot fly QTLs. Similarly, for grain mold disease, the resistant parent, B58586 is diverse from the susceptible parents 27B, 2219B, and 296B; for ergot, IS8525 (resistant) and IMS9B, 296B (susceptible). For terminal drought, R16, M35-1, CSV216R and B35, E36-1 are useful in developing mapping populations and detection of QTLs.

Mapping populations for shoot fly (296B × IS18551, 27B × IS2122), grain mold (296B ×B58586), drought and charcoal rot (M35-1×B35; CSV216R × B35), ergot (2219B × IS8525) and foliar diseases (IS10284 × IS26866) are being developed at DSR, Hyderabad, India.

The parents in these mapping populations were genetically diverse with a minimum genetic distance of 50. We were successful in identifying important QTLs for shoot fly resistance (Satish et al*.,* 2009) and their validation (Aruna et al*.,* 2011) employing two RIL populations (296B × IS18551 and 27B × IS2122). We have initiated marker-assisted introgression of shoot fly resistance QTLs for trichome density, non-preference and deadhearts using SSR markers Xnhsbm1044 and Xnhsbm1013 on SBI-10 and for leaf glossiness on SBI-05 using Xtxp65 and Xtxp30 SSRs into commercially important but shoot fly susceptible seed parents, 296B and 27B from the resistance donor parent IS18551.

Besides their applications in genetic diversity studies, microsatellites can be effectively utilized in DNA fingerprinting, purification of impure genetic stocks and in identification and authentication of genotypes of disputed pedigree. DNA fingerprinting, being inexpensive and decisive, has important role in characterization of genetic stocks and cultivars in view of plant varietal registration and breeders rights. A minimum of three SSR markers were able to successfully differentiate all the 31 genotypes very categorically in the present study indicating the use of marker technology to identify the genotypes. The markers will be of great value in identifying duplicates and protecting plant breeder's rights.

Conclusion

The present study demonstrated the utility of sorghum SSR markers in understanding the genetic similarities and affiliations among a set of sorghum genotypes. The

SSRs were useful in discriminating all the genotypes without ambiguity. Ample genetic diversity exists in the material at genetic level. The study has identified pairs of genotypes, which are potential parents to create mapping populations to map QTLs controlling economically important traits. Utility of diversity assay among a set of genotypes for selecting an appropriate parental pairs for developing trait based mapping populations was proved with the results of two shoot fly mapping populations.

REFERENCES

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

Alamnza-Pinzon MI, Khairullah M, Fox PN, Warburton ML (2003). Comparison of molecular markers and co-efficients of parentage for the analysis of genetic diversity among spring bread wheat accessions. Euphytica, 130: 77-86.

Ali ML, Rajewski JF, Baenziger PS, Gill KS, Eskridge KM, Dweikat I (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Mol. Breed. 2: 497-509.

Aruna C, Bhagwat VR, Madhusudhana R, Sharma V, Hussain T, Ghorade RB, Khandalkar HG, Audilakshmi S, Seetharama N (2011). Identification and validation of genomic regions that affect shoot fly resistance in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Gen. 122: 1617-1630.

Becelaere GV, Edward LL, Paterson AH, Chee PW (2005). Pedigree-vs. DNA marker-based genetic similarity estimates in Cotton. Crop Sci. 45: 2281-2287.

Bhosale SU, Stich B, Rattunde HFW, Weltzien E, Haussmann, BIG, Hash CT, Melchinger AE, Parzies HK (2011). Population structure in sorghum accessions from West Africa differing in race and maturity class. Genetica, 139: 453-463.

Bohn M, Utz HF, Melchinger AE (1999) Genetic similarities among winter wheat cultivars determined on the basis of RFLPs and SSRs and their use for predicting progeny variance. Crop Sci. 39: 228-237.

Casa AM, Mitchell SE, Hamblin MT, Sun H, Bowers JE, Paterson AH, Aquadro CF, Kresovich S (2005). Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. Theor. Appl. Gen. 111: 23-30.

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.

Cox TS, Murphy JP, Rodgers DM (1986). Changes in genetic diversity in the red winter wheat regions of the United States. PNAS, USA. 83: 5583-5586.

Crasta OR, Xu W, Rosenow DT, Mullet JE, Nguyen HT (1999). Mapping of post flowering drought resistance traits in grain sorghum: Association of QTLs influencing premature senescence and maturity. MGG. 262: 579-588.

Dogget H (1988). Sorghum. Longman Scientific and Technical, Essex

Folkertsma RT, Rattunde FH, Chandra S, Soma Raju W, Hash CT (2005). The pattern of genetic diversity of Guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. Theor. Appl. Genet. 111: 399-409.

Fufa H, Baenziger PS, Beecher BS, Dweikat I, Graybosch, Eskridge KM (2005). Comparison of phenotypic and molecular marker-based

classifications of hard red winter wheat cultivars. Euphytica, 145: 133-146.
Gascuel O (1997). Concerning the NJ algorithm and its unweighted version, UNJ. In: Mathematical Hierarchies and Biology. DIMACS workshop, Series in Discrete Mathematics and Theoretical Computer Science. Am. Mathematical Soc. 37: 149-170.
Ghebru B, Schmidt RJ, Bennetzen JL (2002). Genetic diversity of Eritrean sorghum landraces assessed with simple sequence repeat (SSR) markers. Theor. Appl. Genet. 105: 229-236.
Grenier C, Deu M, Kresovich S, Bramel-Cox PJ, Hamon P (2000). Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs non-random sampling procedures B. Using molecular markers. Theor. Appl. Genet. 101: 197-202.
Harlan JR, De Wet JMJ (1972). A simple classification of cultivated sorghum. Crop Sci. 12: 172-176.
Innan H, Terauchi R, Miyashita NT (1997). Microsatellite polymorphism in natural populations of wild plant *Arabidopsis thaliana*. Genetics, 146: 1441-1452.
Klein RR, Mullet JE, Jordan, DR, Miller FR, Rooney WL, Menz MA, Franks CD, Klein PE (2008). The effect of tropical sorghum conversion and Inbred development on genome diversity as revealed by high-resolution genotyping. Plant Genome, 1: 12-26.
Kumar AA, Reddy BVS, Sharma HC, Hash CT, Srinivasa Rao P, Ramaiah B, Sanjana Reddy P (2011). Recent advances in sorghum genetic enhancement research at ICRISAT. Am. J. Plant Sci. 2: 589-600.
Kumara CD, Bantilan MCS, Nedumaran S, Deb UK (2011).Tracking of Improved Cultivars adoption in India, paper presented in the 7th ASAE Conference held at Vietnam.
Liu K, Muse SV (2005). PowerMarker: An integrated analysis environment for genetic marker analysis. Bioinformatics, 21: 2128-2129.
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, Unruh NC, Rooney WL, Klein PE, Mullet JE (2004). Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci. 44: 1236-1244.
Muraya MM, De VS, Parzies HK, Mutegi E, Sagnard F, Kanyenji, BM, Kiambi, D, Geiger HH (2011). Genetic structure and diversity of wild sorghum populations (*Sorghum* spp.) from different eco-geographical regions of Kenya. Theor. Appl. Genet. 123: 571-583.
Ng' Uni D, Geleta M, Johansson E, Fatih, M, Bryngelsson T (2011). Characterization of the Southern African sorghum varieties for mineral contents : Prospects for breeding for grain mineral dense lines. J. Food Sci. 5: 436-445.
Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J et al., (Provide Complete Name) (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature, 457: 551-556.
Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In: Hamon P (ed.). Genetic diversity of cultivated tropical *plants*. Enfield, Science Publishers. Montpellier, pp. 43-76.
Powell WW, Machery GC, Provan J (1996). Polymorphism revealed by simple sequence repeats. Trends Gen. 1: 76-83.
Rooney WL (2004). Sorghum improvement-integrating traditional and new technology to produce improved genotypes. Adv. Agron. 83: 37-109.
Saghai-Maroof, MA, Soliman KM, Jorgensen RA, Allard RW (1984). Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci. USA, 81: 8014-8018.
Sagnard F, Deu M, Dembélé D, Leblois R, Touré L, Diakité M, Calatayud C, Vaksmann M, Bouchet S, Mallé Y, Togola S, Traoré PC (2011). Genetic diversity, structure, gene flow and evolutionary relationships within the *Sorghum bicolor* wild-weedy-crop complex in a western African region. Theor. Appl. Genet. 123(7): 1231-1246.
Satish K, Srinivas G, Madhusudhana R, Padmaja PG, Nagaraja Reddy R, Murali Mohan S, Seetharama N (2009). Identification of quantitative trait loci for resistance to shoot fly in sorghum [*Sorghum bicolor* (L.) Moench]. Theor. Appl. Gen. 119: 1425-1439.
Schloss SJ, Mitchell SE, White GM, Kukatla R, Bowers JE, Paterson AH, Kresovich S (2002). Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. Theor. Appl. Gen. 105: 912-920.
Smith CW, Frederiksen RA (2000). Sorghum: Origin, history, technology and production. New York: John Wiley Sons.
Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean 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.
Thudi M, Fakrudin B (2011). Identification of unique alleles and assessment of genetic diversity of rabi sorghum accessions using simple sequence repeat markers. J. Plant Biochem. Biotech. 20: 74-83. Souza E, Sorrells ME (1989). Pedigree analysis of North American Oat cultivars released from 1951 to 1985. Crop Sci. 29: 595-601.
Upadhyaya HD, Pundir RPS, Dwivedi SL, Gowda CLL, Reddy VG, Singh S (2009). Developing a mini core collection of sorghum for diversified utilization of germplasm. Crop Sci. 49: 1769-1780.
Van Beuningen LT, Busch RH (1997). Genetic diversity among North American spring wheat cultivars: III. Cluster analysis based on quantitative morphological traits. Crop Sci. 37: 981-988.
Varshney RK (2011). Application of next generation sequencing and genotyping technologies to develop large-scale genomic resources in SAT legume crops. In : Muralidharan K and Siddiq EA (eds.). Genomics and crop improvement: Revelance and Reservations. Institue of Biotechonology, Acharya NG Ranga Agricultural University, Rajendranagar, Hyderabad 500030, India.
Varshney RK, Graner A, Sorrells ME (2005). Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23: 48-55.
Wang M L, Zhu C, Barkley NA, Chen Z, Erpelding JE, Murray SC, Tuinstra MR, Tesso TR Pederson GA, Yu J (2009). Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. Theor. Appl. Gen. 120: 13-23.
Wang ML, Dean R, Erpelding J, Pederson G (2006a) Molecular genetic evaluation of sorghum germplasm differing in response to fungal diseases: Rust (*Puccinia purpurea*) and anthracnose (*Colletotricum graminicola*). Euphytica, 148: 319-330.
Weber JL (1990). Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics, 7: 524-530.