Genetic diversity among *Sorghum bicolor* L. Moench genotypes as revealed by prolamines and SSR markers

Ramya Vittal[1], Nabarun Ghosh[2*], Yiqun Weng[3, †] B. A. Stewart[1]

[1]Department of Agriculture and [2]Department of Life, Earth and Environmental Sciences, West Texas A&M University, Canyon, Texas 79016, USA. [3]Texas AgriLife Research, Amarillo, Texas 79106, USA.

Sorghum (*Sorghum bicolor* L. Moench) is a leading cereal in the arid and semi-arid regions and ranks fifth in importance among the world's grain crops. Given its importance as a staple food crop, a livestock feed crop and potentially a bioenergy crop, there is a constant need for its genetic improvement and an important step in this regard is the evaluation of genetic diversity in sorghum. In this study, 23 genotypes, including both local and exotic cultivars, were analyzed based on their water-soluble protein and prolamine-protein profiles. DNA-variation was investigated using twenty microsatellite, or simple sequence repeat (SSR) markers. There were no differences in the banding patterns of water-soluble proteins, indicating that it is highly conserved among the genotypes. The prolamine profiles showed differences in their banding patterns among the genotypes. Microsatellites detected a higher degree of genetic variation among the sorghum genotypes compared to the prolamines. All the 20 SSR loci were polymorphic with a total of 94 alleles. The number of alleles per locus ranged from 2 to 8, with an average of 4.7. A genetic diversity index (DI) of 0.63 was consistent with the reports of previous publications. This low DI might be due to the lower number of genotypes used in the study, most of which were from the United States. Our results indicate that the genotypes grouped according to their geographical origins. New alleles were detected for some microsatellite loci. This is important especially if a particular locus evaluated in this study is linked to a desirable trait, the data can be used for the identification of the progeny carrying those characteristics. This would further help in the identification of appropriate parental lines for an efficient plant-breeding program and in the conservation and utilization of genetic resources.

* Corresponding author: Dr. N. Ghosh, e-mail address: nghosh@mail.wtamu.edu

† Present address: USDA-ARS, Vegetable Crops Research Unit, Horticulture Department, University of Wisconsin, Madison, WI 53706, USA.

Abbreviations: SDS-PAGE: Sodium Dodecyl Sulfate - Poly Acrylamide Gel Electrophoresis; RAPD: Random Amplified Polymorphic DNA; SSR: Simple Sequence Repeat; AFLP: Amplified Fragment Length Polymorphisms; UPGMA: Unweighted Pair Group Methods with Arithmetic Averages, SAHN: Sequential, Agglomerative, Hierarchical and Nested, NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System.

Introduction

Sorghum is a leading cereal in the arid and semi-arid regions of the world, ranking fifth in importance among the world's grain crops after wheat, maize, rice and barley [1]. It is a C4 crop particularly adapted to the drought prone areas with hot, semi-arid tropical environments, receiving an annual rainfall of 400-600 mm that are too dry for other cereals like maize [2, 3]. It is the chief cereal of human consumption in parts of Asia and Africa. Because of its global socio-economic importance, there is a constant need for the improvement of sorghum. Conservation of important genetic resources and choosing appropriate parental lines for efficient breeding programs are essential for its improvement.

For efficient conservation and breeding practices, characterization and evaluation of genetic diversity within a species is important [4, 5]. Information on various aspects of sorghum diversity will result in the formation of a sorghum core collection that contains a maximum amount of variation [6]. This data can be used as a tool for mining germplasm collections for genomic regions associated with adaptive or agronomically important traits [7]. It would facilitate in identifying appropriate and diverse parental lines for an efficient plant-breeding program [8]. Techniques for evaluating the genetic diversity include morphological characteristics, seed proteins profiling, amplified fragment length polymorphisms (AFLPs), random amplified length polymorphisms (RAPDs), restriction fragment length polymorphisms (RFLPs), simple sequence repeats (SSRs) etc.

Seed storage proteins are characterized by uniform, stable and additive nature that makes them an important tool for the identification of species and also in elucidating the origin and evolution of cultivated plants [9]. The composition of the seed storage proteins is not influenced by changes in either environment or nutritional factors during the growth of the plant [10, 11]. After the pioneering work of Laemmli [12] in 1970, polyacrylamide gel electrophoresis has been used an important analytical tool for numerous studies related to the identification of species, cultivars and F_1 hybrids [13, 14, 15]. It is a relatively low cost, high throughput method for analyzing even the proteins of cereals [16]. Another technique for evaluating the genetic diversity is based on the polymorphism detected by SSRs or microsatellites. These are relatively new class of molecular markers based on tandem repeats of short (2-6 bp) DNA sequences [17]. They are characterized by great abundance, co-dominance, high variability and uniform distribution in plant genomes [18, 19]. These repeats are highly polymorphic, even among closely related cultivars, due to the mutations causing variation in the number of repeating units [20] besides other reasons. Because of these characteristics, microsatellites are one of the preferred genetic markers in plants and animals. It is difficult to clone and sequence SSR-containing DNA fragments from the plant species of interest but once developed, they provide profiles that are highly discriminative among cultivars for many species including maize [21, 22], rapeseed [23, 24], soybean [25, 26], wheat [27, 28]. With the efforts of investigators like Brown et al. [29], Bhattramakki et al. [30], and Menz et al. [31], SSRs are readily available for use as markers in sorghum crop improvement and conservation studies.

Reports are meager in literature on sorghum genetic diversity studies comparing seed protein profiles and SSRs. The purpose of our study was to determine the genetic diversity among the selected sorghum genotypes and the specific objectives were (i) to assess the genetic diversity among 23 sorghum genotypes using seed storage protein profiles and twenty SSR markers and (ii) to compare the genetic variation data obtained with seed protein electrophoresis versus the SSR markers.

Materials and Methods

a. Seed material

Seeds of 23 sorghum genotypes were obtained from the Texas Agricultural Experiment Station, Lubbock, Texas. The complete list of the genotypes, their origin and distinguishing characteristics are shown in Table 1.

Table 1. List of 23 sorghum genotypes, their origin, race and distinguishing characters (According to the USDA, ARS, National Genetic Resources Program).

No.	Genotype	Origin	Race	Line	Characteristics
1	Tx2880	USA	-	R	Resistant to midge; A1 cytoplasm 25% genes from BTx3197; pollinator
2	SC-35	Ethiopia	Durra	A	Post flowering drought tolerant; Stay green
3	Tx642	USA	Durra	B	Post flowering drought tolerant; Stay green; selection from F2 of 2nd back cross in the conversion of IS12555
4	Tx2752	USA	-	B	Resistant to biotype C greenbugs; A1 cytoplasm; widely used as female parent of hybrids; Derivative of BTx399 (Wheatland)
5	ATx378	USA	-	A	A line version of Tx378
6	CSM63	Mali	Guinea	-	-
7	SC630-11E	Zimbabwe	Cafforum Kafir (work group)	B	Excellent grain weathering Resistance; fertility reaction maintainer
8	SC155-14E	Ethiopia	Durra-bicolor	-	work group is subglabrescens Durra dochna type
9	SC414-12E	Sudan	Caudatum kafir (work group)	-	Resistant to sorghum midge and downy mildew pathotypes
10	Ajabsido	Sudan	Caudatum	-	Pre-flowering drought tolerant
11	Lian Tang Ai	China	Kaoling	-	Contains exotic grain yield genes
12	China 17	China	-	-	-
13	NC+5C35	USA	-	-	Proprietary hybrid
14	Pioneer8699	USA	-	-	Proprietary hybrid
15	Tx7078	USA	Kafir	-	Parent is one of the original Sorghum hybrid (RS 610); some tolerance to pre-flowering drought stress
16	Tx3197	USA	Combine kafir 60	B	Female parent of many early hybrids; some tolerance to pre-flowering drought stress
17	RTx430	USA	-	R	Pre-flowering and drought tolerant; Elite US pollinator with excellent combining ability; Yellow endosperm; partially converted IS12661
18	Redland Tx378	USA	-	B	Pre-flowering drought tolerant; A1 cytoplasm
19	SC325-12E	USA	Nigricans	-	-
20	SC176-14E	Ethiopia	Zera zera	-	-
21	San chi San	China	Bicolor	-	Cold tolerant
22	77CSI/Tx2816	Zimbabwe	Kafir	-	-
23	MB9	-	-	B	-

b. Analysis of proteins using SDS-PAGE

Protein extraction: For the extraction of water-soluble proteins, 20 mg of finely ground seed was suspended in 1 ml extraction and loading buffer (50 mM Tris-HCl, pH 8.5; 2% w/v SDS; 0.1% bromophenol blue; 10% glycerol; 2% 2β-ME) and it was carried out according to Bukhari et al. [32]. For the extraction of prolamines, 20 mg of the seed powder was suspended in 200 µl of 70% ethyl alcohol [33]. The prolamines were suspended in 20 µl extraction and loading buffer (1 M Tris-HCl; 0.5% SDS; 0.1% bromophenol blue; 8% glycerol; and 2% 2β-ME). The samples were stored at -20°C until further use.

SDS-PAGE: The proteins were analyzed by SDS-PAGE using an Owl P9DS Dual Gel System (Thermo Scientific, CA). The denaturing gel comprised of 10% separating and 4% stacking gels [12]. The samples were run along with a broad-range (2-212 kDa) protein marker (New England Biolabs, Beverly, MA) and the gel was stained with 0.25% Coomassie blue G-250 for about 12 h followed a destaining till the bands appeared clear. The gel was visualized for further analysis.

c. Analysis of DNA using microsatellite markers

DNA extraction, PCR and agarose gel electrophoresis: For the extraction of genomic DNA, 200 mg leaf tissue was ground into fine powder in liquid nitrogen and dispersed in 600 µl of CTAB extraction buffer (1 M Tris-HCl, pH 8.0; 5 M NaCl; 0.5 M EDTA; 2% 2β-ME; and 2% CTAB) and the extraction was carried out according to Weng et al. [34]. DNA was suspended in Tris-EDTA buffer and the tubes were stored at 4°C until further use. Twenty SSR markers were used to evaluate the genetic diversity (Table 2). One µl DNA was assayed in a reaction volume of 10 µl (comprising 5 µM of each forward and reverse primer and 1× PCR master mix) using the GeneAmp 2720 thermal cycler (Applied Biosystems, CA). A one-fit-all, touch-down PCR program was designed for all the primer sets and the cycling conditions consisted of a 3-min initial denaturation at 95°C, six cycles of 45 s at 94°C, 5 min at 68°C, 1 min at 72°C, with an annealing temperature being reduced by 2 min at 50°C and 1 min at 72°C. The PCR products were resolved on a 4% agarose gel containing ethidium bromide. DNA extracted from Johnson grass was used as an out-group. The fragment sizes were compared to a 100 bp DNA ladder (New England Biolabs, Beverly, MA). The gels were visualized for further analysis.

d. Data analyses

For both seed protein and microsatellite analyses, the presence of a band was scored at "1" and the absence as "0". The binary matrix data were analyzed using the Similarity for Qualitative Data (SIMQUAL) module to generate Dice similarity coefficients [35]. The similarity coefficients were used to construct dendrograms using the Unweighted Pair Group Methods with Arithmetic averages (UPGMA) employing the Sequential, Agglomerative, Hierarchical and Nested (SAHN) clustering method using the Numerical Taxonomy and Multivariate Analysis System (NTSYS-pc, 36) Version 2.20 (Exeter Software, E. Setauket, NY, USA) program [37].

For the SSR analyses, total number of alleles detected and the polymorphism information content (PIC) values were determined for each marker. It refers to the value of a marker for detecting polymorphism within a population, depending on the number of detectable alleles

Table 2. Microsatellite primer sets, linkage group, repeat motif, and flanking sequences used for the study (Menz et al.[31]; Bhattramakki et al. [30]): their PCR product range, number of alleles per locus, and polymorphic information content (PIC) values.

Locus	LG[†]	Repeat	Flanking sequences (5' to 3') of F[‡] and R[‡] primers	Allelic size (bp)	No. of alleles	PIC
Xtxp4	B	(GA)23	F: AAT ACT AGG TGT CAG GGC TGT G R: ATG TAA CCG CAA CAA CCA AG	173	8	0.672
Xtxp10	I	(CT)14	F: ATA CTA TCA AGA GGG GAG C R: AGT ACT AGC CAC ACG TCA C	145	8	0.834
Xtxp18	H	(AG)21	F: ACT GTC TAG AAC AAG CTG CG R: TTG CTC TAG CTA GGC ATT TC	231	3	0.469
Xtxp20	J	(AG)21	F: TCT CAA GGT TTG ATG GTT GG R: ACC CAT TAT TGA CCG TGG AG	217	5	0.782
Xtxp24	D	(TC)21	F: CCA TTG AGC TTC TGC TAT CTC R: TTC TAA GCC CAC CGA AGT TG	145	4	0.511
Xtxp30	E	(AAT)25	F: AAA AAG GAC GCG CAG CTG R: CTG CTC TCC ACC ATC CGT AG	273	6	0.717
Xtxp31	C	(CT)25	F: TGC GAG GCT GCC CTA CTA G R: TGG ACG TAC CTA TTG GTG C	222	5	0.701
Xtxp37	A	(TC)23	F: AAC CTA AGA GGC CTA TTT AAC C R: ACG GCG ACT ATG TAA CTC ATA G	189	3	0.486
Xtxp41	D	(CT)19	F: TCT GGC CAT GAC TTA TCA C R: AAA TGG CGT AGA CTC CCT TG	278	3	0.533
Xtxp43	A	(CT)28	F: AGT CAC AGC ACA CTG CTT R: AAT TTA CCT GGC GCT CTG C	171	7	0.595
Xtxp72	B	(GA)36	F: TTA TGG AAG CAA AAT GAC R: CGA ATC CTA ATT GAG GTA AGC	123	5	0.661
Xtxp145	F	(AG)22	F: GTT CCT CCT GCC ATT ACT R: CTT CCG CAC ATC CAC	238	7	0.740
Xtxp159	G	(CT)21	F: ACC CAA AGC CCA AAT CAG R: GGG GGA GAA ACG GTG AG	169	2	0.469
Xtxp217	J	(GA)23	F: GGC CTC GAC TAC GGA GTT R: TCG GCA TAT TGA TTT GGT TT	175	5	0.702
Xtxp230	I	(GA)28	F: GCT ACC GCT GCT GCT CT R: AGG GGG CAT CCA AGA AAT	191	4	0.624
Xtxp274	F	(TTC)19	F: GAA ATT ACA ATG CTA CCC CTA AAA GT R: ACT CTA CTC CTT CCG TCC ACA T	331	4	0.579
Xtxp285	C	(CTT)11 CTC (CTT)16	F: ATT TGA TTC TTC TTG CTT TGC CTT GT R: TTG TCA TTT CCC CCT TCT TTC TTT T	249	4	0.727
Xtxp292	H	(AC)12	F: CAT TTG CGA AGT TAC AAC ATT GCT R: CAT TCC TGA CTG CCC TCT CC	332	3	0.635
Xtxp295	G	(TC)19	F: AAA TCA TGC ATC CAT GTT CGT CTT C R: CTC CCG CTA CAA GAG TAC ATT CAT AGC TTA	165	3	0.469
Xtxp303	E	(GT)13	F: AAT GAG GAA AAT ATG AAA CAA GTA CCA A R: AAT AAC AAG CGC AAC TAT ATG AAC AAT AAA	160	5	0.702

[†] LG refers to linkage group

[‡] F refers to forward primer and R refers to reverse primer

and the distribution of their frequency [38]. It was calculated according to the following algorithm [39]:

$$PIC_i = 1 - \sum_{j=1}^{n} X_{ij}^2$$

where X_{ij} is the frequency of the i^{th} allele for the j^{th} locus, and totaled over n alleles. PIC values range from 0 (monomorphic) to 1 (very highly discriminative, with many alleles each in equal and low frequency). Genetic diversity of SSR data were estimated by the number of alleles per locus and the mean DI over all the loci was calculated according to Assar et al. [39] using the algorithm:

$$DI = n_a(1/n_l \sum_j (1 - \sum_i x_{ij}^2))/(n_a - 1)$$

where n_l is the number of loci, x_{ij} is the frequency of the i^{th} allele of the locus j and n_a the total number genotypes included in the study.

Results and Discussion

a. Analysis of proteins using SDS-PAGE

The banding patterns for the water-soluble proteins were similar in all the genotypes and hence were not subjected to further analyses. This might be because they are highly conserved among the genotypes. They constitute a minor portion of the total seed protein. The alcohol-soluble prolamines ranged from 16 kDa to 25 kDa. The majority of the prolamines were of the molecular weight 23 kDa and 25 kDa, which are recognized as α-kafirins. This was in accordance to the reports that α-kafirins make up about 80% of the total kafirins. A total of 55 bands were detected, of which seven were electrophoretically different. Maximum number of bands recorded for a genotype was four and was observed in CSM63, Pioneer 8699 and NC+5C35. Band number 1 was detected in 19 of the 23 genotypes. No unique bands were observed in any of the genotypes, showing that there is not much variation among the genotypes with respect to their prolamines. The similarity coefficients ranged from 0 to 1 indicating no similarity to 100% similarity. The UPGMA dendrogram of the sorghum genotypes based on their prolamine profile is shown in Figure 1. The protein data could be analyzed jointly with SSR data to have dendrogram (Figure 1 and Figure 7). Two genotypes originating from China, China 17 and Lian Tang Ai, clustered together. However, the other genotype San Chi San did not show any bands. Genotypes from the US, clustered together as two different groups. Likewise, Pioneer 8699 and NC+5C35, both hybrids from the US grouped together and were 100% similar. Overall, it looks like the genotypes from the same region of origin clustered together and were less similar to the genotypes from other regions. Thus, the genotypes were clustered mainly according to their regional characteristics.

b. Analysis of DNA using microsatellite markers

The 20 SSR markers used in this study detected a total of 94 alleles with an average of 4.7 alleles per locus. Agarose gel profiles of PCR products amplified from the genomic DNA of sorghum genotypes using three different markers (Xtxp10, 41, 43, 295 and 303) are shown in Figures 2, 3, 4, 5 and 6. The numbers of alleles detected per marker are given in Table 2. The maximum number of alleles detected was 8 using the marker Xtxp10 and the minimum was 2 using Xtxp159. The PIC values ranged from 0.469 (for Xtxp18 and Xtxp252) to

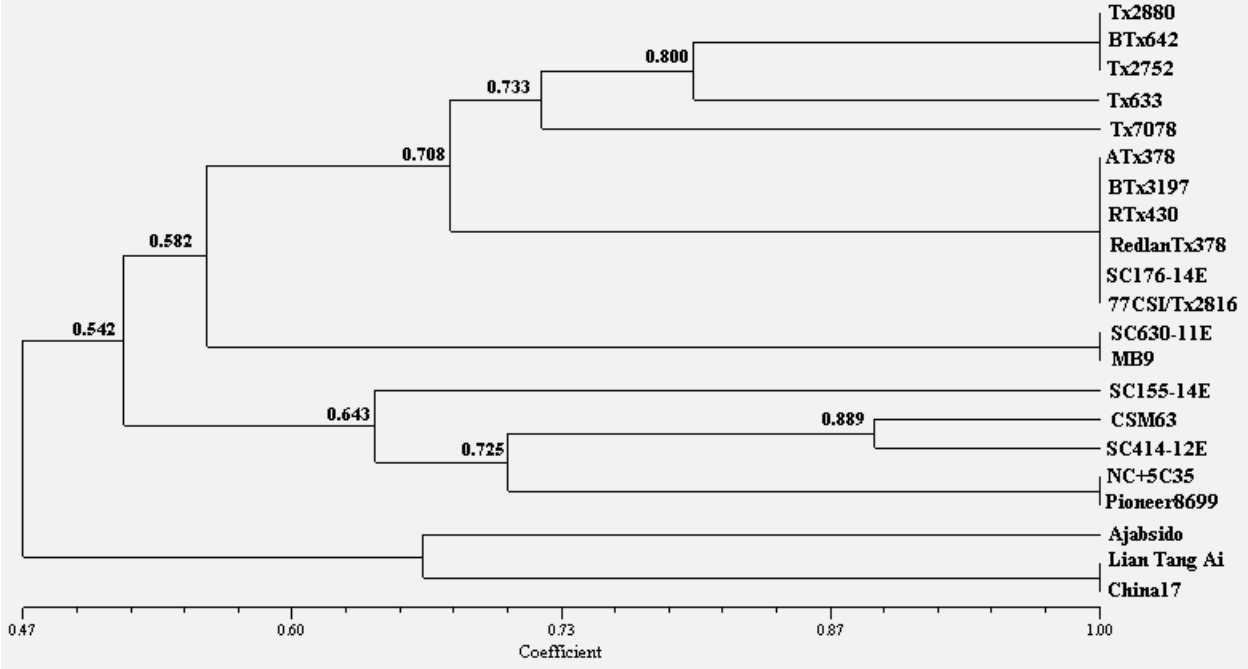
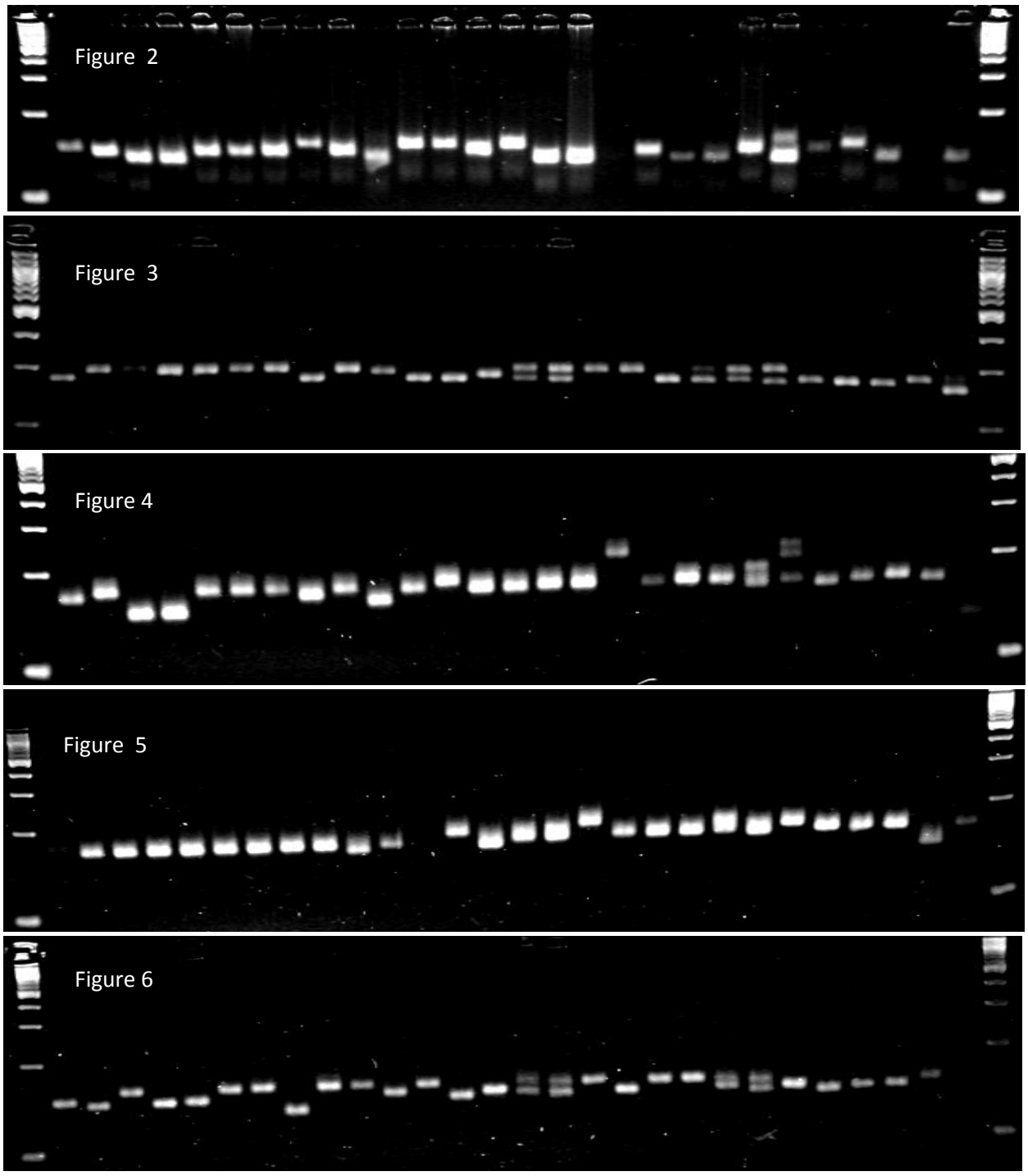


Figure 1. UPGMA dendrogram showing the genetic relationship among 23 sorghum genotypes based on Dice coefficient using prolamine-protein profiles as revealed by SDS-PAGE.

0.839 (for Xtxp10). The number of alleles observed for markers Xtxp4, 10, 145, 292 and 303 was much higher than that reported in other studies but is consistent with the reports of other investigators like Assar et al [39]. The similarity coefficients among the genotypes ranged from 16.3% to 91.4%. The UPGMA dendrogram of the sorghum genotypes based on 20 SSR loci is shown in Figure 7. Hybrids, Pioneer 8699 and NC+5C35, were the most similar among all the genotypes with a coefficient of 0.914. This is consistent with the similarity detected based on their prolamine profiles. The least similar genotypes were MB9 and Johnson grass with a coefficient of 0.121. As expected, Johnson grass had the lowest similarity with most of the sorghum genotypes. Most of the genotypes grouped according to their regions of origin. For instance, SC414-12E and Ajabsido, both from Sudan grouped together and so did China 17 and Lian Tang Ai (both from China); and Pioneer 8699 and NC+5C35 (both hybrids from USA); and the genotypes BTx2752, ATx378, BTx 3197, RTx430, RedlandTx378 (all from USA). The overall DI was 0.63, which is comparable to the DI values obtained by Uptmoor et al. [40] and Assar et al. [39]. However, this is lower than the DI value of 0.897, as obtained by Dje et al. [41]. This relatively low genetic DI might be due to the lower number of genotypes included in this study. In addition, the study conducted by Dje et al. [41] included sorghum germplasm from the world collection, which are expected to be diverse. On the contrary, this study had germplasm mainly from the US. Just two genotypes each from Ethiopia, Sudan, China and Zimbabwe are too less in number to represent the actual diversity found in those regions.

All the 20 SSR markers used in this study detected polymorphism, which allowed the discrimination of each of the 23 sorghum



Figures 2, 3, 4, 5 and 6. Agarose gel profile of PCR products amplified from genomic DNA of sorghum genotypes (in the order as shown in Table 1.) using the primer pair from the 2) SSR locus Xtxp10 3) SSR locus Xtxp41 4) SSR locus Xtxp43 5) SSR locus Xtxp295 6) SSR locus Xtxp303. The extreme two lanes are the DNA markers and the last lane from left is Johnson grass. Note: Genotypes of lanes 2 and 6 are same as lane 7 i.e. they all belong to the genotype ATx378. Genotype of the lane 19 is same as lane 20 i.e. they belong to the genotype RTx430.

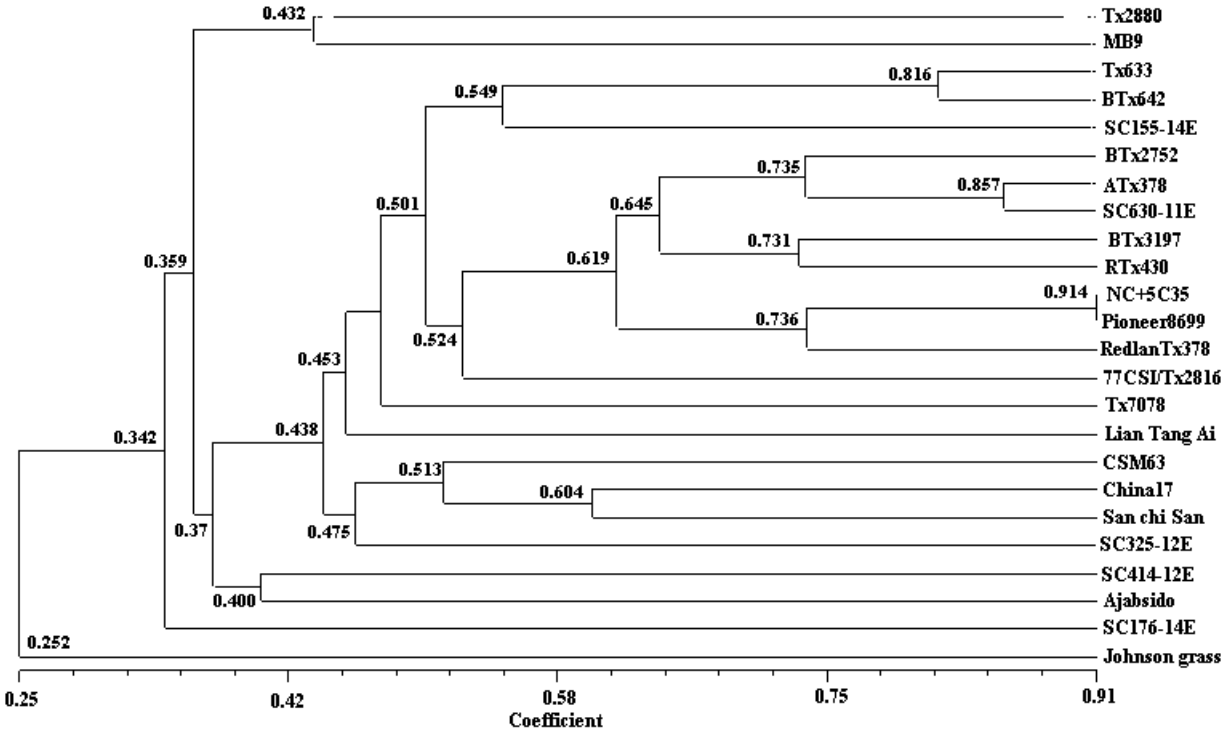


Figure 7. UPGMA dendrogram showing the genetic relationship among 23 sorghum genotypes and Johnson grass based on Dice coefficient as revealed by 20 SSR markers.

genotypes. Even though the number of markers used in this study is comparatively less, they are uniformly distributed across the sorghum genome, with atleast two markers from each of the ten linkage groups. Thus, the selection of 20 SSR markers is justified because it is avoiding both over- and under-representation of certain genomic regions [39]. The UPGMA dendrograms of the prolamine proteins and SSR markers are very different from each other, which is mainly due to the detection of relatively less polymorphism by the prolamines and comparatively high polymorphic nature of the SSRs.

Conclusion

This study provided the first detailed evaluation of sorghum genotypes using water-soluble proteins, prolamine proteins and microsatellites. The water-soluble proteins and the prolamine protein profiles obtained by SDS-PAGE detected no polymorphism and relatively less polymorphism respectively among the sorghum genotypes. Even though these techniques are relatively easy and economic, they are not quite efficient in characterizing the genotypes, especially those belonging to the same species, as is the case here. DNA-based SSRs revealed high genetic diversity among the genotypes and were able to differentiate them successfully. New alleles were detected at some loci, an information that can be exploited by the plant breeders in sorghum improvement. Most of the genotypes selected for this study were assessed for the first time. The genetic diversity information revealed by the SSR makers can be applied in both genetic conservation and crop improvement programs. The data revealed by these markers can be used in marker-assisted selection. If the traits linked to the SSR markers used in this study have already been identified,

these markers can be exploited in the development of sorghum cultivars with desirable characteristics. For example, if a particular locus evaluated in this study is linked to resistance to a disease or a pest, the data can be used for the identification of the progeny carrying the desirable characteristics. This is extremely applicable for characters like greenbug resistance, Striga tolerance etc. which have already been mapped using the SSR markers. This study can facilitate sorghum breeders in the development of inbred lines and hybrids by assisting them in the selection of diverse germplasm with agronomically important traits. This molecular information in combination with the field evaluation results into characteristics for drought stress tolerance and pest resistance will help in an efficient selection of the parents. Future studies aiming at integrating both the genetic and agronomic results are necessary for successful sorghum breeding programs.

References

1. Bantilan MCS, Deb UK, Gowda CLL, Reddy BVS, Obilana AB and Evenson RE. 2004. Sorghum genetic enhancement: research process, dissemination and impacts. [Online]. http://agropedia.iitk.ac.in/openaccess/?q=content/sorghum-genetic-enhancement-research-process-dissemination-and-impacts(Verified 7 May. 2010).
2. Doggett H. 1988. Sorghum, 2nd ed., John Wiley and Sons, Inc., New York.
3. Rooney WL, 2004. Sorghum improvement—integrating traditional and new technology to produce improved genotypes. Advances in Agronomy 83: 37–109.
4. Bretting PK and Widrlechner MP. 1995. Genetic markers and plant genetic resources management. Plant Breed. Rev. 13:11-86.
5. Garcia E, Jamilena M, Alvarez JI, Arnedo T, Oliver JL and Lozano R. 1998. Genetic relationships among melon breeding lines revealed by RAPD markers and agronomic traits. Theor. Appl. Genet. 96:878-885.
6. Maqbool SB, Devi P and Sticklen MB. 2001. Biotechnology: Genetic improvement of sorghum (*Sorghum bicolor* (L.) Moench). In Vitro Cell. Dev. Biol.-Plant 37:504-515.
7. Casa AM, Mitchell SE, Hamblin MT, Sun H, Bowers JE, Paterson AH, Aquadro CF and Kresovich S. 2005. Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. Theor. Appl. Genet. 111:23-30.
8. Ribaut JM and Hoisington D. 1998. Marker-assisted selection: new tools and strategies. Trends in Plant Science 3:236-239.
9. Ladizinsky G and Hymowitz T. 1979. Seed protein electrophoresis in taxonomic and evolutionary studies. Theor. Appl. Genet. 54:145-151.
10. Dunnill M, Patricia, and Fowden L. 1965. The amino acids of seeds of the *Cucurbitaceae.* Phytochemistry 4:933-944.
11. Adriaanse A, Klop W and Robbers JE. 1969. Characterisation of *Phaseolus vulgaris* cultivars by their electrophoretic patterns. J. Sci. Fd Agric. 20647-650.
12. Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.
13. Cherry JP, Katterman FRH and Endrizzi JE. 1970. Comparative studies of seed proteins of species of *Gossypium* by gel electrophoresis. Evolution 24:431-447.
14. Bianchi-Hall CM, Keys RD, Stalker HT, and Murphy JP. 1993. Diversity of seed storage protein patterns in wild peanut (*Arachis, Fabaceae*) species. Plant Syst. Evol. 186:1-15.
15. Ahmad M, Frautrier AG, Burritt DJ and McNeil DL. 1997. Genetic diversity and relationships in *Lens* species and their F_1 interspecific hybrids as determined by SDS-PAGE. New Zealand Journal of Crop and Horticultural Science 25:99-108.
16. Bean SR and Lookhart GL. 2000. Electrophoresis of cereal storage proteins. J. Chromatogr. 881:23-36.
17. Litt M and Luty JA. 1989. A hyper-variable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am. J. Hum Genet. 44:397-401.
18. Lagercrantz U, Ellegren H and Anderson L. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res. 21:1111-1115.
19. Akkaya MS, Shoemaker RC, Specht JE, Bhagwat AA and Cregan PB. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci. 35:1349-1445.
20. Yang, W., A.C. de Oliveira, I. Godwin, K. Schertz, and J. L. Bennetzen. 1996. Comparison of DNA marker technologies in characterizing plant genome diversity: variability in Chinese sorghums. Crop Sci. 36:1669-1676.
21. Smith JSC, Kresovich S, Hopkins MS, Mitchell SE, Dean RE, Woodman WL, Lee M and K. Porter. 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Sci. 40:226-232.
22. Matsuoka Y, Mitchell SE, Kresovich S, Goodman M and Doebley J. 2002. Microsatellites in *Zea* – variability, patterns of mutations, and use for evolutionary studies. Theor. Appl. Genet. 104:436-450.
23. Kresovich S, Szewc-McFadden AK, Blick SM and McFerson JR. 1995. Abundance and characterization of simple-sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). Theor. Appl. Genet. 91:206-211.

24. Charters YM, Robertson A, Wilkinson MJ and Ramsay G. 1996. PCR analysis of oilseed rape cultivars (*Brassica napus* L. spp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet. 92:442-447.
25. Akkaya MS, Bhagwat AA and Cregan PB. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132:1131-1139.
26. Rongwen, J., M.S. Akkaya, A.A. Bhagwat, U. Lavi, and P.B. Cregan. 1995. The use of microsatellite DNA markers for soybean genotype identification. Theor. Appl. Genet. 90:43-48.
27. Plaschke J, Ganal MW and Roder MS. 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theor. Appl. Genet. 91:1001-1007.
28. Roder MS, Plaschke J, Konigh SU, Borner A, Sorrels ME, Tanksley SD and Ganal MW. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. Mol. Gen. Genet. 24:327-333.
29. Brown SM, Hopkins MS, Mitchell SE, Senior ML, Wang TY, Duncan RR, 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.
30. Bhattramakki D, Dong J, Chhabra AK and Hart GE. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* using SSRs and RAPDs. African Journal of Biotechnology 2:334-340.
31. Menz MA, Klein RR, Mullet JE, Obert JA, Unruh NC and 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.
32. Bukhari YM, Koivu K and Kanerva A. 1997. Seed storage protein size variations and their significance in the evolution and the systematics of *Acacia* and *Prosopis (Mimosaceae)*. Silvae Genetica 47:241-249.
33. Krishnan HB, White JA and Pueppke SG. 1989. Immunocytochemical analysis of protein body formation in seeds of *Sorghum bicolor*. Can. J. Bot. 67:2850-2856.
34. Weng, Y., W. Li., R.N. Devkota, and J.C. Rudd. 2005. Microsatellite markers associated with two *Aegilops tauschii*-derived greenbug resistance loci in wheat. Theor. Appl. Genet. 110:462-469.
35. Dice LR. 1945. Measures of the amount of ecological associations between species. Ecology 26: 297-302.
36. Sneath PHA and Sokal RR. 1973. Numerical taxonomy. Freeman, San Fransisco, CA.
37. Rohlf FJ. 2000. NTSYS-pc numerical taxonomy and multivariate analysis system, version 2.20. Applied Biostatistics, Exeter Publ, New York.
38. Ni J, Colowit PM and Mackill DJ. 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci. 42:601-607.
39. Abu Assar AH, Uptmoor R, Abdelmula AA, Salih M, Ordon F, and Friedt W. 2005. Genetic variation in sorghum germplasm from Sudan, ICRISAT, and USA assessed by simple sequence repeats (SSRs). Crop Sci. 45:1636-1644.
40. Uptmoor R, Wenzel W, Friedt W, Donaldson G, Ayisi K and 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.
41. Dje Y, Forcioli D, Ater M, Lefebvre C and Vekemans X. 1999. Assessing population genetic structure of sorghum landraces from Northwestern Morocco using allozyme and microsatellite markers. Theor. Appl. Genet. 99:157-163.