Mol Breeding (2008) 21:497–509
DOI 10.1007/s11032-007-9149-z

Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers

M. L. Ali · J. F. Rajewski · P. S. Baenziger · K. S. Gill · K. M. Eskridge · I. Dweikat

Received: 25 April 2007 / Accepted: 3 December 2007 / Published online: 25 December 2007

Abstract Sweet sorghum (*Sorghum bicolor* L.) is a type of cultivated sorghums and has been recognized widely as potential alternative source of bio-fuel because of its high fermentable sugar content in the stalk. A substantial variation of sugar content and related traits is known to exist in US sweet sorghum. The objectives of the study were to assess the genetic diversity and relationship among the US sweet sorghum cultivars and lines using SSR markers and to examine the genetic variability within sweet sorghum accessions for sugar content. Sixty-eight sweet sorghum and four grain sorghum cultivars and lines were genotyped with 41 SSR markers that generated 132 alleles with an average of 3.22 alleles per locus. Polymorphism information content (PIC) value, a measure of gene diversity, was 0.40 with a range of 0.03–0.87. The genetic similarity co-efficient was estimated based on the segregation of the 132 SSR alleles. Clustering analysis based on the genetic similarity (GS) grouped the 72 sorghum accessions into 10 distinct clusters. Grouping based on clustering analysis was in good agreement with available pedigree and genetic background information. The study has revealed the genetic relationship of cultivars with unknown parentage to those with known parentage. A number of diverse pairs of sweet sorghum accessions were identified which were polymorphic at many SSR loci and significantly different for sugar content as well. Information generated from this study can be used to select parents for hybrid development to maximize the sugar content and total biomass, and development of segregating populations to map genes controlling sugar content in sweet sorghum.

Keywords Cluster analysis · Genetic distance · Genetic diversity · Genetic similarity co-efficient · SSRs · Sweet sorghum

M. L. Ali · J. F. Rajewski · P. S. Baenziger · I. Dweikat (✉)
Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE 68583-0915, USA
e-mail: idweikat2@unl.edu

K. M. Eskridge
Department of Statistics, University of Nebraska, Lincoln, NE 68583, USA

K. S. Gill
Department of Crop and Soil Sciences, Washington State University, Pullman, WA 99164, USA

Introduction

Sweet Sorghum (*Sorghum bicolor* L.) Moench) is one of many types of cultivated sorghum, noted for its high sugar content in the stem juice. Sweet sorghum is adapted to widely differing climatic and soil conditions. It tolerates drought and high-temperature stress better than many crops and has the capability of remaining dormant during the driest periods

(Gnansounou et al. 2005). Ethanol can be produced from any sugary or starchy material and is presently produced mainly from sugarcane (*Saccharum officinarum* L.) and corn (*Zea mays* L.). However, sugarcane and corn are very energy and water intensive. For drier climates, there is a need to produce bio-energy from more water efficient crops, such as sweet sorghum. Sweet sorghum is a very efficient source of bio-energy as it uses C4 photosynthesis to produce sucrose as a storage molecule, which can be directly fermented. Hence, it is highly imperative to breed new cultivars of sweet sorghum with high sugar content in combination with other desirable agronomic traits.

Success of any crop breeding program is based on the knowledge of and availability of genetic variability for efficient selection. Genetic similarity (or genetic distance) estimates among genotypes are helpful in selecting parental combinations for creating segregating populations so as to maintain genetic diversity in a breeding program (Becelaere et al. 2005) and the classification of germplasm into heterotic groups for hybrid crop breeding (Menz et al. 2004). The search for and establishment of heterotic groups can be based on geographical origin, agronomical traits, pedigree data or on molecular marker data (Melchinger 1999). Before the advent of molecular genetic tools, genetic diversity was estimated from pedigree or agronomic and morphological characteristics. However, the estimates based on pedigree information are generally inflated and often found unrealistic (Fufa et al. 2005; Almanza-Pinzon et al. 2003; Cox et al. 1986; Souza and Sorrells 1989). The morphologically based genetic diversity estimates suffer from the drawback that morphological characteristics are limited in number and are influenced by the environment (van Beuningen and Busch 1997). Therefore, neither pedigree-based nor morphologically-based estimates may reflect the actual genetic difference of the studied populations. On the other hand, molecular markers are not influenced by environment and reflect genetic similarity (and differences) and do not require previous pedigree information (Bohn et al. 1999) which is valuable for crops where pedigree information is lacking.

Various types of molecular markers are available for genome analysis in the grass family. Simple sequence repeats (SSRs) in particular have been reported to be very useful to analyze the structure of germplasm collections as these are abundant, codominant, multiallelic, highly polymorphic and chromosome-specific (Ahmad 2002; Huang et al. 2002; Parker et al. 2002). SSR markers have been extensively used in genetic diversity studies in many plants, including wheat (*Triticum aestivum* L.; Fufa et al. 2005; Mahmood et al. 2004), pearl millet (*Pennisetum glaucum* L.; Budak et al. 2003), sorghum (Casa et al. 2005; Smith et al. 2000), triticale (*X-Triticosecale* Wittmack; Kuleung et al. 2006) and maize (Smith et al. 1997).

Sweet sorghum varieties have been growing in various regions of USA for the last 150 years. The first agronomic sorghum was a sweet sorghum, Chinese Amber, introduced from China in 1853 (Smith and Frederiksen 2000). This was followed by a number of sweet and forage type introductions from Africa, China and Australia. The origin and pedigree information for many of the introductions and derived cultivars are not available or poorly documented. Hence, a study on genetic relationship among sweet sorghum cultivars (with and without parentage information) will help determine their genetic relationships. In this study, we collected 68 sweet sorghum cultivars/lines (old and new) that were grown before or are being presently grown in different regions across the USA to assess their genetic diversity and to evaluate their sugar yielding potential as a key aspect for potential biofuel production.

The objectives of the study were (i) to estimate the genetic diversity and relationship present among the sweet sorghum cultivars and lines using SSR markers and (ii) to examine the genetic variability among sweet sorghum accessions for sugar content and to identify the diverse sweet sorghum cultivars and lines.

Materials and methods

Plant materials

Sixty-eight sweet sorghum and four grain sorghum cultivars and breeding lines grown in different states in USA since 1850s were selected to evaluate the genetic diversity and the relationship among them.

The list of the cultivars and lines with their designation or accession number, registration year, possible place of origin or reporting and their pedigree or genetic background information (if available) are in Table 1. Two cultivars, Wheatland and KS78 and two inbred lines, RTx632 and RTx430 were included as reference lines to grain sorghum. Seeds were obtained from USDA-ARS, University of Nebraska-Lincoln, Nebraska; National Center for Genetic Resources Preservation (NCGRP), 111 South Mason, Fort Collins, Colorado; National Plant Germplasm System, Griffin, Georgia; Texas Agricultural System Station, College station, Texas and University of Kentucky, Kentucky.

DNA extraction

Sap extraction method was used to isolate gnomic DNA from fresh tissues of each cultivar/line. About one gram of fresh leaves from 2–3 weeks-old seedlings were placed in between the two rollers of a sap extraction apparatus (Ravenel Specialties, Seneca, SC) and 5 ml of extraction buffer (50 mM Tris–Hcl, 25 mM EDTA, 1 M NaCl, 1% CTAB, 1 mM of 1, 10-phenanthroline, 0.15% 2-mercaptoethanol) was slowly added to the rollers. The solution mixed with extracted sap was collected in a 15 ml Falcon tube. The extraction procedure was followed as described by Kuleung et al. 2004). The extracted DNA was then re-suspended in 500 µl of TE buffer and the DNA concentration quantified by spectrophotometry (TKO100 Fluorometer, Hoefer Scientific Instruments, San Francisco, CA).

PCR conditions and gel electrophoresis

Sorghum simple sequence repeats (SSR) primers were synthesized based on the published sequence information (Schloss et al. 2002). The PCR reaction mix preparation, PCR conditions, gel electrophoresis and gel staining were performed as described by Kuleung et al. (2004). Annealing temperature was varied from 50 to 58°C which was determined after an initial amplification test with all the primers on two accessions. The gel was photographed using Bio-Rad Gel Doc2000 gel documentation system (Bio-Rad, Hercules, CA).

Genetic diversity estimation

The amplified fragments of each SSR marker were scored as "1" and "0", where "1" indicated the presence of a specific allele (band) and "0" indicated its absence. Polymorphism information content (PIC) of SSR markers was calculated using the formula developed by Anderson et al. (1993). $\mathrm{PIC} = 1 - \sum P_{ij}^2$, where relative frequency of the *j*th allele for the *i*th locus summed across all the alleles for the locus over all lines. PIC provides an estimate of the discriminatory power of a locus by taking into account, not only the number of alleles that are expressed, but also the relative frequencies of those alleles. PIC values range from 0 (monomorphic) to 1 (very highly discriminative), with many alleles in equal frequencies. Senior et al. (1998) opined that PIC is synonymous with the term 'gene diversity' as described by Weir (1996).

Genetic diversity estimate related analyses were done using NTSYSpc ver.2.02i (Rohlf 2000). Genetic similarities (GS) between pairs of accessions were measured by the DICE similarity coefficient based on the proportion of shared alleles (Dice 1945; Nei and Li 1979) with SIMQUAL module. Genetic distances between pairs of lines were estimated as GD or $D = 1 - GS$. The clustering of accessions was done based on a similarity matrix using an unweighted pair group method with arithmetic average (UPGMA) algorithm following SAHN module. The clustering result was used to construct a dendrogram following TREE module.

An analysis of variance (ANOVA) was conducted by PROC ANOVA (SAS Institute 1999–2001) to test significant differences among the groups generated by the cluster analysis for brix content. To determine the differences between the pairs of groups (group means) for brix content, t-tests (LSD) were conducted.

Phenotypic evaluation

Sweet sorghum cultivars and lines included in the study were grown in the Lincoln field nursery during spring-summer 2006 for evaluation of anthesis date, plant height, brix content, wet stalk weight, dry stem weight, stem diameter and stalk moisture. The entries were planted on May 10 in single-row plots of 6.0 m long with a spacing of 8.0 cm between the plants and 0.76 m between the rows. The plants were harvested

Table 1 Pedigree and origin of sweet and grain sorghum accessions used in the genetic diversity assessment

Cultivar/line name	Designation/ accession no.	Registration year	Place of origin/ report	Pedigree/Background information
Dale	NSL 74333	1973	Mississippi	Tracy/MN 960 (PI 152857)
Tracy	NSL 4029	1955	Mississippi	White African (Mer. 51-2)/Sumac
Della	PI 566819	1993	Virginia	Dale*2/ATx622
White African	NSL 3985	1936	Mississippi	Introduction from South African and was also designated as Mer.51-2
Wray	NSL 117772	1981	Mississippi	PI 152728 (Mer.57-1)//Brawley /Rio
M81E	NSL 174431	1983	Mississippi	Brawley//Brawley/Rio
Theis	CSR 216	1978	Mississippi	(Wiley/C.P. Special)/(MN1054/White African)/Mn660
Brawley	NSL 4346	1960	California	White Collier/Rex
Kansas Collier			Kansas	Selection from variety Collier (Undendebule), an introduction from South Africa
White Collier			Kansas	Selection from Collier (Undendebule)
African Millet			Kansas	Selection from Neeazana, an introduction from Natal, Africa. Related to Orange sorgos
Colman	NSL 3978	1936	Kansas	Orange/Amber
Heirloom Sugarcane				Centuries-old variety of unknown origin. Mature cane heart can be chewed like candy.
Brandes	NSL 29336	1974	Mississippi	Colllier 706-C/MN1500 (PI 154844)
Bailey	NSL 187557	1984	Georgia	Wiley (Collier/MN822//MN2046)/Tracy
Smith	PI 511355	1988	Texas	Mer. 81-2 (MN 4004/Mer. 61-11)
Ramada	NSL107377	1980	Mississippi	Mer.45-45/MN1056//MN1054/MN1060
Roma				Sudan grass type variety grown in Texas
Top 76-6	PI 583832	1994	Georgia/Mississippi	Mer. 60-2 {PI154844 (MN1500)/PI 152967 (MN1056)}/Brandes
Grassl	PI 154844	1988	Mississippi	Introduction from Uganda and designated as MN1500 (PI 154844)
Rio	NSL 40230	1972	Mississippi	Rex (MN23)/MN1048 (PI 152959)
Keller	NSL 165819	1982	Mississippi	Mer. 50-1/Rio
Norkan	NSL 4002	1942	Kansas	Atlas/Early Sumac
KS73		1981	Kansas	(Highland/Atlas)/Dwf.Jpn.Bromcorn
KS75		1981	Kansas	KS9//KS9/Dwf.Jpn. Broomcorn
KS76		1981	Kansas	KS9//KS9/Dwf.Jpn. Broomcorn
KS78		1981	Kansas	Redlan [(Kafir/Milo)/Blackhull Kafir]//Redlan/ Dwf.Jpn.Broomcorn
Wheatland	CIso 918	1936	Oklahoma	Milo/Kafir
Sugar Drip	NSL 3991	1936		Parentage unknown but it is presumed to belongs to Orange type of sorgo
Early Sumac	NSL 3970	1936	Kansas	Selection from Standard Sumac
Northern sugarcane				Adapted in northern cold climate. Origin unknown
N98	PI 535783	1990	Nebraska	(Waconia //AN39/N4692-Rio) /Fremont
N99	PI 535784	1990	Nebraska	Fremont/Theis
N100	PI 535785	1990	Nebraska	Waconia/wray
N108	PI 535793	1990	Nebraska	Inbred derived from Saccharum-sorgo
N109	PI535794	1990	Nebraska	Inbred derived from White Collier/Grain sorghum line
N110	PI535795	1990	Nebraska	Inbred developed from Red X
N111	PI 535796	1990	Nebraska	Inbred developed from Waconia-L
Waconia -L	NSL3978	1936	Nebraska	Belongs to Orange group of sorgos
Kansas Orange	NSL3977	1936	Kansas	Selection from Orange which was selected from Neeazana, an introduction from South Africa
Red X			Kansas	Resembles Orange sorgos having red glumes
Mennonite			Missouri	Old-fashioned cane sorghum with red-hulled seeds

Table 1 continued

Cultivar/line name	Designation/ accession no.	Registration year	Place of origin/ report	Pedigree/Background information
Rox Orange			Georgia	Known as GA04 Sugar cane belongs to orange group of sorgo
Georgia.Blue Ribbon			Mississippi	Uncertain origin with brown medium-sized seeds
Fremont Cane			Kentucky	Forage sorghum variety of unknown origin
Simon			Kentucky	Early maturing variety of unknown origin adapted in Kentucky
Sumac	PI 35038	1936	Kansas	Selection from an introduction from Natal, Africa. Also was known as Standard Sumac.
Rex	NSL 3988	1936		Also designated as MN23
Honey	NSL 4030	1936		Also known as Honey Drip, Selection from an early introduction
Greenleaf	NSL 4028	1955	Texas	Leoti-Sudan 2/Leoti- Sudan 4
Lahoma	CSR 214	1960	Oklahoma	Leoti Sorgo/*2 Sudangrass
Rancher	NSL 4016	1949	Texas	39-30-S//39-30-S/10-30-S. Both parental lines were selections from Dakota Amber
Leoti	NSL3983	1936	Kansas	Uncertain Origin. Waxy Amber from china was in its parental background
Red Amber (TX)	PI 17548	1936	Texas/Kansas	Selection from introduction from Australia having dark red glumes
Chinese Amber	PI 22913	1936	Indiana	Introduction from china via France. Served as progenitor for many early sorghum varieties/lines
Dakota Amber	NSL 3974	1936	Texas	Selection from Early Amber
Minnesota Amber	NSL 3972	1936	Texas	Selection from Early Amber
Black Ambercane			Kansas	Selection from Black Amber, an early introduction
Black Amber (NE)			Nebraska/Indiana	Originally selected from Chinese Amber sorgo. It was also known as Early Amber
Kansas Amber			Kansas	Selection from early Amber
Early Folger	NSL 3984	1936	Kansas	Selection from Early Amber (introduction from Africa, first recognized variety)
Cowley	NSL 189405	1985	Texas	Mer.64-7/Mer.64-6
Williams			Kentucky/Georgia	Adapted in Kentucky and Georgia
Umbrella			Kentucky/Vermont	Widely grown in Kentucky and Vermont
Tx632	PL-161	1986	Texas	(SC120-6/RTx7000)-1-2-1
RTx430	PL 140	1984	Texas	Tx2536/SC170-6-51-E2
Blacktop			Kentucky	Variety with black dense heads on stiff stems
Snowflake				
Hastings				
Iceberg				
White Orn	PI563429		Nebraska	Nebraska selection (IS10715) from Togo, Africa

during the first week of October. Data was recorded from three randomly sampled plants from each entry. Brix determinations (a measure of sugar content) were made with a hand-held refractometer by reading juice samples obtained from the second basal internode of the stalk. The degrees brix scale measures percent total solids in the stalk juice including sucrose, reducing sugars and other dissolved solids which are neither crystallisable nor fermentable.

Results and discussion

Allelic diversity at SSR loci

Forty-one SSR markers that generated 132 alleles were used to estimate the genetic diversity among 72 entries. The number of alleles revealed by each marker ranged from 2 to 10 with an average of 3.22 per marker. The polymorphism information content

(PIC) value for the SSR loci ranged from 0.03 to 0.87 with an average of 0.40.

The mean number of alleles per SSR locus (3.22) detected on 72 sorghum accessions was similar to that detected by Schloss et al. (2002) on 25 sorghum lines (3.4) but lower than that reported by Agrama and Tuinstra (2003) and Smith et al. (2000) with mean allele per locus of 4.3 and 5.9, respectively. The gene diversity observed in this population (Mean PIC value = 0.40) that is lower than the diversity value (0.46, 0.62, 0.58) reported by Schloss et al. (2002), Agrama and Tuinstra (2003) and Smith et al. (2000), respectively. The SSR loci which produced a higher number (4–10) of alleles, such as, Xcup01, Xcup5, Xcup14, Xcup47, Xcup50, Xcup67, Xcup73 and Xcup74 revealed a high gene diversity (PIC value) that ranged from 0.58 to 0.87 per locus. Results for these markers corresponded with the findings of Schloss et al. (2002). All these markers except Xcup50 contained di-nucleotide repeats. Smith et al. (2000) used only 15 SSR markers of which 13 were di-nucleotide repeats while Agrama and Tuinstra used SSR markers that had only di-nucleotide repeats. In our study, only 19 out of 41 SSRs contained di-nucleotide repeats. The low number of SSR markers with di-nucleotide repeats used in this study might be one of the reasons for the low allelic diversity in this population. In general, SSRs with di-nucleotide repeats displayed a higher number of alleles than tri- and tetra-nucleotide repeats, and a direct relationship exists between marker information content and the number of repeat units (Weber 1990; Senior et al. 1998; Innan et al. 1997; Smith et al. 1997). Another possible reason for relatively low gene diversity in the present set of cultivars and lines could be that many of the cultivars and lines used in this study were either developed from crosses involving parents that were selected from the early introductions of sweet sorghum cultivars or were direct selections from early introductions with narrow genetic base. Many of the accessions (Table 1) had common parents in their genetic backgrounds. This common parentage among the cultivars and lines was also a reason for lower gene diversity that was observed in the present study. Agrama and Tuinstra (2003) stated that the high levels of allelic diversity of SSR markers observed in their study was probably due to the presence of an extensive genetic diversity in their sorghum genotypes that represented different races and geographic areas. Similarly, Smith et al. (2000) deliberately chose inbreds that encompassed relatively broad array of germplasm diversity in grain sorghum.

Two SSR loci, Xcup33 and Xcup63 were identified to possess rare alleles and exhibited gene diversity index of 0.03 and 0.05, respectively. Rare alleles are defined as a frequency of <0.05 (Somers et al. 2007; Casa et al. 2005). These rare alleles could be of particular interest as they are uniquely linked to some particular genotypes. Such alleles are important because they may be diagnostic for particular genotypes or for particular regions of the genome specific to a particular type of sorghum (Arama and Tuinstra 2003).

Cluster analysis and genetic diversity

Based on the 132 shared alleles, genetic similarity co-efficient was estimated for each pair of the 72 sorghum genotypes which ranged from 0.26 to 1.00. Except for N110 and N111, the dendrogram (Fig. 1) clearly discriminated all the 72 sorghum cultivars and lines. N110 and N111 were remained together with a genetic similarity co-efficient of 1.00. Both inbreds were developed in Nebraska in 1990. N110 was developed from Red-X while N111 was developed from variety Waconia-L and both parental cultivars had their origin from Orange group of sorgos cultivars descended from early introductions. The greatest genetic diversity in this study was observed between Greenleaf and Bailey followed by Grassl vs Blacktop with genetic distance (D) value of 0.735 and 0.732, respectively. Greenleaf originated from a Sorgo cross with Sudangrass while Bailey had its origin from Collier, a saccharine sorghum variety.

The cluster analysis grouped the 72 sorghum cultivars and lines into 2 main groups. Group A included 63 and group B included only 9 accessions. Group A was again sub-divided into sub-groups I through VIII while two sub-clusters of group B were named as group IX and X (Fig. 1). Eighteen genotypes were clustered in group I, which were further grouped into 4 sub-subclusters. Five genotypes, namely, Dale, Tracy, Della, White African and Wray were clustered in group Ia. Dale, Tracy, Della and

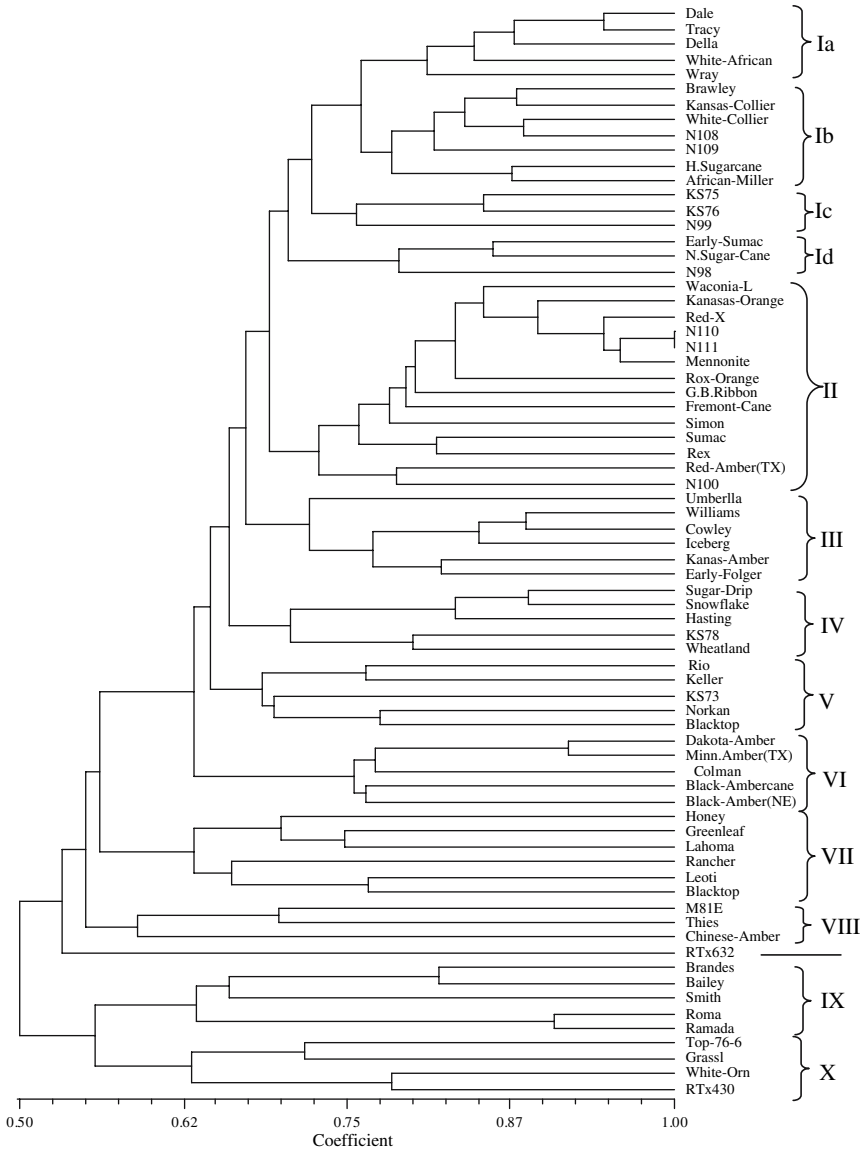

Fig. 1 Dendrogram of 68 sweet and 4 grain sorghum accessions revealed by cluster analysis of genetic similarity estimates generated by Nei and Li coefficient based on 41 SSR markers

White African shared a common genetic background while Wray is unrelated from rest of the entries with regard to the apparent pedigree relationship. Group Ib included accessions Brawley, Kansas Collier, White Collier, N108, N109, Heirloom (H) Sugarcane and African Millet. Out of these seven, four cultivars, such as, Brawley, Kansas Collier, White Collier and N109 had common a parent, Collier (Table 1). In group Ic, out of three, two cultivars, KS75 and KS76, developed in Kansas were clustered together as they shared the common parent in their pedigree. Inclusion of N99 in this group could not be explained by parentage information. In group 1d, three cultivars, Early Sumac, Northern Sugar Cane (no pedigree information was available) and N98 were clustered together. N98, an inbred developed in Nebraska, has Waconia, an Orange type of sorgos, in its genetic background. Similarly, Early Sumac, a selection from Standard Sumac, which in turn, was selected from early introductions and possibly, was related to the Orange group of sorgos in their ancestry.

In group II, 14 genotypes were clustered together (Fig. 1). Out of 14, 8 cultivars/lines, namely, N100, N110, N111, Kansas Orange, Waconia-L, Red X, Rox Orange and Rex belong to the orange group of

sorgos, and therefore, a relationship among them was expected. Pedigree information for four cultivars in this group, namely, Mennonite, Georgia Blue Ribbon (G. B. Ribbon), Simon and Fremont Cane were not available. Out of six cultivars in group III, Kansas Amber and Early Folger belong to the Amber group of sweet sorghum and thus, an association was expected between them. Cowley, a sweet sorghum variety developed in Texas in 1984 was in this group with three other cultivars, Icerberg, Williams and Umbrellla. No parentage information was available for Iceberg, Williams and Umbrella. Possibly, they were the descendants from early introductions having ancestral relationship with Amber group of sorgos. In group IV, three sweet sorghum cultivars, Sugar Drip, Snowflake, and Hastings and two grain sorghum cultivars, KS78 and Wheatland were clustered together. No parentage information was available for Sugar Drip, Snowflake and Hastings. KS78 and Wheatland shared the common genetic backgrounds of Milo and Kafir in their pedigree and paired together with a similarity co-efficient of 0.79.

In group V, five cultivars were included, out of which Rio and Keller developed in Mississipi were closely related because Rio was one of the parents of Keller (Table 1, Fig. 1). Similarly, grouping of two Kansas varieties, Norkan and KS73 together could be ascribed to their parental relationship where cultivar Atlas was common in their parentage. The parentage information for the fifth cultivar Blacktop was not known. The group VI cluster included five cultivars of which four were Amber type of sweet sorghum, Dakota Amber, Minnesota Amber, Black Ambercane and Black Amber (NE) and all had Early Amber in their ancestry. The fifth cultivar Colman was a hybrid between Amber and Orange types of sorghum. Therefore, the grouping of these five cultivars together was expected. Group VII cluster included six cultivars, Honey, Greenleaf, Lahoma, Rancher, Leoti and Blacktop (ent.3) with genetic similarity in between 0.77 and 0.63. Out of six cultivars, Lahoma, Greenleaf and Leoti shared the genetic background of Leoti (Table 1). Leoti is presumed to have Amber in parental background while Rancher was originated from parents having Amber in their genetic background. Two Mississippi cultivars, M81E and Theis and the first introduced variety, Chinese Amber, were clustered together in Group VIII. Both M81E and Theis had Collier in their parental backgrounds. RTx632, a grain sorghum inbred, grouped alone at the extreme end of this group.

The two sub-clusters of group B were related with a similarity coefficient of 0.55. Five cultivars, namely, Brandes, Bailey, Smith, Roma and Ramada were clustered in group IX of which Brandes and Bailey are related based on parental relationship. The reason for grouping of other widely grown sweet sorghum varieties, Ramada, Smith and Roma could not be explained based on their available pedigree information. In group X, only four cultivars or lines, Top 76-6, Grassl, White Orn and RTx430 were clustered together. Top 76-6 and Grassl shared a common parent MN1500, an introduction from Uganda. Brandes, Ramada and Grassl were developed in Mississippi while Bailey was developed in Georgia. Cultivar Top 76-6 was jointly released by the University of Georgia and Mississippi State University. RTx430, a grain sorghum parental line, was positioned at the extreme end on the dendrogram (Fig. 1). RTx430 paired with White Orn with a similarity co-efficient of 0.76. Dean et al. (1999) observed RTx430 to be grouped alone from rest of the 'Orange' accessions of sweet sorghum included in his study. No background information about White Orn was available.

Some accessions that shared a common parentage were found to be clustered in two or more different groups instead of being grouped together. For example, Wray (group I) and M81E (group VIII) shared Brawley in their parentage but did not cluster together. Similarly, Rex (group II) was in the parentage of Rio (group V) but they did not group together. This separation is most likely due to selections for different traits. A number of studies (Bohn et al. 1999; Tams et al. 2005; Fufa et al. 2005) have demonstrated that genetic relationships based on molecular markers do not always agree with those estimated by pedigree information because of unrealistic assumptions for estimating co-ancestry coefficient (f). Tams et al. (2005) reported a low but significant correlation between f and genetic similarity based on molecular markers (0.32 and 0.33 for SSR and AFLP markers, respectively) in European winter triticale. Fufa et al. (2005) reported a low correlation ($r = 0.28$ for SSR and 0.15 for SRAP markers) while Bohn et al. (1999) reported a moderate correlation ($r = 0.45$) between molecular marker based similarity and f in winter wheat cultivars.

Association of cultivars of unknown parentage with the cultivars of known parentage

Sixty-eight sweet sorghum cultivars and lines were selected without any regard to their origin or their genetic backgrounds. There have been some traditional groupings of sweet sorghums since the beginning of their introduction from African and other countries, (e.g., Amber and Orange group of sorgos) but no previous attempts have been made for systematic classification of the historic and important sweet sorghum cultivars and lines We observed that most of the accessions which were related based on their parentage clustered closely or grouped in the same group as was expected. Twenty cultivars included in the study did not have any genetic background information. These cultivars were distributed in different groups revealing their genetic relationship with other cultivars and lines. For example, Northern Sugar Cane, a variety with high sugar content with unknown origin/parentage, grouped with Early Sumac and N98 with known parentage in group Id. Similarly, Rex, Mennonite, Georgia Blue Ribbon, Simon and Fremont Cane with unknown parentage were grouped in group II with eight other cultivars having Orange in their genetic backgrounds. Roma, a high sugar yielding sorghum variety with unknown parentage, clustered with Ramada, Smith and Bailey (cultivars with known parentage). This clustering analysis based on the segregation at SSR loci can resolve the genetic background issues of the cultivars with unknown pedigree. These relationships of sweet sorghums of unknown genetic origin with cultivars with known parentage will help sorghum breeders to select appropriate parents in their breeding programs to maximize yield as well as to maintain genetic diversity.

Evaluation of sugar content in diverse cultivars/lines

Brix values were estimated for each cultivar and line at maturity (Table 2). The highest brix readings were obtained from cultivars, N. Sugar Cane, Smith, N98, Kansas Collier and Wray with values of 19.29, 18.60, 18.33, 18.20 and 18.00, respectively. The lowest brix readings were from Blacktop, White Orn, Theis, Red Amber (TX) and KS76 with brix values of 7.13, 10.60, 10.93, 11.87 and 11.86, respectively. Out of the possible 25 pairs between these high and the low brix yielding cultivars/lines, only one combination (contrasting pairs), Kansas Collier/White Orn, was diverse with regard to brix content and at DNA level as well with genetic distance (D) value of 0.67. The other pairs (combinations) showed lower genetic distances with D values ranging from 0.29 to 0.56.

Highly significant variation ($P = 0.004$) was observed among the 13 groups (including sub subclusters) for brix content. It could be noted that some varieties with high sugar content, such as, Smith, Ramada and Roma (average sugar content of 17.73) were clustered together in group IX while their diverse counterpart low sugar yielding cultivars, such as, Greenleaf and Honey (average sugar content of 12.43) were clustered in group VII revealing their distant relationship with an average genetic distance of 0.66 (Table 3, Fig. 1). These two groups were found to be significantly different as detected by t tests (LSD) for brix content. Similarly, Kansas Collier, White collier and Brawley were clustered together in group 1b (average sugar content of 17.71, Table 3) were also found to be significantly different from the cultivars white Orn, Top 76-6 and Grassl (average sugar content of 12.47) grouped in group X (Fig. 1) and the contrasting groups also showed a genetic distance of 0.64.

The cultivars and lines which were high sugar yielding and at the same time diverse at SSR loci from their counterpart low sugar yielding cultivars and lines, may possess alleles for increased amount of sugar content in the stem juice. These diverse (contrasting) pairs of cultivars and lines could be used as potential parents for developing mapping populations to map genes that control sugar accumulation in sweet sorghum. It will be relatively easier to develop high density molecular linkage maps and to map genes influencing increased amount of sugar content using segregating populations derived from the crosses between these diverse cultivars/lines. Anderson et al. (1993) identified most diverse genotype pairs based on genetic distance as determined on the basis of PIC values of the RFLP clones and suggested that populations derived from these diverse pairs would map more informative polymorphic markers compared with a population derived from the most polymorphic potential parents based on mere phenotype.

In conclusion, the present study has revealed valuable information on the relationships among a

Table 2 Brix (sugar content) and other agronomic traits values of Sweet sorghum cultivars/lines

Cultivar/line	Days to anthesis	Plant height (cm)	Brix reading	Wet Stalk (g/stalk)	Dry stalk (g/stalk)	Stalk moisture (%)
Dale	100	288	16.07	831.67	241.33	71.04
Tracy	105	266	15.20	715.00	209.33	70.65
Della	78	270	16.00	878.33	292.33	66.70
White African	83	266	14.87	720.00	204.33	71.62
Wray	79	216	18.00	383.33	132.66	64.97
M81E	110	374	12.67	1,173.33	306.00	73.87
Theis	111	348	10.93	1,026.6	263.00	74.46
Brawley	78	240	17.33	518.33	163.33	68.51
Kansas-Collier	75	194	18.20	393.00	134.00	65.92
White-Collier	71	180	17.80	328.33	107.33	67.32
African-Millet	79	180	12.93	493.33	145.66	70.55
Colman	79	242	12.87	528.33	158.00	70.07
H. sugarcane	69	160	15.86	545.00	164.66	70.19
Brandes	113	276	12.97	695.00	197.00	71.55
Bailey	105	290	14.60	544.67	146.33	73.17
Smith	96	248	18.60	990.00	309.33	68.60
Ramada	105	286	17.93	815.00	260.66	67.97
Roma	98	280	16.66	813.33	248.66	69.41
Top 76-6	110	302	12.67	840.00	217.66	74.10
Grassl	110	288	14.13	935.67	276.66	70.18
Keller	99	304	16.13	1,041.6	322.66	68.99
Norkan	68	172	16.30	348.33	108.00	68.53
KS73	71	116	13.60	191.67	63.00	66.27
KS75	70	132	12.13	223.33	75.33	66.30
KS76	68	116	11.86	323.33	110.00	64.08
KS78 (control)	63	138	7.400	265.00	88.66	66.29
Sugar Drip	82	236	16.73	810.00	241.00	70.23
Early Sumac	71	188	15.33	446.67	129.00	71.11
N. Sugar Cane	85	218	19.26	561.67	207.66	60.28
N98	68	104	18.33	316.67	99.00	68.67
N99	62	166	14.73	240.00	76.33	67.38
N100	79	230	17.20	623.33	202.66	67.27
N108	68	122	12.60	440.00	128.00	70.93
N109	68	122	17.00	216.67	73.33	66.03
N110	67	188	16.13	443.33	110.00	73.17
N111	67	184	15.00	458.33	147.00	67.98
Waconia-L	67	180	17.77	408.33	131.00	67.95
Kansas orange	68	192	17.73	498.33	160.33	67.63
Red-X	67	190	15.53	431.67	129.66	70.00
Mennonite	68	184	17.66	335.00	112.33	65.83
Rox orange	67	194	15.00	390.00	121.33	68.84
G. B. Ribbon	71	206	13.30	486.67	148.66	69.36
Fremont cane	59	174	12.07	325.00	92.66	71.53

Table 2 continued

Cultivar/line	Days to anthesis	Plant height (cm)	Brix reading	Wet Stalk (g/stalk)	Dry stalk (g/stalk)	Stalk moisture (%)
Simon	63	198	17.33	415.00	134.66	67.66
Sumac	72	200	17.60	498.00	146.00	70.93
Rex	85	210	17.27	390.00	131.66	66.09
Honey	78	222	12.60	481.67	152.67	68.45
Greenleaf	70	188	12.27	95.00	44.00	52.07
Lahoma	79	218	12.33	386.00	125.00	67.67
Rancher	57	166	13.40	138.67	49.33	64.59
Leoti	63	170	13.67	353.33	119.33	66.19
Red Amber	62	190	11.87	305.00	100.00	67.06
Chinese Amber	68	220	16.93	406.67	133.00	67.09
Dakota Amber	67	192	14.27	255.00	87.33	65.64
Minnesota Amber	68.	174	14.00	305.00	108.66	63.70
Black Ambercane	59	190	13.20	268.33	93.00	64.97
Black-Amber	62	200	12.13	275.00	95.00	65.42
Early Folger	70	216	16.47	363.33	125.33	65.50
Cowley	75	202	16.67	508.33	162.33	68.05
Williams	78	224	15.93	670.00	202.00	69.85
Umbrella	71	226	15.47	596.67	172.66	70.92
Blacktop	57	168	7.13	258.33	75.66	70.62
Snowflake	85	216	17.06	583.33	187.66	67.61
Hastings	77	252	14.40	503.33	155.33	69.14
Iceberg Amber	81	204	15.73	606.67	179.667	70.24
White Orn	85	230	10.60	673.00	190.00	70.89

large number of US sweet sorghum cultivars and lines, especially, for the cultivars of unknown parentage with the cultivars of known parentage. This genetic relationship information will be very helpful in future sweet sorghum breeding programs to improve sugar yield and maintain broad genetic diversity. The study has also identified pairs of cultivars which are diverse for sugar content and for genetic make-up, and these cultivars could be used as potential parents to create mapping populations to map genes influencing high sugar content in sweet sorghum. However, these cultivars should be further

Table 3 Average brix and genetic distance (D) values for diverse groups of cultivars/lines (and particular contrasting pairs of cultivars/lines)

Diverse groups of cultivars /contrasting pairs of cultivars						D value
Cultivar/line	Group	Brix reading	Cultivar/line	Group	Brix reading	
Roma, Smith and Ramada	IX	17.73	Greenleaf and Honey	VI	12.43	0.66
Kansas Collier, White Collier and Brawley	1b	17.77	White Orn, Top 76-6 and Grassl	X	12.47	0.64
Simon and Waconia-L	II	17.55	White Orn, Top 76-6 and Grassl	X	12.47	0.63
Sumac, Rex, Simon and Waconia-L	II	17.54	Greenleaf and Honey	VI	12.43	0.56
Snowflake	IV	17.06	Greenleaf	VII	12.27	0.70
Simon	II	17.33	Theis	VIII	10.93	0.68
Cowley	III	16.67	White Orn	X	10.60	0.63

evaluated for their sugar content to confirm their value as parents for tagging such genes.

Acknowledgements This research was supported by USDA-CSREES-NRI Grant 2004-35300-14700. Authors extend special thanks to Dr. J. F. Pederson and Mr. J. J. Toy, USDA-ARS, University of Nebraska-Lincoln, Nebraska; National Center for Genetic Resources Preservation (NCGRP), 111 South Mason, Fort Collins, Colorado; National Plant Germplasm System, Griffin, Georgia; Dr. W. L. Rooney, Texas Agricultural System Station, College station, Texas and Dr. M. Bitzer, University of Kentucky, Kentucky for providing seeds of sweet sorghum cultivars and lines.

References

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

Ahmad M. (2002) Assessment of genomic diversity among wheat genotypes as determined by simple sequence repeats. Genome 45:646–651

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

Anderson JA, Churchill GA, Autrique JE, Sorrellls ME, Tanksley SD (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

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

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

Budak H, Pedraza F, Cregan PB, Baenziger PS, Dweikat I (2003) Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearlmillet germplasm. Crop Sci 43:2284–2290

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

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

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich (1999) Genetic redundancy and diversity among Orange accessions in the US national sorghum collection as assessed with simple sequence (SSR) markers. Crop Sci 39:1215–1221

Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26:297–302

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

Gnansounou E, Dauriat A, Wyman C (2005) Refining sweet sorghum to ethanol and sugar: economic trade-offs in the context of North China. Bioresoure Tech 96:985–1002

Huang XQ, Börner A, Röder MS, Ganal MW (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor Appl Genet 105:699–707

Innan H, Terauchi R, Miyashita T (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana*. Genetics 146:1441–1452

Kuleung C, Baenziger PS, Dweikat I (2004) Transferability of SSR markers among wheat, rye and triticale. Theor Appl Genet 108:1147–1150

Kuleung C, Baenziger PS, Kachman SD, Dweikat I (2006) Evaluating the genetic diversity of triticale with wheat and rye SSR markers. Crop Sci 46:1692–1700

Mahmood A, Baeeziger PS, Budak H, Gill KS, Dweikat I (2004) The use of microsatellite markers for the detection of genetic similarity among winter bread wheat lines for chromosome 3A. Theor Appl Genet 109:1494–1503

Melchinger AE (1999) Genetic diversity and heterosis. In: The genetics and exploitation of heterosis in crops. American Society of Agronomy, Madison, Wisc., pp 99–118

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

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl acad Sci USA 76:5269–5273

Parker GD, Fox PN, Langridge P, Chalmers K, Whan B, Ganter PF (2002) Genetic diversity within Australian wheat breeding programs based on molecular and pedigree data. Euphytica 124:293–306

Rohlf FJ (2000) NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket

SAS Institure (1999–2001) The SAS system for windows V8. Release 8.2 (TS2MO)

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 Genet 105:912–920

Senior ML, Murphy JP, Goodman NM, Stuber CW (1998) Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci 38:1088–1098

Smith CW, Frederiksen RA (2000) Sorghum: origin, history, technology and production. Wiley, New York. Preface, pp vii–ix

Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S, Ziegle J (1997) An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. Theor Appl Genet 95:163–173

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

Somers DJ, Banks T, Depauw R, Clarke J, Fox S, Pozniak C (2007) Genome-wide linkage disequilibrium in wheat. Plant and Animal Genome XV conference, 13–17 Jan., 2007, San Diego, CA, Poster abstract P268

Souza E, Sorrells ME (1989) Pedigree analysis of North American Oat cultivars released from 1951 to 1985. Crop Sci 29:595–601

Tams SH, Melchinger AE, Bauer E (2005) Genetic similarity among European winter triticale elite germplasms assessed with AFLP and comparisons with SSR and pedigree data. Plant Breed 124:154–160

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

Weber JL (1990) Informativeness of human (dC-dA)n (dG-dT)n poylmorphisms. Genomics 7:524–530

Weir BS (1996) Genetic data analysis: II. Methods for the discrete population genetic data, 2nd edn. Sinnauer Associates, Sunderland