Genetic diversity of sweet sorghum germplasm in Mexico using AFLP and SSR markers

Víctor Pecina-Quintero[1], José Luis Anaya-López[1], Alfredo Zamarripa-Colmenero[2], Noe Montes-García[3], Carlos Nuñez-Colín[1], Jose Luis Solis-Bonilla[2], María Rocío Aguilar-Rangel[1] and Louis Prom[4]

[1]Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP), Campo Experimental Bajío, Km 6.5, Carretera Celaya-San Miguel de Allende, Apartado Postal 112, CP 38110 Celaya, Guanajuato, Mexico. E-mail: pecina.victor@inifap.gob.mx, anaya.jose@inifap.gob.mx, nunez.carlos@inifap.gob.mx, rociobacks@hotmail.com [2]INIFAP, Campo Experimental Rosario Izapa, Tuxtla Chico, Chiapas, Mexico. E-mail: zamarripa.alfredo@inifap.gob.mx, solis.joseluis@inifap.gob.mx [3]INIFAP, Campo Experimental Río Bravo, Río Bravo, Tamaulipas, Mexico. E-mail: montes.noe@inifap.gob.mx [4]United States Department of Agriculture, Agricultural Research Service, 2765 F&B Road, College Station, TX, USA. E-mail: louis.prom@ars.usda.gov

Abstract – The objective of this work was to evaluate the diversity and genetic relationships between lines and varieties of the sweet sorghum (*Sorghum bicolor*) germplasm bank of the National Institute for Forestry, Agriculture and Livestock Research, Mexico, using AFLP and SSR markers. The molecular markers revealed robust amplification profiles and were able to differentiate the 41 genotypes of sweet sorghum evaluated. Analysis of the frequency and distribution of polymorphic fragments allowed for the detection of unique (AFLP) and rare (SSR) alleles in several genotypes (RBSS-8, RBSS-9, RBSS-25, RBSS-32, and RBSS-37), indicating that these markers may be associated with a feature that has not yet been determined or may be useful for the identification of these genotypes. The genetic relationships indicated the presence of at least two types of sweet sorghum: a group of modern genotypes used for sugar and biofuel production, and another group consisting of historic and modern genotypes used for the production of syrups. Sweet sorghum genotypes may be used to develop new varieties with higher sugar and juice contents.

Index terms: *Sorghum bicolor*, biofuels, genetic improvement, molecular marker, rare allele.

Diversidade genética em germoplasma de sorgo sacarino no México com uso de marcadores AFLP e SSR

Resumo – O objetivo deste trabalho foi avaliar a diversidade e as relações genéticas entre linhagens e variedades de sorgo sacarino (*Sorghum bicolor*) do banco de germoplasma do Instituto Nacional de Investigações Florestais, Agrícolas e Pecuárias, México, por meio de marcadores AFLP e SSR. Os marcadores moleculares revelaram robustez dos perfis de amplificação e permitiram diferenciar os 41 genótipos de sorgo sacarino avaliados. A análise da frequência e da distribuição dos fragmentos polimórficos permitiu detectar alelos únicos (AFLP) e raros (SSR) em diversos genótipos (RBSS-8, RBSS-9, RBSS-25, RBSS-32 e RBSS-37), o que indica que estes marcadores podem estar associados a uma característica que ainda não foi determinada ou podem ser úteis para a identificação destes genótipos. As relações genéticas indicaram a presença de pelo menos dois tipos de sorgo sacarino: um grupo de genótipos modernos usados para a produção de açúcar e biocombustíveis, e outro grupo formado por genótipos antigos e modernos utilizados para a produção de xaropes. Genótipos de sorgo sacarino podem ser usados para o desenvolvimento de novas variedades com maior conteúdo de açúcar e caldo.

Termos para indexação: *Sorghum bicolor*, biocombustíveis, melhoramento genético, marcador molecular, alelos raros.

Introduction

Sweet sorghum [*Sorghum bicolor* (L.) Moench] belongs to the same species as grain sorghum and forage sorghum; however, it has been agriculturally selected to accumulate high levels of sucrose in the stem parenchyma. The juice extracted from sweet sorghum cane contains high levels of sucrose and invert sugar that are easily fermented to produce ethanol (Prasad et al., 2007). This juice has been used to produce syrup in the USA and alcohol in Brazil and India. It is estimated that, under favorable conditions, sweet sorghum can produce around 43 Mg ha^{-1} per year of juice, which contains 11.8% of fermentable sugars (Kim & Day, 2011). Some varieties have been reported to produce sugar yields similar to those of sugarcane (Ratnavathi et al., 2010). Additionally, sweet sorghum is highly efficient in water use, even in areas where there are frequent periods of drought and high temperatures. The cultivation costs are also three times lower than those of sugarcane (Reddy et al., 2005), making it an ideal raw material for bioenergy.

Several sweet sorghum varieties from the USA have been introduced into Mexico, and new genotypes have been developed from these varieties. However, according to Murray et al. (2009), only six African genotypes were used in the development of many of the new varieties, indicating that sweet sorghum cultivars in the USA may have a narrow genetic base. If the genetic base is too narrow, the development of new genotypes for bioenergy production will be difficult. Therefore, it is necessary to determine the genetic diversity of this crop and to identify groups of similar genotypes for the purposes of conservation and proper use of genetic resources, in addition to the protection of property rights.

Genetic diversity has been estimated using DNA molecular markers, and the quality of this estimation depends on the number of markers generated for each species and on their coverage of the genome. In sweet sorghum, genetic diversity has been determined using amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) markers. Due to their high values for polymorphic information content (PIC) and the Shannon diversity index (Geleta et al., 2006), these marker types are adequate for use in this species.

Although some analyses of the sweet sorghum germplasm in the USA have indicated that sweet sorghum genotypes cannot be separated from the grain sorghum lines because of their similar racial origin within *S. bicolor* ssp. *bicolor* (Ritter et al., 2007; Ali et al., 2008). A more recent study using SSR and single nucleotide polymorphism (SNP) markers showed that sweet sorghum accessions can be classified into three main groups: historical and modern genotypes used for syrup production, modern genotypes used for sugar or energy production, and amber genotypes (Murray et al., 2009). In Mexico, the National Institute for Forestry, Agriculture and Livestock Research (Inifap) has introduced lines and varieties from the USA and the International Crops Research Institute for the Semi-Arid Tropics (Icrisat); however, the level of genetic diversity required to develop new genotypes is unknown.

Regarding SSR markers, different studies have recommended the use of these markers in analyses of genetic diversity due to their high degree of polymorphism (Geleta et al., 2006; Ali et al., 2008; Shehzad et al., 2009).

The objective of this work was to evaluate the diversity and genetic relationships between lines and varieties of the sweet sorghum germplasm bank of Inifap using AFLP and SSR markers.

Materials and Methods

A representative set of 41 lines and cultivars of sweet sorghum (Table 1) from Inifap was used for diversity analyses. Total genomic DNA was extracted from the young leaves of five plants of each line following the standard cetyltrimethylammonium bromide method (CTAB) with minor modifications (Doyle & Doyle, 1990). Leaf tissue (2 g) was ground to a fine powder in liquid nitrogen, and 100 mg of this powder were homogenized in 800 µL of extraction buffer [100 mmol L^{-1} Tris, pH 8.0; 150 mmol L^{-1} EDTA, pH 8.0; 2.1 mol L^{-1} NaCl; 3% (w v^{-1}) CTAB; 1% (v v^{-1}) β-mercaptoethanol].

AFLP analysis was performed as described by Vos et al. (1995). Genomic DNA (10 ng $µL^{-1}$) was digested with 5 U each of EcoRI and MseI at 37°C for 3 hours. Digested samples were incubated at 70°C for 15 min to deactivate the restriction enzymes. EcoRI and MseI adapters (5 and 50 pmol, respectively) were ligated to the digested DNA fragments using 1 U of T4 DNA ligase in ligation buffer (1X T4 DNA ligase buffer), and incubated at 37°C overnight. Taq DNA polymerase was used in all polymerase chain reactions (PCR). Selective amplification was performed using six different AFLP primer combinations (Table 2). Pre-amplification of ligated DNA (diluted 10-fold) was carried out with

Table 1. Genotypes of sweet sorghum (*Sorghum bicolor*) used in the study.

Number	Accession	Origen	Number	Accession	Origen
1	RBSS-1	ICSSH-21	22	RBSS-22	SP-4504
2	RBSS-2	ICSSH-22	23	RBSS-23	S-23
3	RBSS-3	ICSSH-23	24	RBSS-24	SP-4495
4	RBSS-4	ICSSH-24	25	RBSS-25	RP-21607
5	RBSS-5	ICSSH-25	26	RBSS-26	RP-21613
6	RBSS-6	ICSSH-26	27	RBSS-27	RP-21616
7	RBSS-7	ICSSH-38	28	RBSS-28	RP-21620
8	RBSS-8	ICSSH-324	29	RBSS-29	RP-21628
9	RBSS-9	ICSSH-474	30	RBSS-30	RP-21632
10	RBSS-10	ICSSH-675	31	RBSS-31	Fortuna
11	RBSS-11	ICSSH-749	32	RBSS-32	Dale
12	RBSS-12	NTJ-2	33	RBSS-33	RP-21634
13	RBSS-13	SPV-422	34	RBSS-34	URJA
14	RBSS-14	ICSV-700	35	RBSS-35	Topper
15	RBSS-15	E-36	36	RBSS-36	Sureño
16	RBSS-16	Mazatlán	37	RBSS-37	Theis
17	RBSS-17	SSV-74	38	RBSS-38	Maravilla
18	RBSS-18	SP-4481	39	RBSS-39	RB-Cañero
19	RBSS-19	SP-4511	40	RBSS-40	INIFAP16
20	RBSS-20	SS-2016	41	RBSS-41	INIFAP23
21	RBSS-21	SP-4487	-	-	-

primers complementary to the EcoRI and MseI adapters with one selective nucleotide (adenine and cytosine, respectively) in a Px2 thermal cycler (Thermo Electron Corporation, Milford, MA, USA), using the following cycling parameters: 20 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 60 s. The second amplification was carried out with six selective primer combinations of EcoRI (700 and 800 nm) and MseI with three selective nucleotides (Table 2).

Each PCR product was electrophoresed independently on 6% denaturing polyacrylamide gels. The PCR products were fractionated on a sequencing system (Li-Cor IR2, Li-Cor, Lincoln, NE, USA) equipped with two infrared lasers with the ability to read at 700 and 800 nm wavelengths. Only bright, clearly distinguishable bands between 50 and 700 bp were recorded for analysis. Negative controls containing no template were performed for each PCR.

The SSR primers used in the present study were developed for *S. bicolor* (Taramino et al., 1997; Bhattramakki et al., 2000; Kong et al., 2000) and were selected based on their clear polymorphic patterns and on their position in the sorghum genome, covering ten linkage groups or chromosomes. In total, 29 SSR primers were used (Table 3). The microsatellite PCR contained 50 ng μL^{-1} template DNA, 1X PCR buffer, 0.2 mmol L^{-1} each dNTP, 4 mmol L^{-1} $MgCl_2$, 1 μmol L^{-1} each primer (F and R), and 1 unit of Taq DNA polymerase in a total reaction volume of 25 μL. Amplification was performed in a Px2 thermal cycler (Thermo Electron Corporation, Milford, MA, USA) using a program with

an initial denaturation at 94°C for 60 s followed by 35 cycles at 94°C for 60 s, 57°C for 60 s, and 72°C for 60 s with a final hold at 72°C for 10 min. Formamide dye was added to the appropriate PCR products, which were then subjected to electrophoretic separation on 8.5% denaturing polyacrylamide gels.

To determine which of the AFLP primer combinations were the most informative, the following parameters were calculated: polymorphism information content (PIC), marker index (MI), resolving power (RP), and diversity index (DI). The PIC value for each AFLP primer combination was calculated according to the formula $PICi = 2fi (1 - f_i)$, in which f_i is the frequency of the marker fragments that were present, and $1 - f_i$ is the frequency of the marker fragments that were absent. PIC was averaged over all fragments for each primer combination. PIC for each SSR locus was calculated using a freely available online calculator (Kemp, 2012).

The marker index (MI) was calculated according to the formula $MI = PIC \times EMR$, in which EMR is the effective multiplex ratio, defined as the product of the total number of loci per fragments per primer (n) and the fraction of polymorphic loci fragments (β) ($EMR = n\beta$).

The resolving power (RP) of each primer was calculated as $RP = \Sigma I_b$, in which I_b represents fragment informativeness. The I_b can be represented on a 0–1 scale using the formula $I_b = 1 - [2 \times (0.5 - p)]$, in which p is the proportion of the 88 accessions that contain the fragment.

Table 2. Polymorphism information, genetic diversity index (DI), marker index (MI), effective multiple ratio (EMR), polymorphic information content (PIC), and resolving power (RP) given for each of the amplified fragment length polymorphism (AFLP) combinations used in the study.

Primer combination	Number of fragments					DI	MI	EMR	PIC	RP
	Amplified	Polymorphic	Unique	Rare	Similar					
E-AAG + M-CAC	108	40	2	7	81	0.671	11.077	40	0.277	15.610
E-AGA + M-CAC	122	60	8	6	74	0.712	18.196	60	0.303	28.244
E-AAC + M-CAA	140	64	4	6	98	0.633	18.706	64	0.292	27.854
E-AGA + M-CAA	108	54	7	4	64	0.723	16.952	54	0.314	26.683
E-ACA + M-CAT	120	48	5	5	88	0.668	13.459	48	0.28	19.707
E-AGG + M-CAT	124	47	2	6	91	0.679	14.506	47	0.309	21.756
E-ACA + M-CTG	111	54	5	3	73	0.679	14.682	54	0.272	21.366
E-AGG + M-CTG	121	55	4	8	82	0.675	15.455	55	0.281	22.537
E-AAG + M-CAG	137	58	1	9	91	0.738	18.211	58	0.314	26.780
E-AGG + M-CAG	128	60	1	11	82	0.722	17.989	60	0.300	26.537
Total	1,219	540	39	65	824	6.900	159.230	540	2.942	237.074
Average	121.9	54	3.9	6.5	82.4	0.690	15.920	54	0.294	23.707

The diversity index, which indicates the genetic diversity of the germplasm, was calculated using the formula $DI = 1 - \Sigma Pi^2$, in which Pi is the allele frequency i^n (each individual allele is considered a unique fragment amplification).

Each AFLP primer combination received a score (1, for presence, and 0, for absence of bands in each accession), and a binary matrix was generated. A dendrogram was constructed using the Dice similarity coefficient (DSC) (Nei & Li, 1979), the unweighted pair group method with arithmetic average (UPGMA), and the jackknife method for corroboration. A bootstrap resampling method was performed to determine the robustness of the dendrogram, and 1,000 bootstrap replicates were obtained from the original data of 88 accessions. From these 1,000 matrices, confidence

Table 3. Number of alleles (NA), expected number of alleles (NAE), Shannon index (SI), marker index (MI), polymorphic information content (PIC), resolving power (RP), and genetic diversity index (DI) calculated for each simple sequence repeat (SSR) locus.

Locus	NA	NAE	SI	MI	PIC	RP	DI
mSbCIR286	5.00	3.99	1.49	0.401	0.713	2.88	0.810
Xcup01	8.00	2.84	1.46	0.232	0.621	2.44	0.942
Xtxp32	5.00	3.58	1.36	0.434	0.672	2.63	0.833
Msbcir238	6.00	3.31	1.34	0.249	0.643	2.05	0.816
Xcup40	4.00	2.61	1.10	0.309	0.559	2.00	0.691
Xgap84	7.00	5.04	1.73	0.345	0.773	3.66	0.887
Xgap236	5.00	3.39	1.37	0.329	0.663	2.39	0.764
Xtxp285	7.00	3.59	1.49	0.226	0.662	2.24	0.953
mSbCIR323	5.00	2.07	1.05	0.213	0.489	1.32	0.775
Xtxp12	6.00	4.27	1.58	0.353	0.732	3.07	0.843
Xtxp21	3.00	2.76	1.05	0.367	0.560	1.56	0.517
Xtxp15	7.00	5.61	1.78	0.350	0.796	3.71	0.910
Xtxp23	4.00	2.54	1.05	0.219	0.528	1.12	0.616
Xtxp057	4.00	3.86	1.36	0.479	0.693	3.27	0.746
Xtxp17	5.00	3.37	1.33	0.326	0.650	2.24	0.763
Xtxp265	5.00	3.97	1.48	0.358	0.709	2.59	0.920
Xtxp6	7.00	5.14	1.72	0.349	0.777	3.95	0.889
Xcup52	3.00	2.41	0.95	0.228	0.499	0.88	0.448
Xcup47	4.00	3.14	1.23	0.363	0.621	1.95	0.681
Xtxp18	6.00	3.73	1.49	0.311	0.690	2.49	0.826
Xtxp273	6.00	2.98	1.27	0.220	0.606	1.76	0.777
Xtxp321	7.00	4.56	1.67	0.319	0.749	3.27	0.877
Xcup2	3.00	1.31	0.45	0.159	0.218	0.54	0.706
Xtxp289	7.00	3.94	1.57	0.246	0.709	2.39	0.952
Xtxp67	6.00	4.34	1.58	0.360	0.733	3.41	0.847
Xcup07	2.00	1.76	0.62	0.433	0.339	1.27	0.717
Xcup67	6.00	4.53	1.63	0.373	0.748	3.27	0.853
Sb03g027710	5.00	3.17	1.35	0.326	0.644	2.20	0.899
Sb03g039090	6.00	4.38	1.59	0.257	0.738	2.00	0.962
Average	5.31	3.53	1.36	0.314	0.639	2.36	0.800
Standard deviation	1.49	1.02	0.32	-	-	-	-

limits for each pair-wise comparison were determined. All calculations were performed using FreeTree 0.9.1.50 (Pavlícek et al., 1999), and the dendrogram was drawn using TreeView 1.6.6. (Page, 2001). Moreover, a principal coordinate analysis (PCooA), using the DSC (Nei & Li, 1979), was conducted with NTSYS PC version 2.2 (Rohlf, 2000). This analysis was performed to generate a three-dimensional representation of the accessions to confirm the similarity among them.

Results and Discussion

Robust amplification profiles were obtained from AFLP and SSR markers, which were able to differentiate between genotypes. The polymorphism rate of the AFLP markers was 54%, which was similar to that observed by Ritter et al. (2007); however, other studies have reported polymorphism rates between 85 and 93% using AFLP and RAPD markers (Geleta et al., 2006; Kachapur et al., 2009). In the present study, AFLP markers were polymorphic enough to distinguish between sweet sorghum genotypes and to detect high levels (69%) of genetic diversity (Table 2), which is consistent with the observations of Geleta et al. (2006), Ritter et al. (2007), Ali et al. (2008), and Kachapur et al. (2009).

The analysis of the frequency and distribution of polymorphic AFLP fragments showed 39 specific fragments – each present in a single individual – and 65 rare fragments present in up to 10% of the genotypes (Table 2). It is probable that the presence of such fragments in a gene is related to a loss of genetic diversity, possibly due to selection (Foster et al., 2010). However, specific and rare alleles are of great interest because they may be linked to a particular genotype and may serve to distinguish a genotype or a specific region of the genome (Agrama & Tuinstra, 2003). The genotypes with the highest number of unique and rare AFLP fragments were RBSS-8, RBSS-9, RBSS-25, RBSS-32 (Dale), RBSS-37 (Theis); three of them are genotypes introduced by Icrisat, and two are improved USA varieties, confirming the data on rare and unique markers (Agrama & Tuinstra, 2003; Foster et al., 2010).

However, five AFLP combinations were highly informative (E-AGA + M-CAC, E-AGG + M-CAT, E-AAG + M-CAG, E-AGA + M-CAA, and E-AGG + M-CAG) with an average PIC of 0.3 (Table 2).

Therefore, these AFLP combinations are recommended for use in germplasm diversity analysis of *S. bicolor* and other species. In addition, the correlation of PIC with MI and RP (0.83 and 0.99, respectively) confirms the usefulness of these AFLP combinations (Tatikonda et al., 2009).

The total number of alleles and size of SSR markers varied from the ones originally reported, indicating that the source of the material tested differed from that of the previously characterized collections (Ali et al., 2008; Shehzad et al., 2009). An average of five alleles per locus and a polymorphism rate of 98% were observed (Table 3); the latter being similar to that found by Agrama & Tuinstra (2003) and Pei et al. (2010), but higher than the ones observed by Ali et al. (2008) and Shehzad et al. (2009). Moreover, 11 tested loci (mSbCIR286, Xgap84, Xtxp12, Xtxp15, Xtxp265, Xtxp6, Xtxp321, Xtxp289, Xtxp67, Xcup67, and Sb03g039090) had PIC values greater than 0.7, which

is consistent with the results found for some of these loci by Shehzad et al. (2009) and Pei et al. (2010). The high average genetic diversity observed in the present study (0.79) was similar to that found by Smith et al. (2000) and Agrama & Tuinstra (2003), who obtained values of 0.62 and 0.58, respectively, but was higher than the values reported by Schloss et al. (2002), Ali et al. (2008), and Shehzad et al. (2009), who obtained values of 0.46, 0.40, and 0.217, respectively. Rare alleles were also observed in 65% of the evaluated loci (Table 4). Rare alleles have a frequency less than 0.05 (Casa et al., 2005; Somers et al., 2007), and their importance lies in the fact that they are unique and associated with a particular genotype (Agrama & Tuinstra, 2003). In the present study, the genotypes with the highest number of rare SSR alleles also contained unique AFLP alleles.

The Xtxp265 marker was highly polymorphic and had a high PIC value of 0.7 (Table 3). Murray et al.

**Table 4.** Number and frequency of simple sequence repeat (SSR) alleles observed in sweet sorghum.

Locus	Alleles per locus							
	A	B	C	D	E	F	G	H
mSbCIR286	0.1585	0.1220	0.3902	0.1220	0.2073			
Xcup01	0.5610	0.1098	0.0610	0.0488	0.0224	0.0732	0.1098	0.0122
Xtxp32	0.1585	0.3780	0.2927	0.1585	0.0122			
Msbcir238	0.1951	0.0488	0.3902	0.3293	0.0244	0.0122		
Xcup40	0.5366	0.0366	0.1829	0.2439				
Xgap84	0.1707	0.2927	0.1585	0.0854	0.2195	0.0488	0.0244	
Xgap236	0.4512	0.1585	0.1463	0.2073	0.0366			
Xtxp285	0.1951	0.4390	0.1829	0.0366	0.1098			
mSbCIR323	0.6707	0.0244	0.0976	0.0976	0.1098			
Xtxp12	0.0244	0.0854	0.3659	0.1951	0.1585	0.1707		
Xtxp21	0.4146	0.3902	0.1951					
Xtxp15	0.0122	0.1463	0.2439	0.1707	0.1707	0.1829	0.0732	
Xtxp23	0.4878	0.0244	0.3780	0.1098				
Xtxp057	0.3171	0.1951	0.2195	0.2683				
Xtxp17	0.3780	0.1707	0.3415	0.0854	0.0244			
Xtxp265	0.0732	0.1829	0.2195	0.3780	0.1463			
Xtxp6	0.0244	0.0244	0.1220	0.2195	0.2561	0.2073	0.1463	
Xcup52	0.3902	0.1098	0.5000					
Xcup47	0.3659	0.1585	0.3902	0.0854				
Xtxp18	0.0366	0.3415	0.1098	0.0732	0.3537	0.0854		
Xtxp273	0.4512	0.0244	0.0244	0.1463	0.3293	0.0244		
Xtxp321	0.1707	0.2927	0.0366	0.1098	0.2927	0.0610	0.0366	
Xcup2	0.8659	0.1098	0.0244					
Xtxp289	0.0488	0.2317	0.3902	0.1951	0.0732	0.0366	0.0244	
Xtxp67	0.1098	0.2683	0.2195	0.3049	0.0610	0.0366		
Xcup07	0.3171	0.6829						
Xcup67	0.1098	0.3415	0.0488	0.1098	0.2195	0.1707		
Sb03g027710	0.0854	0.0732	0.2073	0.4878	0.1463			
Sb03g039090	0.3415	0.2439	0.1463	0.1220	0.1220	0.0244		

(2009) reported that this marker had a significant association with plant height and flowering time. Xtxp265 is located on chromosome 6, within the gene more closely linked with photoperiod sensitivity (Murray et al., 2008). Similarly, Ji et al. (2011) found a positive association between the Xtxp18 marker and an allele of 270 bp located on chromosome SB1-08. Polymorphisms were also observed using the marker Sb03g027710, which corresponds to an SNP marker derived from sugarcane (Calvino et al., 2009). These authors recommend the use of this marker along with five additional SNP markers for the genetic mapping of genes linked to sugar content or for the selection of parents in sweet sorghum. Therefore, this marker may be useful for the genetic improvement of this species in the Mexican germplasm bank.

The diversity indices for AFLP and SSR markers were 69 and 79%, respectively. Although a more accurate estimate of genetic relationships would theoretically be obtained using a greater number of markers, the distribution of these markers in the genome is equally important. Menz et al. (2002) indicated that markers with a low level of genomic coverage may provide a different classification for a particular gene. For example, many of the AFLP markers are generated using the enzymes EcoRI and MseI, and the recognition sites for these enzymes are more densely located in the centromeric regions of chromosomes; therefore, centromeric regions have a high weight in the classification of germplasm. With the creation of genetic maps for many crop species, markers can be selected to adequately cover the entire genome without any particular region being over- or underrepresented. Consequently, a more accurate estimate of the dissimilarity between genotypes can be achieved by selecting markers that are uniformly distributed throughout the genome. In the present study, 29 SSR markers distributed throughout the ten chromosomes of sorghum were selected, which provided a better estimate of genetic relationships among the studied genotypes compared to the AFLP markers. These genetic relationships (Figure 1) indicated the presence of at least two types of sorghum. The first group, cluster 1, was divided into two subgroups (1A and 1B). Cluster 1A contained most of the genotypes, including the grain sorghums Sureño and Mazatlán, and both subgroups of this cluster contain historical and modern sorghum lines used for syrup production (Murray et al.,

2009). These lines have tall stems that are large in diameter and very juicy, with high brix content, which is lower than that of sorghum lines developed for sugar and energy production (bicolor type). The second group, cluster 2, includes four modern genotypes of sweet sorghum used for bioenergy production, i.e., Dale, Theis, Maravilla, and RB-Cañero (Murray et al., 2009); the latter is derived from the variety M81E (kafir/bicolor type).

Deu et al. (2006) found that the bicolor race has high genetic diversity and many rare alleles, which is not surprising considering that this race is considered to be the oldest and most widely distributed geographically due to its several uses (fodder, brooms, and sweet stems). Additionally, genotypes RBSS-7,
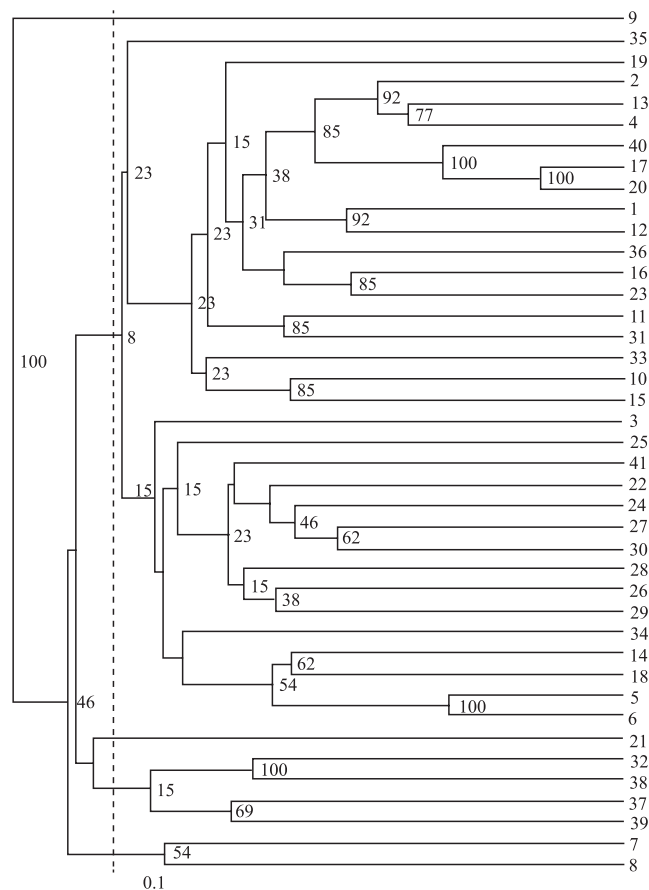
Figure 1. Dendrogram of the 41 evaluated genotypes of sweet sorghum (*Sorghum bicolor*) obtained from amplified fragment length polymorphism (AFLP) and simple sequence repeat (SSR) molecular markers, using the similarity index of Nei & Li and the unweighted pair group method with arithmetic mean (UPGMA).

RBSS-8, RBSS-9, and RBSS-21 were very divergent. Casa et al. (2008) observed that the genotypes of amber varieties of sweet sorghum are very similar to those of the bicolor type, although amber varieties are more divergent. Therefore, it is likely that the four genotypes previously mentioned correspond to the amber variety. Furthermore, bootstrap confidence intervals indicating the similarity of the genotypes showed that, at multiple nodes, the association was not as robust because the values were less than 30%.

To validate these results, PCooA was performed, which determined that the first three principal coordinates accounted for 26% of total variation, which agreed with the bootstrap values. A three-dimensional representation (Figure 2) showed that the dispersion of the genotypes of the detected clusters corresponded to the one observed in the dendrogram, and also illustrated that the genotypes of cluster 1B were more heterogeneous. Moreover, outlier genotypes (RBSS-7, RBSS-8, RBSS-9, and RBSS-21) could clearly be distinguished.
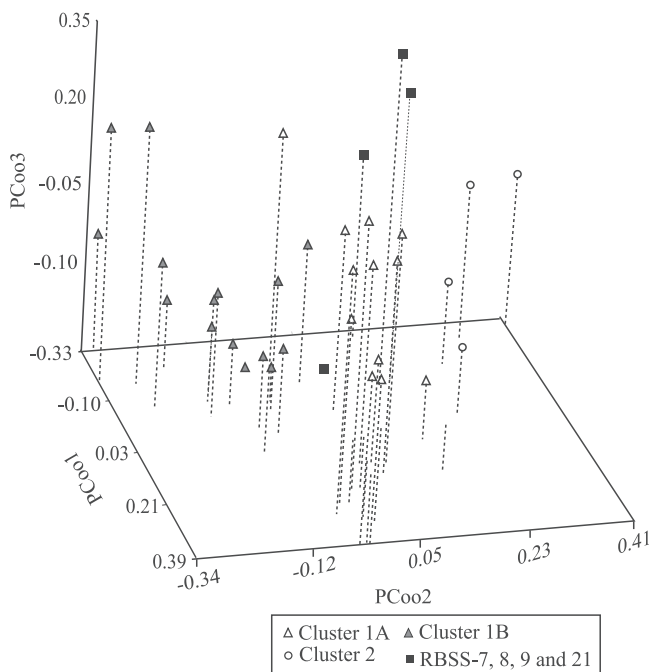
Figure 2. Dispersion of sweet sorghum (*Sorghum bicolor*) germplasm from the National Institute for Forestry, Agriculture and Livestock Research, based on the first three principal coordinates.

Conclusions

1. The genotypes RBSS-8, RBSS-9, RBSS-25, RBSS-32 (Dale), and RBSS-37 (Theis) show unique (AFLP) and rare (SSR) alleles, indicating that the AFLP and SSR markers in these genotypes may be associated with a particular feature that has not yet been determined.

2. In general, AFLP and SSR markers show a high level of genetic diversity.

3. The Inifap germplasm bank contains a great diversity of genotypes, which may be used in breeding programs for the development of new varieties of sweet sorghum with high sugar and juice content.

Acknowledgements

References

AGRAMA, H.A.; TUINSTRA, M.R. Phylogenetic diversity and relationships among sorghum accessions using SSRs and RAPDs. **African Journal of Biotechnology**, v.2, p.334-340, 2003.

ALI, M.L.; RAJEWSKI, J.F.; BAENZIGER, P.S.; GILL, K.S.; ESKRIDGE, K.M.; DWEIKAT, I. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. **Molecular Breeding**, v.21, p.497-509, 2008.

BHATTRAMAKKI, D.; DONG, J.M.; CHHABRA, A.K.; HART, G.E. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. **Genome**, v.43, p.988-1002, 2000.

CALVIÑO, M.; MICLAUS, M.; BRUGGMANN, R.; MESSING, J. Molecular markers for sweet sorghum based on microarray expression data. **Rice**, v.2, p.129-142, 2009.

CASA, A.M.; MITCHELL, S.E.; HAMBLIN, M.T.; SUN, H.; BOWERS, J.E.; PATERSON, A.H.; AQUADRO, C.F.; KRESOVICH, S. Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. **Theoretical and Applied Genetics**, v.111, p.23-30, 2005.

CASA, A.M.; PRESSOIR, G.; BROWN, P.J.; MITCHELL, S.E.; ROONEY, W.L.; TUINSTRA, M.R.; FRANKS, C.D.; KRESOVICH, S. Community resources and strategies for association mapping in sorghum. **Crop Science**, v.48, p.30-40, 2008.

DEU, M.; RATTUNDE, F.; CHANTEREAU, J. A global view of genetic diversity in cultivated sorghums using a core collection. **Genome**, v.49, p.168-180, 2006.

DOYLE, J.J.; DOYLE, J.L. Isolation of plant DNA from fresh tissue. **Focus**, v.12, p.13-15, 1990.

FOSTER, J.T.; ALLAN, G.J.; CHAN, A.P.; RABINOWICS, P.D.; RAVEL, J.; JACKSON, P.J.; KEIM, P. Single nucleotide polymorphism for assessing genetic diversity in castor bean (*Ricinus communis*). **BMC Plant Biology**, v.10, p.1-11, 2010.

GELETA, N.; LABUSCHAGNE, M.T.; VILJOEN, C.D. Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho-agronomical markers. **Biodiversity and Conservation**, v.15, p.3251-3265, 2006.

JI, G.S.; SONG, Y.F.; LIU, G.Q.; DU, R.H.; HAO, F.W. Genetic analysis of sorghum resources from China using SSRs. **Journal of SAT Agricultural Research**, v.9, p.1-7, 2011.

KACHAPUR, R.M.; SALIMATH P.M.; REDDY, B.V.S. Genetic diversity within sweet sorghum (*Sorghum bicolor* (L.) Monech) accessions as revealed by RAPD markers. **Journal of Maharashtra Agricultural Universities**, v.34, p.38-42, 2009.

KEMP, S. **PIC calculator**. Available at: <http://www.liv.ac.uk/~kempsj/pic.html>. Accessed on: 6 Aug. 2012.

KIM, M.; DAY, D.F. Composition of sugar cane, energy cane, and sweet sorghum suitable for ethanol production at Louisiana sugar mills. **Journal of Industrial Microbiology and Biotechnology**, v.38, p.803-807, 2011.

KONG, L.; DONG, J.; HART, G.E. Characteristics, linkage-map positions, and allelic differentiation of *Sorghum bicolor* (L.) Moench DNA simple-sequence repeats (SSRs). **Theoretical and Applied Genetics**, v.101, p.438-448, 2000.

MENZ, M.A.; KLEIN, R.R.; MULLET, J.E.; OBERT, J.A.; UNRUH, N.C.; KLEIN, P.E. A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. **Plant Molecular Biology**, v.48, p.483-499, 2002.

MURRAY, S.C.; ROONEY, W.L.; HAMBLIN, M.T.; MITCHELL, S.E.; KRESOVICH, S. Sweet sorghum genetic diversity and association mapping for brix and height. **The Plant Genome**, v.2, p.48-62, 2009.

MURRAY, S.C.; SHARMA, A.; ROONEY, W.L.; KLEIN P.E.; MULLET, J.E.; MITCHELL, S.E.; KRESOVICH, S. Genetic improvement of sorghum as a biofuel feedstock: I. QTL for stem sugar and grain nonstructural carbohydrates. **Crop Science**, v.48, p.2165-2179, 2008.

NEI, M.; LI, W.H. Mathematical model for studying genetic variation in terms of restriction endonucleases. **Proceedings of the National Academy of Sciences of the United States of America**, v.76, p.5269-5273, 1979.

PAGE, R.D.M. **TreeView**. Version 1.6.6. Glasgow: University of Glasgow, 2001. Available at: <http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>. Accessed on: 6 Aug. 2012.

PAVLÍČEK, A.; HRDÁ, S.; FLEGR, J. FreeTree – freeware program for construction of phylogenetic trees on the basis of distance data and bootstrap/jackknife analysis of the tree robustness. Application in the RAPD analysis of genus *Frenkelia*. **Folia Biologica**, v.45, p.97-99, 1999.

PEI, Z.; GAO, J.; CHEN, Q.; WEI, J.; LI, Z.; LUO, F.; SHI, L.; DING, B.; SUN, S. Genetic diversity of elite sweet sorghum genotypes assessed by SSR markers. **Biologia Plantarum**, v.54, p.653-658, 2010.

PRASAD, S.; SINGH, A.; JAIN, N.; JOSHI, H.C. Ethanol production from sweet sorghum syrup for utilization as automotive fuel in India. **Energy and Fuels**, v.21, p.2415-2420, 2007.

RATNAVATHI, C.V.; SURESH, K.; VIJAY KUMAR, B.S.; PALLAVI, M.; KOMALA, V.V.; SEETHARAMA, N. Study on genotypic variation for ethanol production from sweet sorghum juice. **Biomass and Bioenergy**, v.34, p.947-952, 2010.

REDDY, B.V.S.; RAMESH, S.; SANJANA REDDY, P.S.; RAMAIAH, B.; SALIMATH, P.M.; KACHAPUR, R. Sweet sorghum – a potential alternate raw material for bio-ethanol and bio-energy. **Journal of SAT Agricultural Research**, v.1, p.1-8, 2005.

RITTER, K.B.; MCINTYRE, C.L.; GODWIN, I.D.; JORDAN, D.R.; CHAPMAN, S.C. An assessment of the genetic relationship between sweet and grain sorghums, within *Sorghum bicolor* ssp. bicolor (L.) Moench, using AFLP markers. **Euphytica**, v.157, p.161-167, 2007.

ROHLF, F.J. **NTSYS-Pc**: numerical taxonomy and multivariate analysis system. Version 2.1. Setauket: Exeter Software, 2000. 44p.

SCHLOSS, S.J.; MITCHELL, S.E.; WHITE, G.M.; KUKATLA, R.; BOWERS, J.E.; PATERSON, A.H.; KRESOVICH, S. Characterization of RFLP probe sequences for gene discovery and SSR development in *Sorghum bicolor* (L.) Moench. **Theoretical and Applied Genetics**, v.105, p.912-920, 2002.

SHEHZAD, T.; OKUIZUMI, H.; KAWASE, M.; OKUNO, K. Development of SSR-based sorghum (*Sorghum bicolor* (L.) Moench) diversity research set of germplasm and its evaluation by morphological traits. **Genetic Resources and Crop Evolution**, v.56, p.809-827, 2009.

SMITH, J.S.C.; KRESOVICH, S.; HOPKINS, M.S.; MITCHELL, S.E.; DEAN, R.E.; WOODMAN, W.L.; LEE, M.; PORTER, K. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. **Crop Science**, v.40, p.226-232, 2000.

SOMERS, D.J.; BANKS, T.; DEPAUW, R.; FOX, S.; CLARKE, J.; POZNIAK, C.; MCCARTNEY, C. Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. **Genome**, v.50, p.557-567, 2007.

TARAMINO, G.; TARCHINI, R.; FERRARIO, S.; LEE, M.; PE, M.E. Characterization and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. **Theoretical and Applied Genetics**, v.95, p.66-72, 1997.

TATIKONDA, L.; WANI, S.P.; KANNAN, S.; BEERELLI, N.; SREEDEVI, T.K.; HOISINGTON, D.A.; DEVI, P.; VARSHNEY, R.K. AFLP-based molecular characterization of an elite germplasm collection of *Jatropha curcas* L., a biofuel plant. **Plant Science**, v.176, p.505-513, 2009.

VOS, P.; HOGERS, R.; BLEEKER, M.; REIJANS, M.; VAN DE LEE, T.; HORNES, M.; FRIJTERS, A.; POT, J.; PELEMAN, J.; KUIPER, M.; ET AL. AFLP: a new technique for DNA fingerprinting. **Nucleic Acids Research**, v.23, p.4407-4414, 1995.