CROP BREEDING AND APPLIED BIOTECHNOLOGY

NOTE

Crop Breeding and Applied Biotechnology
18: 426-431, 2018
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332018v18n4n62

Genetic diversity trends in sugarcane germplasm: Analysis in the germplasm bank of the RB varieties

Dennis Crystian[2], João Messias dos Santos[1], Geraldo Veríssimo de Souza Barbosa[1] and Cícero Almeida[2*]

Abstract: *Brazil is the largest sugarcane producer in the world and, the main varieties grown in Brazil, known as RB cultivars, were developed by the Interinstitutional Network for the Development of the Sugar and Alcohol Sector (RIDESA) and are used in 58.9% of the planted area in Brazil. These varieties were obtained through intercrosses between genotypes from the Serra do Ouro germplasm bank and successive crosses with related genotypes may have increased the level of genetic similarity. The aim of the present study was to analyze the genetic base of the Serra do Ouro germplasm bank over the past decades using microsatellite molecular markers. The genetic similarity among varieties using all the markers ranged from 0.166 to 0.823, and regression analysis showed an increase in genetic similarity in the 1970s; however, a narrowing of the genetic base over the last five decades was not observed.*

Keywords: *Simple sequence repeat (ssr), microsatellites, breeding, germplasm, saccharum.*

INTRODUCTION

In the early twentieth century, researchers in India and Java developed interspecific hybrids between the polyploid species *Saccharum officinarum* and *S. spontaneum* while maintaining noble and rustic features, respectively (Grivet et al. 2004). After successive backcrosses, modern hybrids have approximately 80% of the *S. officinarum* genome, 10% from *S. spontaneum,* and 10% recombinant between the two genomes (D'Hont et al. 1996). Sugarcane currently is very important to the world economy because of the production of sugar, ethanol, and energy from its biomass. In Brazil, sugarcane cultivation occupies a total area of 9.8 million hectares, with an output of 739 million tons (FAO 2013). The development of new varieties has contributed to expanding cultivation because the new genotypes are resistant to biotic and abiotic factors. The Interinstitutional Network for the Development of the Sugar and Alcohol Sector (RIDESA) and the Sugarcane Technology Centre (CTC) are currently the main centers for the development of new varieties of sugarcane in Brazil. They are responsible for developing the cultivars RB (República do Brasil) and SP (São Paulo), respectively. According to Barbosa et al. (2012), RB cultivars comprise 58.9% of the varieties that are planted and cultivated in Brazil, while the varieties developed by the Sugarcane Technology Centre comprise 35.8%.

***Corresponding author:**
E-mail: cicero@arapiraca.ufal.br

Received: 20 November 2016
Accepted: 29 August 2017

[1] Universidade Federal de Alagoas, Centro de Ciências Agrárias, Campus Delza Gitaí, BR104 - Norte, km 85, 57.072-900, Rio Largo, AL, Brazil
[2] Universidade Federal de Alagoas, Campus Arapiraca, Avenida Manoel Severino Barbosa s/n, Rodovia AL 115, km 6,5. 57.300-970, Bairro Bom Sucesso, AL, Brazil

Cultivars developed in the latter half of the twentieth century by RIDESA and CTC programs, although genetically distinct, have very high parental similarity because of the small number of hybrids used as a breeding base (Santos et al. 2012). In subsequent decades, intercrossing between modern hybrids became the main characteristic for developing new varieties. This plant group is characterized as aneuploid and polyploid hybrids, where recombination is the factor responsible for the genetic diversity of modern sugarcane accessions. Considering the genetic narrowing that has been caused in other crops by the use of a few clones as the genetic base (Wouw et al. 2010), there is the possibility of genetic narrowing in modern sugarcane hybrids.

Microsatellite markers are prominent among molecular tools for the genetic analysis of germplasm because they are highly polymorphic, allowing the analysis of genetic similarity between morphologically similar plant materials. This type of marker has been used in crops such as rice (Wu and Tanksley 1993), barley (Saghai-Maroof et al. 1984), and wheat (Roder et al. 1995). In the genus *Saccharum*, these markers can be used for the preselection of progenies in a breeding program to germplasm studies (Cordeiro et al. 2000, Cordeiro et al. 2001, Cordeiro et al. 2003, Santos et al. 2012, Silva et al. 2012). The objective of the present study was to analyze the basis of the genetic pool of sugarcane over the years using microsatellite molecular markers.

MATERIAL AND METHODS

Plant material

A set of 47 parents of sugarcane (Table 1) was obtained from the germplasm bank of "*Serra do Ouro*." The genotypes were classified according to the decades in which they were developed (year of crossing) and comparative evaluations were performed by means of contrasts restricted to cultivars of the same decade (Table 1).

Table 1. Sugarcane cultivars differentiated according to the decade of development and used to assess the genetic pool

Decade	Cultivars
1960 and earlier	H59-1966 L60-14 NA56-79 Q107
1970	RB72199 RB72454 RB75126 RB765418 SP70-1143 SP70-1284 SP71-1406 SP77-5181
1980	RB83160 RB835054 RB835089 RB842021 RB845197 RB845210 RB845257 RB855036 RB855113 RB855206 RB855453 RB855463 RB855511 RB855536 RB855546 RB863129 RB865230 RB867515 IAC86-2210
1990 and beyond	RB925211 RB925268 RB925345 RB92579 RB928064 RB931011 RB931530 RB935744 RB951541 RB956911 RB962962 RB965902 RB965917 RB966928 RB99395 RB002504

DNA extraction and amplification of the SSR Loci

DNA was isolated from the tissues of young leaves stored at −80 °C, using the procedure described by Saghai-Maroof et al. (1984). The integrity and concentration of each DNA sample was determined by spectrophotometry (absorbance at 260 and 280 nm) and gel electrophoresis in 1% agarose, respectively. Working stocks containing 50 µL were prepared at a concentration of 25 ng μL^{-1}. We used four SSR markers (SCC03, SCC05, SCC06 and SCC93) developed by the laboratory of genetic resources at the Arapiraca *Campus* of the Federal University of Alagoas. The names of the primers are listed in Table 2 and the sequences of the primers are found in Duarte Filho et al. (2010) and Silva et al. (2012).

For PCR amplification, a final volume of 50 µL was used containing: 25 ng of genomic DNA, 10× buffer, 2.0 mM MgCl2, 0.2 mM dNTP, 2.5 U of Taq-DNA polymerase, 30 pmol of each primer (forward and reverse) and sterile distilled water. The PCR amplifications were carried out in a thermal cycler (Applied Biosys-tems), with the following conditions: one cycle at 94 ^{0}C over 3 min for pre-denaturation, followed by 35 cycles at 94 ^{0}C for 1 min, 62 ^{0}C for 1 min and 72 ^{0}C for 1 min and a final extension at 72 ^{0}C for 12 min. Forward primers for each pair were labeled with different fluorescent dyes (6FAM or HEX) and arranged in duplex for analysis in an automated analyzer for DNA fragment analysis (ABI-3100, Applied Biosystems). The ROX 500 (Applied Biosystems, Foster City, CA) was used to accurately determine the

size of the fragments detected by capillary electrophoresis. These were then visualized in the form of peaks with the respective sizes and intensities and analyzed by Peak Scanner Software (Applied Biosystems).

DATA ANALYSIS

Despite being considered co-dominant SSR markers, in this study they were considered as dominant markers, because in highly polyploid genomes such as that of sugarcane, the SSR markers have difficulty distinguishing the alleles of homologous chromosomes, making it difficult to determine heterozygosity or homozygosity at any particular locus (Cordeiro et al. 2003, Oliveira et al. 2009). From this assumption, all possible alleles detected in the varieties have been converted to a binary system. For each clear and distinct peaks were classified as absent (0) or present (1) to form the matrix that was used to estimate the following variables:

Number of alleles with absent or present among genotypes (Np).

Number of alleles without absent among genotypes (Nnp).

Polymorphism information content (PIC) obtained through the expression: $PIC = 2fi(1 - fi)$, where *fi* is the frequency of the amplified fragments and *1-fi* is the frequency of the non-amplified fragments (Roldan-Ruiz et al. 2000) and the PIC for each primer was obtained using the average from all fragments.

Genetic similarity (GS) between pairs of genotypes of the same decade (Table 1) was calculated using the Jaccard coefficient, obtained by the expression: $GS = \frac{a}{(a + b + c)}$, where GS is the measure of genetic similarity between genotype i and j. For each pair of accessions, "a" represents the number of coincidences of the type 1−1, "b" the number of coincidences 1–0 and "c" the type of 0–1 (Reif et al. 2005).

Linear regression between genetic similarity and decade of the genotypes was applied for best-fit model.

RESULTS AND DISCUSSION

Sugarcane hybrids are polyploid and aneuploid and may have variable numbers of chromosomes; therefore, they exhibit high levels of allelic variation among individuals in the same *locus* (Santos et al. 2012). In the present study, the number of alleles per locus among genotypes analyzed varied between 3 and 16, with a mean of 7.5 (Table 2). The four *loci* amplified a total of 90 alleles, ranging in size from 81 to 236 bp. The number of alleles for all individuals at each *locus* ranged between 17 and 27 with a total average of 22.5 alleles (Table 2). The four SSR markers used had a number of alleles for discriminating all the subjects analyzed. This ability is the result of sugarcane's aneuploid and polyploid nature, with many alleles in the same individual, which increases the discriminatory capacity of the microsatellite markers (Cordeiro et al. 2001, Pinto et al. 2004, Silva et al. 2012). Previous studies using the *Serra do Ouro* germplasm bank have shown that microsatellite markers are highly efficient in distinguishing different genotypes (Santos et al. 2012, Silva et al. 2012).

Table 2. Polymorphic Information Content (PIC), Jaccard coefficient, Number of alleles with absent or present among genotypes (N_p), Number of alleles only with present among genotypes (N_{np}), allelic interval (bp) and number of alleles per individual

SSR	GenBank Number	**Primer** (5´-3´)	**Tm** (ºC)	**PIC**	**Jaccard coefficient**			N_p	N_{np}	**Allelic range** (bp)	**Alleles per individual**
					Max	**Min**	**Mean**				
SCC03	CA082520	F: CTCCGCATTAGCCATTTCC R: TGGTACTCGTCCATGTCGTC	57	0.26	0.937	0.167	0.450	27	0	125-180	7-16
SCC05	CA205346	F: CGGAATCCAATTCGTACGTT R: CATTGGTTGCACCACAGTTC	56	0.30	0.889	0.058	0.357	27	0	81-200	3-16
SCC06	CA207738	F: TATTCCACCGGGAACAAGAA R: GGGATTGTAGCGACGAGTTG	57	0.21	0.900	0.182	0.556	19	0	179-236	4-10
SCC93	CA210595	F:AATCCCAGCCCCGATGAT R: GCCACACCTTGACCTTGAC	58	0.29	0.923	0.143	0.500	17	0	158-196	3-13
All	-			0.27	0.823	0.166	0.448	22.5	0	-	7.5

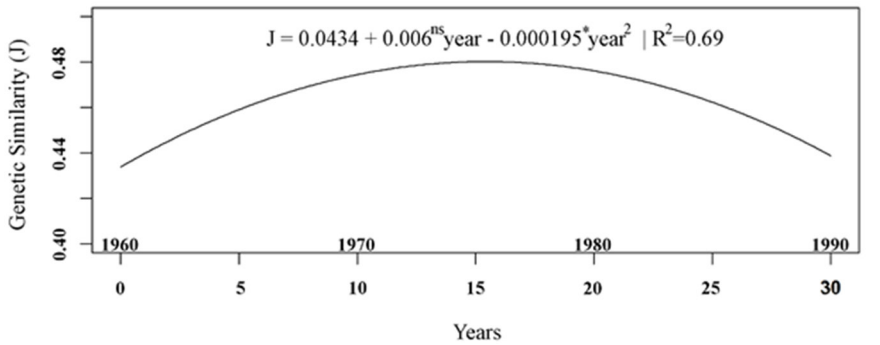

Figure 1. Genetic similarity index versus time (* Significant at the 5% significance level).

Table 3. Genetic Similarity among genotypes by decade

SSR	Jaccard coefficient											
	60s and earlier			70s			80s			90s and beyond		
	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
SCC03	0.75	0.263	0.402	0.625	0.187	0.42	0.937	0.2	0.5	0.846	0.067	0.416
SCC05	0.444	0.153	0.316	0.642	0.2	0.384	0.8	0.062	0.363	0.778	0.067	0.363
SCC06	0.667	0.454	0.55	0.857	0.4	0.556	0.9	0.333	0.625	0.875	0.2	0.5
SCC93	0.818	0.5	0.641	0.923	0.2	0.416	0.888	0.167	0.5	0.909	0.2	0.5
All	0.556	0.428	0.475	0.667	0.325	0.453	0.823	0.279	0.473	0.718	0.241	0.431

The polymorphic information content (PIC) using all the markers was 0.925, while individually, the microsatellites SCC03, SCC05, SCC06, and SCC93 showed values of 0.942, 0.938, 0.897, and 0.913, respectively (Table 2). The genetic similarity between varieties using all the markers ranged from 0.166 to 0.823 (Table 2). The genetic similarity among genotypes within decades showed that the 1980s had the highest similarity coefficient, corresponding to 0.823, while the lowest genetic similarity, with a value of 0.241, was found in the 1990s.

To analyze the genetic similarity as a function of time, a regression model was fitted to the Jaccard similarity coefficient. The results showed quadratic regression as the most suitable statistical model for describing the data, in view of the significance of the model parameters. The results indicated an increase in genetic similarity during the period 1970-1980, while the varieties developed in the 1990s have levels of genetic similarity close to those from the 1960s (Figure 1).

The use of just a few genotypes in the sugarcane breeding programs during the 1970s and 1980s, which mainly aimed at increasing sucrose levels, was one of the factors that raised the levels of genetic similarity. In the present study, the average similarity was 0.448, indicating values near those found by Lima et al. (2002), Pinto et al. (2006) who reported moderate genetic similarity coefficients (0.48 and 0.62, respectively) among accessions used in three sugarcane genetic breeding programs. From the 1990s onwards, there was a change in sugarcane breeding strategy because of the emergence of new pests and diseases and the need for cultivars that were more adapted than the existing ones. Given the new scenario, the breeding programs included new parent varieties, contributing to a lower genetic similarity index among the varieties developed after 1990.

Analyses of levels of genetic diversity in various crops showed a narrowing trend of the genetic base during the 1960s, 1970s, and 1980s (Wouw et al. 2010). This behavior was observed in the present study, with the highest similarity levels in the 1970s and 1980s (significant quadratic parameter); however, there was not a linear increase or decrease of the genetic bases. The increase in similarity levels during the 1970s and 1980s resulted from interbreeding with few parents; after the 1990s, with the introduction of new parents, there was a decrease in genetic similarity levels. It should be noted that the levels of genetic diversity in sugarcane breeding have remained the same over the past five decades, and that

the varieties developed resulted from the genetic diversity of the first hybrids. These were obtained by interbreeding *S. officinarum* and *S. spontaneum,* resulting in hybrids with different genomic constitution ratios (Piperidis et al. 2010) because of different chromosomal proportions (10-20% of *S. spontaneum* and 80-85% of *S. officinarum*). This condition allows a large number of combination ratios of the two genomes during chromosomal segregation in meiosis, which can be exploited in breeding. However, given the new requirements for sugarcane, such as the use of fiber for second generation ethanol, coproduction of electricity, and use of fiber as an organic fuel, new parents with different ratios of *S. spontaneum* should be incorporated into crossings.

ACKNOWLEDGEMENTS

We thank the Federal University of Alagoas for the laboratories and scientific support and the Fundação de Apoio à Pesquisa de Alagoas (FAPEAL) for funding this Project.

REFERENCES

Barbosa MHP, Resende MDV, Dias LAS, Barbosa GVS, Oliveira RA, Peternelli LA and Edelclaiton D (2012) Genetic improvement of sugar cane for bioenergy: the Brazilian experience in network research with RIDESA. **Crop Breeding and Applied Biotechnology 12**: 87-98.

Cordeiro GM, Taylor GO and Henry RJ (2000) Characterisation of microsatellite markers from sugarcane (Saccharum sp), a highly polyploid species. **Plant Science 155**: 161-168.

Cordeiro GM, Casu R, McIntyre CL, Manners JM and Henry RJ (2001) Microsatellite markers from sugarcane (Saccharum spp) ESTs cross transferable to erianthus and sorghum. **Plant Science 160**: 1115-1123.

Cordeiro GM, Pan YB and Henry RJ (2003) Sugarcane microsatellites for the assessment of genetic diversity in sugarcane germplasm. **Plant Science 165**: 181-189.

D'Hont A, Grivet L, Feldmann P, Rao S, Berding N and Glaszmann JC (1996) Characterisation of the double genome structure of modern sugarcane cultivars (Saccharum spp) by molecular cytogenetics. **Molecular & General Genetics 250**: 405-413.

Duarte Filho LSC, Silva PP, Santos JM, Barbosa GVS, Ramalho-Neto CE, Soares L, Andrade JCF and Almeida C (2010) Genetic similarity among genotypes of sugarcane estimated by SSR and coefficient of parentage. **Sugar Tech 12**: 145-149.

FAO (2013) **Food and Agriculture Organization of the United Nations.** Available at <http://faostatfaoorg/site/567/DesktopDefaultaspx?PageID=567 2013>. Accessed on Oct 1th.

Grivet L, Daniels C, Glaszmann JC and D'hont A (2004) A Review of Recent Molecular Genetics Evidence for Sugarcane Evolution and Domestication. **Ethnobotany Research & Applications 2**: 9-17.

Lima, MLA, Garcia AAF, Oliveira KM, Matsuoka S, Arizono H, Souza CL and Souza AP (2002) Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum* spp). **Theoretical and Applied Genetics 104**: 30-38.

Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas SM, Falco MC, Burnquist W, Garcia AAF and Souza AP (2009) Characterization of new polymorphic functional markers for sugarcane. **Genome 52**: 191-209

Pinto LR, Oliveira KM, Ulian EC, Garcia AAF and Souza AP (2004) Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. **Genome 47**: 795-804.

Pinto LR, Oliveira KM, Marconi T, Garcia AAF, Ulian EC and Souza AP (2006) Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. **Plant Breeding 125**: 378-384.

Piperidis G, Piperidis N and D'Hont A (2010) Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. **Molecular Genetics and Genomics 284**: 65-73.

Reif JC, Melchinger AE and Frisch M (2005) Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. **Crop Science 45**: 1-7.

Roldan-Ruiz I, Dendauw J, Vanbockstaele E, Depicker A and De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). **Molecular Breeding 6:** 125-134.

Roder MS, Plaschke J, Konig SU, Borner AE and Sorrels ME (1995) Abundance, variability and chromosomal location of microsatellites in wheat. **Molecular and General Genetics 246**: 327-333.

Saghai-Maroof MA, Soliman KM, Jorgensen RA and Allard RW (1984) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. **Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 81**: 8014-8018.

Santos JM, Duarte Filho LSC, Soriano ML, Silva PP, Nascimento VX, Barbosa GVS, Todaro AR, Ramalho Neto CE and Almeida C (2012) Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. **Industrial Crops and Products 40**: 145-150.

Silva DC, Duarte Filho LSC, Santos JM, Barbosa GVS and Almeida C (2012) DNA fingerprinting based on simple sequence repeat (SSR) markers in sugarcane clones from the breeding program RIDESA. **African Journal of Biotechnology 11:** 4722-4728.

Wouw M, Hintum T, Kik C, Treuren R and Visser B (2010) Genetic diversity trends in twentieth century crop cultivars: a meta analysis. **Theoretical and Applied Genetics 120**: 1241-1252.

Wu KS and Tanksley SD (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. **Molecular and General Genetics 241**: 225-235.