Euphytica (2012) 185:491–510
DOI 10.1007/s10681-012-0661-9

Genetic diversity assessment and genotype identification in sugarcane based on DNA markers and morphological traits

M. F. Perera · M. E. Arias · D. Costilla · A. C. Luque · M. B. García · C. Díaz Romero · J. Racedo · S. Ostengo · M. P. Filippone · M. I. Cuenya · A. P. Castagnaro

Received: 28 June 2010 / Accepted: 8 March 2012 / Published online: 24 March 2012

Abstract Sugarcane is known for its highly complex genetics and more knowledge is needed for better use and conservation of genetic materials. In order to identify genotypes and to assess genetic diversity, diverse data sets such as morphological and molecular markers are used as a general approach. To evaluate the usefulness of different markers, important sugarcane genotypes in Argentina were characterized by AFLP, SSR and morphological traits. All genotypes characterized were grouped in one main cluster in dendrograms using two independent softwares. Interestingly, local genotypes grouped together with USA varieties and no clear genetic differentiation could be found probably due to intensive germplasm exchange between these breeding programs. The molecular markers tested were useful for genetic diversity assessment as well as for genotype identification. These markers should be included in the internationally established characters for sugarcane variety protection as they give a better view on whole genome complexity. Additionally, genetic similarities obtained from molecular markers will provide more accurate information to breeders than the pedigree method, especially when considering the asymmetric genetic inheritance of sugarcane. Morphological traits are valuable tools to identify genotypes since they reflect external resemblance more than genetic relatedness. When they were combined with molecular markers the dendogram obtained revealed genetic relationships and the genetic diversity was better estimated. In summary, both methods appear to be useful, complementing each other and should be used together to assist sugarcane breeders in estimating genetic diversity, electing parents for crossings, identifying superior lines and to protect intellectual property rights.

Keywords AFLP · Breeding program · Sugarcane germplasm bank · Morphological descriptors · Similarity coefficient · SSR

M. F. Perera · J. Racedo · M. P. Filippone · A. P. Castagnaro (✉)
Sección Biotecnología - Unidad Asociada al Instituto Superior de Investigaciones Biológicas; Consejo Nacional de Investigaciones Científicas y Técnicas - Universidad Nacional de Tucumán (INSIBIO; CONICET-UNT), Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Av. William Cross 3150, T4101XAC Las Talitas, Tucumán, Argentina
e-mail: atilio@eeaoc.org.ar; atiliocastagnaro@gmail.com

M. E. Arias · A. C. Luque
Facultad de Ciencias Naturales e Instituto Miguel Lillo, Universidad Nacional de Tucumán, Miguel Lillo 205, 4000 San Miguel de Tucumán, Tucumán, Argentina

D. Costilla · M. B. García · C. D. Romero · S. Ostengo · M. I. Cuenya
Subprograma Mejoramiento Genético de la Caña de Azúcar, Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Av. William Cross 3150, T4101XAC Las Talitas, Tucumán, Argentina

Commercial sugarcane varieties all belong to *Saccharum*, a complex genus characterised by a high degree of polyploidy and frequent aneuploidy (Cordeiro et al. 2000; Da Silva and Bressiani 2005). These characteristics and the cytogenetic complexity of sugarcane cultivars, involving varying chromosome sets and complex recombinational events, imposes difficulties in accomplish effective breeding programs (Vettore et al. 2001). The sugarcane breeding process starts with cross hybridizations either between two (bi-parental cross) or a few (polycross) elite clones (Breaux and Legendre 1983; Heinz 1987). Clones are classified as male or female parents based on the relative amounts of viable pollen produced (McIntyre and Jackson 2001).

In Argentina, the sugarcane industry started in the late nineteenth century and the first sugarcane breeding program was formally established 44 years ago by the Estación Experimental Agroindustrial Obispo Colombres (EEAOC) in the Province of Tucumán. It takes at least 11 years to complete a sugarcane breeding cycle, starting with a crossing between two elite clones (bi-parental), evaluating the progeny to identify true hybrids, several stages of testing and clonal selection and finally ending with a new variety release. In order to broaden the genetic base of the commercial varieties, parents more genetically diverse should be identified and used in breeding programs (Salem et al. 2008). Besides, exchange of elite clones and breeding lines occur regularly between different breeding programs. For these reasons and for protecting new sugarcane varieties and intellectual property rights, an accurate varietal identification is essential (Wagih et al. 2004). In addition, the knowledge of the genetic diversity in sugarcane will provide useful information concerning the genotype value to breeders and will contribute to the improved use and conservation of genetic resources.

Molecular markers are powerful tools to estimate genetic diversity and to better understand the complex genetics of sugarcane as they are accurate, abundant and not affected by the environment (D'Hont et al. 1997). In terms of monitoring genetic diversity and identification of germplasm, simple sequence repeats (SSR) markers are one of the best choices (Smith et al. 1994). Their use is preferred over other marker techniques because of their abundance, sensitivity and high accuracy in detecting polymorphism even between closely related genotypes (Powell et al. 1996). However, as sugarcane has a very large and complex genome a vast marker number is necessary to determine its diversity (Lima et al. 2002). In addition to SSR markers, the use of AFLP (amplified fragment length polymorphisms) marker reveals high polymorphic band numbers (Vos et al. 1995) and they have successfully been used to determine the genetic diversity in several other plant species (Hill et al. 1996; Maughan et al. 1996; Paul et al. 1997; Angiolillo et al. 1999; Selvi et al. 2003).

Additionally to the recently DNA-based marker data, morphological traits are of great help for genetic resource evaluation (Demey et al. 2003) and the availability of an internationally adopted descriptor set would enable an international measurement of genetic distance and cultivar label maintenance both within and between research stations (Gallacher 1997). Thus, morphological traits proposed by the International Union for the Protection of New Varieties of Plants (UPOV 2005) allow for variety characterization. Although morphological characters sometimes do not give clear answers due to ambiguous differences or phenotypic modifications caused by environmental factors (García et al. 2002), the traits proposed by UPOV are known to be very stable under environmental conditions and would therefore be able to be employed not only as an identification tool but also as a source of information for genetic diversity.

Despite the economic importance of sugarcane production in Argentina, to date, no major descriptive study has been carried out to identify genotypes and to estimate the genetic diversity of our breeding materials. It is the aim of this work to evaluate the usefulness of different characters for genotype identification and genetic diversity assessment, by characterizing 36 sugarcane genotypes most widely used as parents in the breeding program conducted by EEAOC, using two different molecular techniques, two data analysis softwares and morphological traits only in the eight most important cultivars.

Sugarcane genotype characterization by using different marker systems will allow us to identify genotypes as well as to assess the sugarcane genetic diversity better and this will provide important information to assist breeders.

Materials and methods

Sugarcane genotypes

Thirty-six sugarcane genotypes (Table 1) kept at EEAOC and used as parents in its local breeding program, were examined by DNA fingerprinting studies. Out of the 36 genotypes, 17 are commercial cultivars in Argentina or in the United States, while the rest has not been commercially released. Two young leaves from each genotype were collected and stored at −70 °C until extraction of total plant DNA.

DNA extraction

Two-hundred mg of frozen tissue from each genotype was placed in liquid nitrogen and ground in a mortar. Total DNA was extracted by using the protocol described by Aljanabi et al. (1999). DNA concentration and quality were determined by measuring the OD_{260} in a spectrophotometer and by running the samples in agarose gel electrophoresis (0.7%), respectively.

Molecular markers

Analysis of AFLPs were performed according to Vos et al. (1995). Sixty-four primer combinations were evaluated by using 10 genotypes randomly chosen. Sixteen pairs were selected based on the presence of scorable bands and/or on the high numbers of polymorphic bands (Table 2), and used for the genotype characterization.

Fifteen SSR primers were used for genetic diversity studies (Cordeiro et al. 2000) (D′Hont, unpublished). The DNA amplification reaction mix contained: 1× buffer, 100 μM dNTPs, 10 ng DNA; primers, $MgCl_2$ and Taq DNA polymerase as shown in Table 3. Cycling parameters were: 1 cycle at 94 °C (5 min); 30 cycles at 94 °C (45 s), appropriate annealing temperature and time (Table 3) and 72 °C (1 min); and 1 cycle at 72 °C (3 min).

All DNA amplifications were performed on a Model PTC-100TM Programmable Thermal Controller, Peltier-Effect Cycling (MJ Research, Inc.). AFLP and SSR products were separated by electrophoresis on 6% polyacrylamide denaturing gel and visualized through silver staining (Caetano-Anollés and Gresshoff 1994).

Polymorphic locus proportion (PLP95), amplification percentage (data percentage = 1), and polymorphism information content (PIC) (Botstein et al. 1980) were calculated by using the Info-Gen software (Balzarini and Di Rienzo, 2003). The former considers a locus as polymorphic if it has population variations and the most common allele frequency does not exceed 95%. Genetic diversity for SSR primers was estimated by using calculations according to Anderson et al. (1993), that has frequently been used to determine the marker value in detecting polymorphism, i.e., PIC.

Morphological traits

Only a subgroup of eight genotypes were morphologically characterized (Table 4). These genotypes were selected because three of them (LCP 85-384, RA 87-3 and TUCCP 77-42) are the most widely planted varieties in the sugarcane growing area in Tucumán (Argentina) (Cuenya et al. 2009); another three varieties (TUC 89-28, TUC 95-37 and TUC 97-8) were released in 2009; while the rest are advanced breeding clones at the final testing stage before a possible commercial release (TUC 97-7 and TUC 95-24; Cuenya, personal communication).

All plant materials were characterized using plant cane (10–12 months of age). Six and twenty-four plants from the same variety were evaluated for qualitative and quantitative characters, respectively, by using the test guidelines applied to all vegetative propagated varieties of *Saccharum* L. proposed by UPOV (2005) (Table 4).

Out of the characters proposed by UPOV, nine are compulsory to identify sugarcane varieties. In order to determine if the UPOV characters, are stable under different environmental conditions, eight of the compulsory characters were evaluated during two years (2010–2011) and three different growth locations (Cevil Pozo, Las Talitas and Santa Ana) in Tucumán, Argentina.

Data analysis

No assumption on the genetic nature of the alleles was made due to the polyploid nature of sugarcane and the absence of a segregation analysis (Gillet 1991). Hence, each molecular and morphological allele was scored in a dominant manner and transformed into either a 0 (absent) or 1 (present) matrix. Although SSRs are classified as co-dominant type markers, they

Table 1 Thirty-six sugarcane genotypes from EEAOC breeding program, their origin and parents

Genotypes	Origin	Varieties	Parents	
			Female	Male
HO 95-888	Houma (USA)	Commercial	CP 86-941	US 89-12
HOCP 85-845	Houma-Canal Point (USA)	Commercial	CP 72-370	CP 77-403
HOCP 91-552			LCP 81-10	CP 72-356
HOCP 91-555			CP 83-644	LCP 82-94
HOCP 92-618			CP 78-304	LCP 81-30
HOCP 92-624		Non commercial	CP 81-325	CP 71-1038
HOCP 92-631			CP 81-325	CP 71-1038
HOCP 92-675			CP 83-644	CP 70-321
HOCP 93-746			LCP 81-10	CP 82-513
HOCP 94-806			CP 81-325	CP 71-1038
L 91-281	Louisiana (USA)	Commercial	CP 78-317	LCP 81-30
L 97-128			LCP 81-10	LCP 85-384
L 99-226			HOCP 89-846	LCP 81-30
L 99-233			CP 79-348	HOCP 91-552
L 00-266		Non commercial	HOCP 89-846	L 93-386
L 89-113			CP 78-317	LCP 81-30
L 94-424			LCP 81-10	LCP 82-89
L 98-209			LCP 86-454	LCP 85-384
LCP 82-89	Louisiana-Canal Point (USA)	Commercial	CP 52-68	CP 72-370
LCP 85-384[a]			CP 77-310	CP 77-407
LCP 86-454			CP 77-310	CP 69-380
LHO 83-153	Louisiana-Houma (USA)	Commercial	CP 77-405	CP 74-339
RA 87-3[a]	República Argentina	Commercial	TUC 75-25	CP 57-614
TUC 89-28[a]	Tucumán (Argentina)	Commercial	TUCCP 77-42	TUCCP 77-42
TUC 95-37[a]			CP 65-357	S 87-1756
TUC 97-8[a]			TUC 87-21	TUCCP 77-42
TUCCP 77-42[a]			CP 71-321	US 72-019
TUC 92-10		Non commercial	TUC 83-1	TUC 77-37
TUC 94-61			TUC 77-52	TUCCP 77-42
TUC 95-24[a]			CP 79-348	LCP 85-358
TUC 95-34			TUC 84-24	TUC 84-13
TUC 96-52			LCP 85-384	TUC 83-8
TUC 97-7[a]			LCP 85-384	LCP 86-454
TUC 97-19			NA 63-90	TUCCP 77-42
TUC 97-21			TUC 87-21	TUCCP 77-42
TUC 97-30			TUC 87-21	TUCCP 77-42

[a] Eight sugarcane genotypes selected for morphological analysis

have been treated as dominant markers in this study. This novel approach was necessary to analyse the highly complex genome of *Saccharum*.

Genetic similarity was calculated by using Jaccard (*Sj*) (Sneath and Sokal 1973) and Dice coefficients (S_D) (Dice 1945). $S_j = A/(A + B + C)$ and $S_D = 2A/(2A + B + C)$, where A is the number of bands common to the first and second genotypes, B is the number of bands unique to the first genotype, and C the number of bands unique to the second genotype.

Table 2 AFLP primer information of sugarcane

Primers		Amplified fragments	Polymorphic fragments	PLP[c]	PIC[d]	AMP[e]
MseI[a]	EcoRI[b]					
CTT	AGG	73	8	0.04	0.08	90.79
	ACG	66	7	0.06	0.11	90.07
CAT	ACT	55	18	0.20	0.12	81.57
	AAC	47	3	0.04	0.19	94.74
CAC	AGC	63	32	0.32	0.12	65.04
	AAG	66	3	0.03	0.08	99.79
	ACA	58	13	0.10	0.01	92.34
CAA	ACC	55	5	0.09	0.28	93.38
CTC	ACC	30	4	0.10	0.16	91.20
	AAG	79	17	0.22	0.19	84.32
CTA	AGC	59	13	0.14	0.13	84.23
	ACG	77	18	0.21	0.23	84.13
CAG	ACT	70	16	0.19	0.18	83.57
	ACA	53	11	0.06	0.09	90.57
	AAC	69	18	0.22	0.23	83.33
CTG	AGG	75	7	0.07	0.20	92.30

[a, b] MseI (GACTGCGTACCAATTC), EcoRI (GATGAGTCCTGAGTAA) sequence primer, respectively, plus 3 bases

[c] PLP: polymorphic locus proportion

[d] PIC: polymorphism information content (Botstein et al. 1980)

[e] AMP: amplification percentage

Cluster analyses were carried out using McQuitty similarity analysis (Unweighted Pair Group Method with Arithmetic Mean, UPGMA) (Sneath and Sokal 1973). The FIND module was used to identify all trees and they were compiled by the CONSEN module to test the robustness of the tree topology. All bands, monomorphic and polymorphic, were included in the analysis. All calculations were carried out by using the two software packages NTSys (Rohlf 1993) and InfoStat (Di Rienzo et al. 2009) in order to compare their robustness. The NTSys program is a commonly used program that presents the result as similarity and the InfoStat, a new and user friendlier one, as distance (1-S, where S is genetic similarity value).

In addition, Spearman′s correlation coefficients were calculated in order to compare genetic distances obtained from the AFLP, SSR and morphological trait.

Coefficient of parentage

The calculation of the coefficient of parentage (f), between two genotypes, as defined by Kempthorne (1957), was carried out and corresponds to the probability that alleles in a locus are identical by descent to alleles in the same locus in another cultivar. The f values were calculated using the procedure "proc inbreeding" of the software SAS (version 9.1., SAS Institute, 2003). The assumptions suggested by Cox et al. (1985) were adopted, and f was considered 0 among the remote ancestors. For each genotype it was assumed that the inbred coefficient was 0, due to the heterozygous character of the genotypes of this crop (Chang and Lo 1993; Deren 1995).

To determine the correlation level among genetic similarity coefficients obtained from AFLP, SSR and morphological traits and, analysis of correlations were carried out using Pearson's coefficient (r).

Hybridity determination

The emasculation treatment implied pollen sterilization by immersion of the panicle of two commercial varieties commonly used as males, LCP 85-384 and RA 87-3, in a hot water tank at 50 °C for 5 min. The combinations studied were, a: LCP 85-384 emasculated × RA 87-3, b: RA 87-3 emasculated × LCP

Table 3 SSR primer information from sugarcane

Primer	[$MgCl_2$] mM	Taq units	[Primer] μM	Annealing conditions	Repeat motifs	Amplified fragments	Polymorphic fragments	Genetic diversity[c]	PLP[d]	PIC[e]	AMP[f]
SMC222CG[a]	2	0.5	1.2	63 °C 30 s	$(CA)_{24}$	9	9	0.94	0.89	0.32	55.86
SMC226CG	2	0.5	1.2	58 °C 30 s	$(CA)_{10}$	10	10	0.92	0.80	0.25	46.67
SMC248CG	2.5	0.3	1	63 °C 30 s	$(TTA)_6$	3	2	0.61	0.67	0.28	68.52
SMC319CG	2.5	0.5	1.2	59 °C 30 s	$(CA)_{17}$	15	12	0.94	0.80	0.30	60.93
SMC477CG	2.5	0.3	1	64 °C 30 s	$(CA)_{31}$	6	5	0.70	0.67	0.24	56.94
SMC863CG	2.5	0.3	1	55 °C 30 s	$(TC)_9$	11	6	0.91	0.55	0.28	83.33
MSCIIR1[b]	1.6	0.5	1	58 °C 1 min	$(GT)_{18}(GA)_{31}$	10	6	0.95	0.60	0.31	70.56
MSCIIR12	2.5	0.8	1.2	56 °C 30 s	$(GA)_{13}$	7	7	0.77	1.00	0.21	58.73
MSCIIR14	2.5	0.5	1.5	61 °C 30 s	$(GA)_{22}$	6	5	0.81	0.83	0.23	46.30
MSCIIR15	2.5	0.5	1.2	65 °C 30 s	$(GA)_{22}$	15	5	0.63	0.27	0.18	76.11
MSCIIR16	2.5	0.8	1.2	56 °C 1 min	$(GA)_{18}$	10	7	0.79	0.60	0.22	56.11
MSCIIR19	2.5	0.5	1.2	52 °C 30 s	$(GA)_{23}$	10	9	0.95	0.80	0.28	55.83
MSCIIR27	3	0.8	1.2	56 °C 30 s	$(GA)_{18}$	7	5	0.80	0.71	0.23	41.27
MSCIIR34	2.5	0.8	1	57 °C 30 s	$(GA)_{20}$	9	6	0.90	0.67	0.29	59.26
MSCIIR68	2.5	0.3	1	56 °C 30 s	$(GT)_{26}GC(GT)_{12}$	8	7	0.91	0.88	0.31	52.43

[a] SMC primers (Cordeiro et al. 2000)

[b] MSCIIR: unpublished sequence primers (D′Hont unpublished)

[c] Genetic Diversity (Anderson et al. 1993)

[d] PLP: polymorphic locus proportion

[e] PIC: polymorphism information content (Botstein et al. 1980)

[f] AMP: amplification percentage

Table 4 Morphological characteristic determination of eight sugarcane genotypes

Characteristics	Sugarcane genotypes							
	LCP 85-384	TUCCP77-42	RA87-3	TUC 97-8	TUC 95-37	TUC89-28	TUC 95-24	TUC97-7
Stool growth habit	Erect	Erect	Inter-mediate	Erect	Erect	Erect	Semi-erect	Erect
Leaf sheath adherence[a]	Strong	Medium	Weak	Medium	Medium	Strong	Weak	Medium
Tillering	Strong	Medium	Weak	Medium	Medium	Medium	Medium	Medium
Sucker number	Very few	Very few	Many	Very few	Very few	Very few	Few	Very few
Leaf canopy	Dense	Dense	Very sparse	Dense	Medium	Dense	Dense	Medium
Green color intensity of leaf canopy	Medium	Dark	Light	Dark	Light	Dark	Medium	Light
Stem: culm height	Medium	Very long	Very long	Long	Very long	Very long	Medium	Mediim
Internode								
Length on bud side	Short	Long	Long	Medium	Long	Long	Very short	Medium
Diameter[a]	Thin	Thick	Thick	Thin	Thick	Medium	Medium	Medium
Shape[a]	Cylindrical	Cylindrical	Cylindrical	Conoidal	Conoidal	Conoidal	Conoidal	Conoidal
Cross-section	Circular	Circular	Circular	Circular	Ovate	Circular	Circular	Circular
Color where exposed to sun (Group:)[a]	63C	60C	79C	59C	58D	66C	57B	60A
Color where not exposed to sun (Group:)[a]	151B	146D	178A	146C	146D	151A	144A	144C
Growth crack depth	Absent	Shallow	Absent	Shallow	Absent	Absent	Absent	Absent
Zigzag alignment expression[a]	Weak	Absent	Absent	Moderate	Weak	Absent	Strong	Medorate
Appearance (rind surface)	Smooth	Smooth	Smooth	Smooth	Rough	Medium	Medium	Smooth
Waxiness	Medium	Medium	Strong	Weak	Strong	Medium	Strong	Medium
Node								
Root band width	Narrow	Medium	Broad	Broad	Broad	Broad	Medium	Broad
Wax ring	Narrow	Wide	Medium	Narrow	Medium	Medium	Narrow	Narrow
Bud shape[a]	Oval	Round	Ovate	Round	Oval	Ovate	Ovate	Ovate
Bud width, excluding wings	Narrow	Medium	Very wide	Medium	Wide	Wide	Wide	Wide
Bud prominence (on second senescent leaf from top)	Medium	Weak	Strong	Strong	Medium	Medium	Medium	Strong
Bud groove depth	Absent	Absent	Absent	Absent	Absent	Absent	Absent	Absent
Bud groove length	Absent	Absent	Absent	Absent	Absent	Absent	Absent	Absent
Bud tip position in relation to growth ring	Clearly below	Clearly below	Inter-mediate	Clearly below	Clearly below	Clearly below	Inter-mediate	Clearly below
Pubescence on the bud	Present	Present	Present	Present	Present	Present	Present	Present
Pubescence on the bud position	Lateral	Apice	Base	Lateral	Apice	Lateral	Lateral	Lateral
Bud cushion (space between base of bud and leaf scar)	Absent	Narrow	Medium	Absent	Absent	Absent	Absent	Narrow

Table 4 continued

Characteristics	Sugarcane genotypes							
	LCP 85-384	TUCCP77-42	RA87-3	TUC 97-8	TUC 95-37	TUC89-28	TUC 95-24	TUC97-7
Bud wing distribution	Median	Apical	Sub-median	Apical	Median	Apical	Median	Apical
Bud wing width	Narrow	Wide	Medium	Narrow	Medium	Medium	Narrow	Narrow
Leaf sheath								
Length	Short	Long	Medium	Long	Medium	Short	Short	Very short
Hair number (group 57)	Very many	Absent	Absent	Many	Many	Many	Few	Very many
Hair number (group 60)	Absent	Absent	Absent	Medium	Medium	Medium	Very many	Few
Hair length (group 57)	Long	Absent	Absent	Long	Short	Medium	Short	Short
Hair length (group 60)	Absent	Absent	Absent	Medium	Medium	Long	Long	Short
Hair distribution	Lateral and dorsal	Absent	Absent	Lateral and dorsal	Lateral and dorsal	Lateral and dorsal	Lateral and Dorsal	Lateral and dorsal
Ligule shape	Deltoid	Deltoid	Assymmetrica, horizontal	Deltoid	Crescent-shaped	Crescent-shaped	Deltoid	Crescent-shaped
Ligule width	Wide	Wide	Medium	Medium	Medium	Narrow	Wide	Medium
Ligule hair length (group 61)	Short	Medium	Long	Short	Short	Short	Long	Short Medium
Ligule hair density (group 61)	Dense	Medium	Dense	Medium	Sparse	Medium	Dense	Dense
Underlapping auricle shape	Lanceolate	Deltoid	Lanceolate	Deltoid	Deltoid	Deltoid	Transitional	Deltoid
Underlapping auricle size	Small	Small	Large	Small	Medium	Medium	Small	Small
Overlapping auricle shape	Lanceolate	Transitional	Transitional	Transitional	Transitional	Transitional	Transitional	Deltoid
Overlapping auricle size	Small	Small	Small	Small	Small	Small	Small	Small
Curvature	Straight	Straight	Curved at base	Straight	Arched	Straight	Curved tips	Curved tips
Width at the longitudinal mid-point[a]	Narrow	Broad	Narrow	Medium	Broad	Broad	Narrow	Broad
Leaf								
Midrib width	Narrow	Narrow	Narrow	Wide	Very wide	Medium	Wide	Very wide
Ratio leaf blade width/midrib width	High	High	High	Low	Low	Medium	Low	Low
Leaf blade								
Length	Long	Very long	Long	Short	Medium	Medium	Very short	Short
Pubescence on margin	Dense	Medium	Absent or very sparse	Absent	Absent	Absent or very sparse	Absent or very sparse	Absent or very sparse
Margin serration	Present	Present	Present	Absent	Present	Present	Present	Present
Cane top								
Length	Medium	Medium	Long	Short	Short	Short	Long	Long
Cross-section shape	Ovate	Ovate	Circular	Ovate	Ovate	Circular	Circular	Ovate
Waxines	Medium	Weak	Strong	Weak	Medium	Medium	Medium	Medium

[a] Eight compulsory characters to identify sugarcane varieties

85-384, c: LCP 85-384 selfed, d: RA 87-3 selfed, e: LCP 85-384 emasculated and selfed and f: RA 87-3 emasculated and selfed. Each cross was made by duplicating. Seeds were kept at −18 °C. Seedlings were held under natural light conditions in a greenhouse, temperatures ranging from 25 to 28 °C. In e and f no viable seeds were obtained.

Fresh leaves from 120 samples belonging to the first four cross combinations were collected and frozen at −70 °C. Then, leaves were ground to a fine powder in liquid nitrogen and DNA was extracted as it was described earlier.

Out of the 15 SSR primer pairs previously checked, MSCIIR19 was chosen because it produces an appropriate molecular profile in the studied genotypes. All samples were amplified and segregation analysis was performed with χ^2 test ($P \leq 0.05$).

Results

AFLPs

It was possible to discriminate each genotype with the 16 primer combinations (Table 2) which generated 62 bands on average. A total of 995 fragments, out of which 193 were polymorphic, were identified. An average of 19.40% polymorphism was obtained. PLP95 ranged from 0.03 to 0.32. PIC values and amplification percentage ranged from 0.01 to 0.28 and 65.04 to 99.79%, respectively (Table 2).

Dendograms obtained with the two softwares showed the same clusters; where related genotypes tended to group together. All cultivars grouped in one main cluster, divided into at least seven subgroups, presenting a high similarity degree and similar complex profiles. Genotype similarities ranged between 0.94 and 0.99, and 0.97 and 0.99 by using Jaccard (Fig. 1a) and Dice coefficients, respectively. Dendograms of the eight genotypes (selected for morphological analysis), with both programs were the same and similarities ranged between 0.95 and 0.99 by using Jaccard coefficient (Fig. 1b).

SSRs

All 15 primers are applicable for sugarcane identification because they produced unique and reproducible amplification profiles. The allele number produced per marker across the 36 genotypes varied from three to 15 and total number produced was 136, out of which, 101 were found to be polymorphic. A minimum of three markers were required to correctly identified and distinguish all genotypes.

SSR markers had a genetic diversity value between 0.61 and 0.95 (mean value: 0.84), calculated by using the Anderson et al. (1993) formula. PIC values ranged from 0.18 to 0.32. PLP95 and amplification percentage, ranged from 0.27 to 1.00 and 41.27 to 83.33%, respectively (Table 3).

Dendograms of all 36 genotypes showed the same clusters using both softwares. Related genotypes tended to group together, as well as those from the same origin. Genotype similarities ranged between 0.57 and 0.91 and 0.73 and 0.96, by using Jaccard and Dice coefficients, respectively. When the eight selected genotypes were grouped, similarities ranged between 0.57 and 0.82 by using Jaccard coefficient; however, both softwares did not show identical clusters. The NTSys dendogram better reflected genotype pedigree because TUC 89-28 was obtained by selfing TUCCP 77-42 and they shared 83% of similarity, while they shared less similarity (81%) with InfoStat. Besides, TUCCP 77-42 is the male parent of TUC 97-8 and their similarity was 67% with NTSys and with InfoStat was only 38%. Furthermore, TUC 95-24 and TUC 97-7 belong to the same cluster as they have related parents, crossing and selecting in the same environment (Table 1; Fig. 2); however, these genotypes did not cluster together with InfoStat. Thus, a tendency for cultivars to group together with others obtained from the same or related cross was observed by using NTSys. Moreover, SSR data reflected genetic relations more clearly than AFLP (Fig. 1b vs Fig. 2).

When both, AFLP and SSR, were used to generate dendogram for the 36 genotypes, similarity ranged between 0.91 and 0.98 by using Jaccard coefficient, indicating that there is low genetic diversity between these varieties. In the dendogram of the eight genotypes selected for morphological characterization, similarities ranged between 0.92 and 0.95 when using Jaccard coefficient, and clusters differed from those obtained with AFLP.

Morphological traits

Fifty-two morphological characters describing various plant parts out of those proposed by UPOV (2005) were

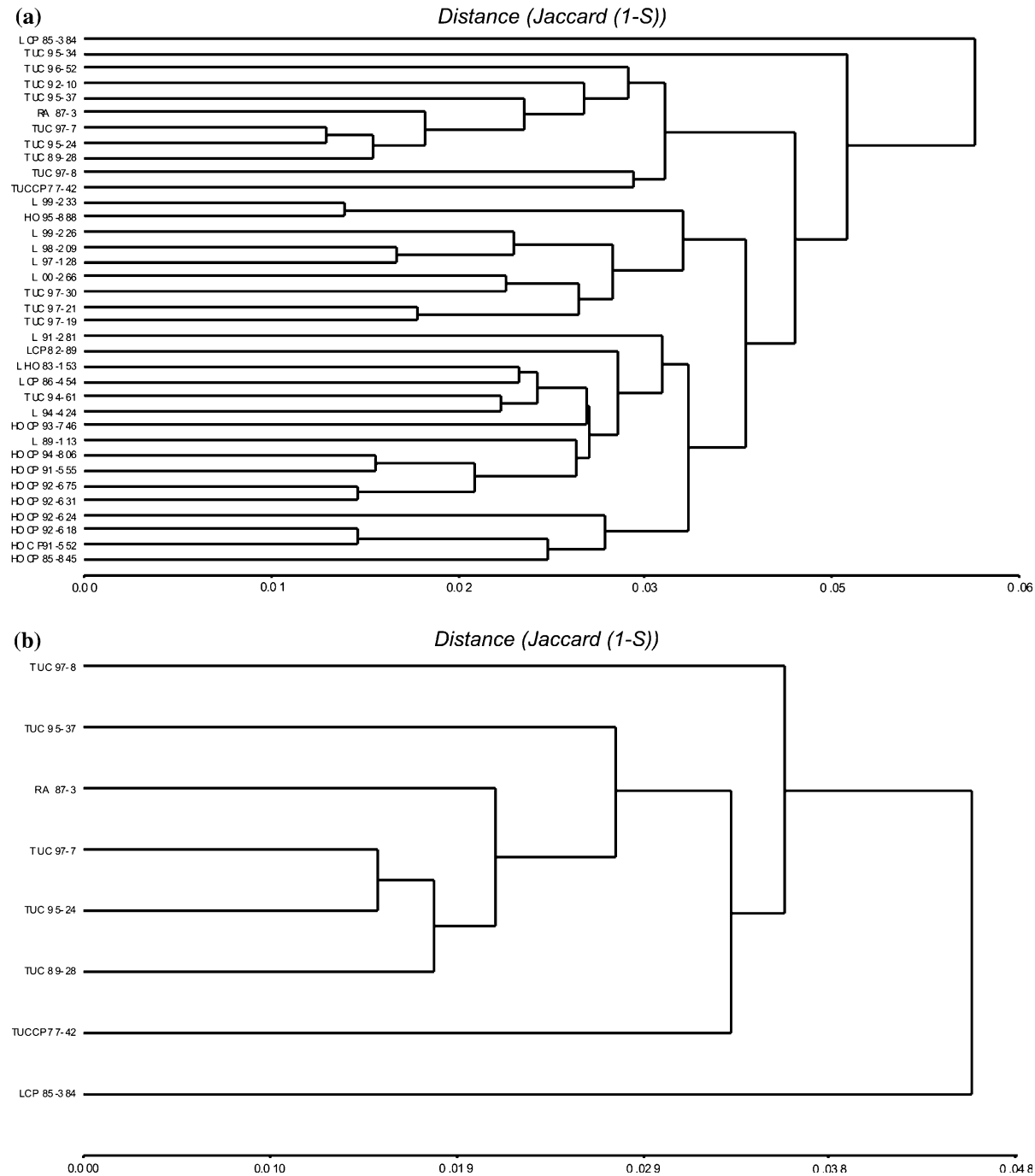

Fig. 1 Phenogram of 36 (**a**) and eight selected (**b**) sugarcane genotypes based on 995 allele analysis from 16 AFLP primer combinations when using Jaccard coefficient and UPGMA clustering method with InfoStat, presented as distance (1-S, S: similarity)

analyzed. Number and length of leaf hairs were divided to facilitate character assessment. The eight genotypes greatly varied in several characters. However, they showed similarity in four characters (depth and length of bud groove, by virtue of lacking a groove, pubescence on the bud, and size of underlapping auricle) (Table 4).

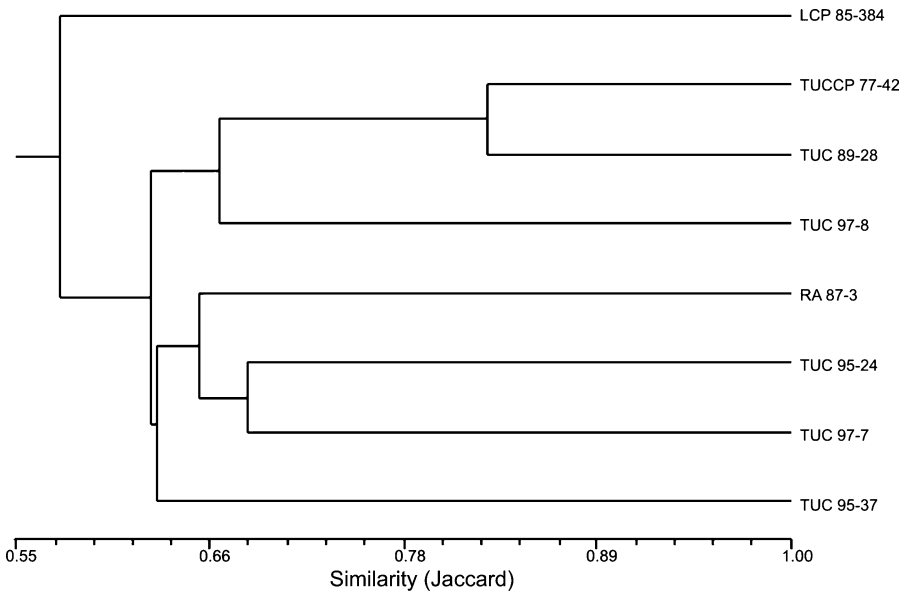

Fig. 2 Phenogram of eight selected sugarcane genotypes based on 136 allele analysis from 15 SSR primers when using Jaccard coefficient and UPGMA clustering method with NTSys shown as similarity

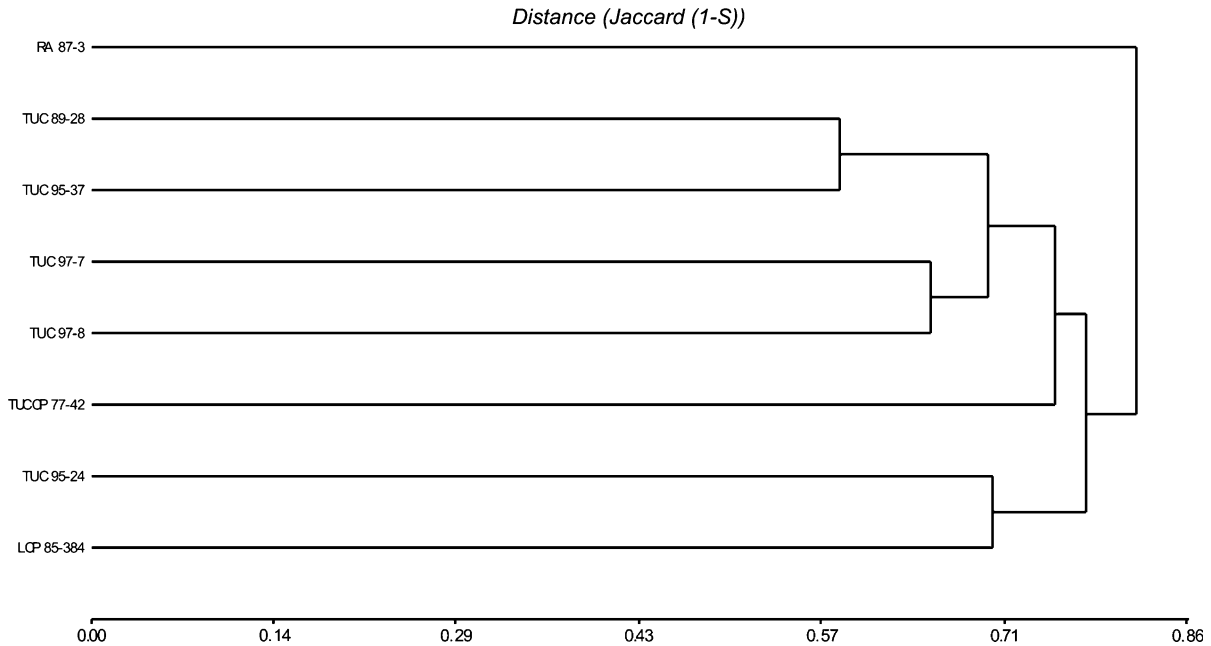

Fig. 3 Phenogram of eight selected sugarcane genotypes based on the analysis of 170 alleles produced from all evaluated morphological traits when using Jaccard coefficient and UPGMA clustering method with InfoStat program, presented as distance (1-S; S: similarity)

Dendograms obtained with both programs showed the same clusters. Genotype similarities ranged between 0.19 and 0.41 when using Jaccard coefficient (Fig. 3). This dendogram did not clearly reflect genotype pedigree; it merely revealed external genotype resemblance which is not directly associated with genetic relationships.

Nine morphological characters are compulsory to identify sugarcane varieties. Only eight were evaluated; and the dendogram constructed with both

Table 5 Characteristics of the different locations (Cuenya et al. 2011) where eight compulsory characters proposed by UPOV (2005) were evaluated

Locations (Tucumán, Argentina)	Annual precipitation (mm)			Soil		
	Normal[a]	2010	2011[b]	Texture	Organic matter	Drainage
Cevil Pozo	1,141	1,030	ND	Silty loam	Moderate	Very good
Las Talitas	1,064	1,017.6	ND	Loam	Moderate	Good
Santa Ana	1,194	940.9	ND	Sandy loam	High	Good

[a] Determined as the average of the values of the years between 1973 and 2009

[b] *ND* not determined yet

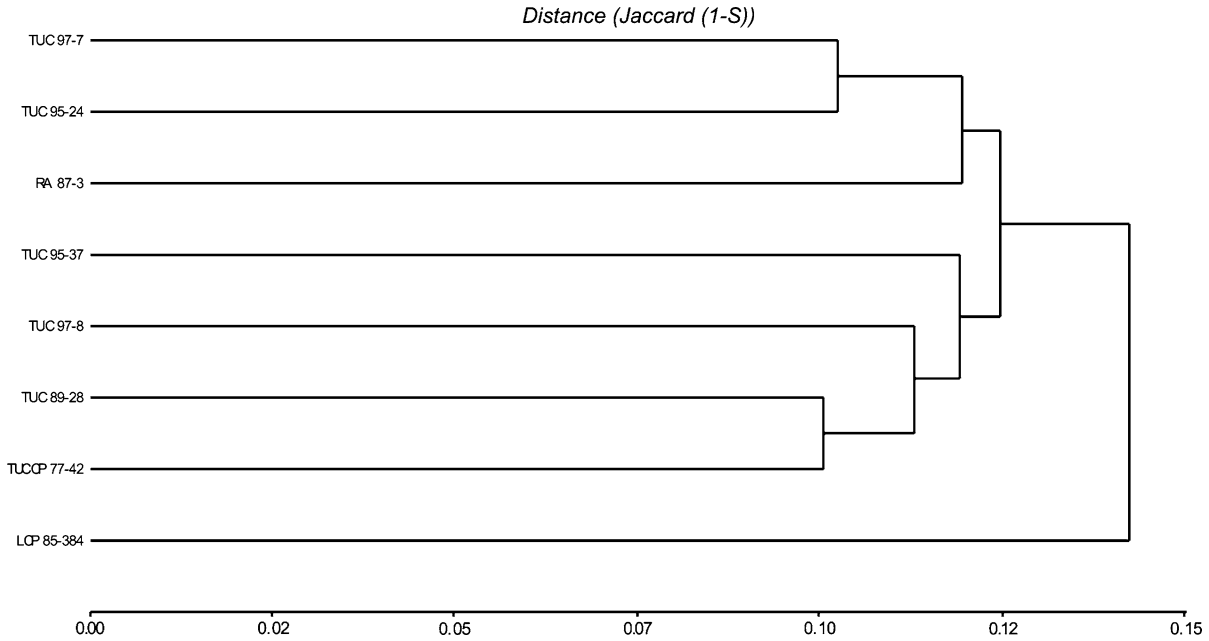

Fig. 4 Phenogram of eight selected sugarcane genotypes based on 170 allele analysis from all morphological traits, 995 alleles produced from 16 AFLP primer combinations and 136 alleles produced from 15 SSR primers when using Jaccard coefficient and UPGMA clustering method InfoStat, presented as distance (1-S; S: similarity)

softwares did not show the same clusters, because two genotype positions were exchanged (TUC 89-28 and TUC 95-37) and the similarity ranged between 0.08 and 0.24 when using Jaccard coefficient. Our results support the use of these characters to differentiate between sugarcane genotypes as they were enough to discriminate all varieties tested.

These eight compulsory characters were evaluated during two years and three different locations and actually did not vary under different environmental conditions (Table 5). This suggests that the characters proposed by UPOV (2005) enable cultivar label maintenance both within and between research stations, as they are very stable under different environmental conditions.

Morphological data were combined with AFLP and SSR data of the eight genotypes and similarity ranged between 0.89 and 0.94, and 0.43 and 0.55, respectively, when using Jaccard coefficient.

Although morphological traits only revealed external resemblance and not genetic relationships, when there were combined with molecular markers the topology of the dendogram obtained reflected genotype pedigree as when only molecular markers were used. In addition, a more accurate estimation of the genetic diversity could be obtained due to genotype

similarities ranged between 0.86 and 0.90 by using Jaccard coefficient (Fig. 4), compared with that obtained by only using molecular markers.

Correlation between AFLP and SSR dendograms of the eight genotypes was 0.58 ($P < 0.05$); this indicates that there would be a concordance between AFLP and SSR, despite the difference in topology of both dendograms. The correlations between morphological traits and AFLP and SSR were not significant and low at 0.29 and 0.11, respectively. However, when the three types of markers were combined, including the morphological ones, high and significant correlations were obtained with molecular markers (Table 6).

On the other hand, dendograms obtained with both softwares did not always show the same cluster. When the data number was high, both programs showed the same clusters; however, when genotype or character numbers were reduced, dendograms differed (for example: eight genotypes with 136 SSR markers or eight genotypes with 34 morphological markers belonging to the eight compulsory characters). Nevertheless, character numbers were more important than genotype numbers in order to obtain the same cluster with both programs. With AFLP data the minimum number of characters needed to obtain the same cluster with both softwares, was tested. At least 150 characters had to be included in the analysis to obtain the same dendograms with both programs (data not shown).

Coefficient of parentage

Coefficient of parentage (f) ranged from 0 to 0.5 with a mean of 0.026. The correlations (r) between genetic similarity coefficients obtained with AFLP, SSR and AFLP plus SSR, and f were highly significant ($P < 0.001$); although, the value obtained was relatively low. The correlation (r) between genetic similarity coefficient obtained from morphological trait and f showed also a low value but it was not significant (Table 7).

Table 6 Spearman's correlation coefficients among genetic distances obtained using AFLP, SSR and morphological traits

	Molecular marker		Morphological traits
	AFLP	SSR	
SSR	0.58[a]		
Morphological traits	0.29	0.11	
AFLP + SSR + Morphological traits	0.55[a]	0.79[a]	0.01

Significance level: $P \leq 0.05$

Table 7 Pearson correlation coefficients among Jaccard genetic similarity coefficients obtained using AFLP, SSR and morphological traits and coefficient of parentage

	Coefficient of parentage	P
AFLP	0.12	$4\ e^{-4}$ [a]
SSR	0.20	$3.7\ e^{-6}$ [a]
AFLP + SSR	0.17	$9.5\ e^{-5}$ [a]
Morphological traits	0.16	0.51

[a] Significance level: $P \leq 0.001$

Hybridity determination

This experiment was carried out in order to assure that both the hybrid character of the progeny and that the emasculation treatment routinely applied for parents in the local breeding program is reliable.

After the emasculation treatment, total pollen absence was confirmed using magnifying glass for the two varieties, RA 87-3 and LCP 85-384, whereas presence of pollen was detected in the same varieties without treatment. Crosses were performed successfully and the number of seedlings obtained/gr of seed is shown in Table 8. No viable seeds were obtained when LCP 85-384 and RA 87-3 were emasculated and selfed. Total DNA from progeny of each cross-combination was screened with the MSCIIR19 primer that amplified seven polymorphic and three monomorphic fragments between the two varieties.

A segregation analysis was conducted to detect hybrids and possible distortions, especially taking into account the polyploid structure of sugarcane. As SSR bands were considered as dominant, and sugarcane is highly heterozygous, only half the hybrid progeny (1:1) on average inherit each male or female-specific SSR band in the crossings and in 3:1 ratio in the selfings (McIntyre and Jackson 2001). For monomorphic bands in the crossings the segregation is expected to be 3:1, considering that sugarcane is highly heterozygous. Thus, it was important to use a sufficient number of male-specific bands to have a reasonable probability of detecting the hybrids. Segregation analysis showed that each marker segregated in a Mendelian fashion (as evaluated by χ^2 tests, $P \leq 0.05$) for each cross (Table 9)

Table 8 Sugarcane cross combinations studied in the emasculation experiment

	Combinations	Average number of seedlings obtained/gr of seed	Number of evaluated seedlings
Crosses	LCP 85-384 emasculated × RA 87-3	129	43
	RA 87-3 emasculated × LCP 85-384	139	44
Selfings	LCP 85-384	23	22
	RA 87-3	14	11
	LCP 85-384 emasculated	0	0
	RA 87-3 emasculated	0	0

and selfing (Table 10). Taking into consideration how markers segregated in the hybrids, out of the six markers presented in LCP85-384, four were heterozygous (markers 5, 6, 8 and 10) and two were homozygous (markers 2 and 3), because the latter two did not segregate in the hybrids when LCP 85-384 was selfed (Table 10). Out of the seven markers present in RA 87-3, five were heterozygous (markers 1, 4, 7, 8 and 9) and two were homozygous (markers 3 and 5) (Table 10). Out of the three monomorphic bands between the two varieties, two did not segregate (markers 3 and 5) because, as it was aforementioned, marker 3 was homozygous in both, LCP 85-384 and RA 87-3, while marker 5 was homozygous in the latter. The monomorphic marker 8 segregated in a 3:1 ratio as expected.

Discussion

Determining genetic diversity and evaluating different approaches to assess such diversity are equally important in order to be able to better identify genetically distant sugarcane germplasm to be used as parents in breeding programs and to generate better performing hybrids. In addition, obtaining such information will enable a more sustainable management of the available resources and provide a better basis for breeding program success (Queme et al. 2005). Moreover, determining genetic diversity of the evaluated genotypes may contribute to the development of a mapping population, which could serve as a tool for the identification and localization of QTL corresponding to yield and fiber content, traits of great future interest taking into account the great value of sugarcane as a bioenergy feedstock. Diversity can be determined by means of morphological characters and/or molecular markers; the latter present some advantages over the morphological traits because they are not affected by the environment, the plant physiological state and other factors. AFLP markers allow for simultaneous screening of many different genome regions distributed randomly throughout the genome (Mueller and Wolfenbarger 1999). In this study we found a high number of AFLP bands but revealed a low polymorphic band number (19.40%) reflecting a very limited genetic diversity in the breeding germplasm evaluated (Jaccard coefficient mean value: 0.96). Lima et al. (2002) found a higher diversity (Jaccard coefficent mean value: 0.47) when analyzing different *Saccharum* species using only polymorphic AFLP bands. A study of Besse et al. (1998) revealed a lower diversity (Dice coefficient mean value: 0.69) when using AFLP markers within *S. officinarum* and *S. Spontaneum*. However, the AFLP technique still provides a useful alternative for diversity estimation as well as for genotype identification because the 16 primer combinations were enough to differentiate between all 36 sugarcane genotypes tested in this study.

Microsatellite (SSR) markers are one of the best choices to determine genetic diversity because of their abundance, high polymorphism between individuals (Cordeiro et al. 2001) (75% between the evaluated genotypes) and reproducibility. This type of markers was useful for the genotypic identification and in establishing relationships between the different sugarcane genotypes used in this study, manifesting the pedigree information as earlier described (Cordeiro et al. 2001). In addition, using SSR markers makes it possible to determine self-pollination accurately and easily identify true hybrid progeny, even at early stages of a breeding program (Zhang et al. 2004). The polymorphism information content (PIC) term, originally introduced into human genetics (Botstein et al.

Table 9 Segregation analysis (χ^2 test, $P \leq 0.05$) for each cross combination, between two sugarcane varieties

Genotyping			LCP 85-384 emasculated × RA 87-3				RA 87-3 emasculated × LCP 85-384			
Marker	LCP 85-384	RA 87-3	Expected segregation	Observed segregation	χ^2	Probability (%)	Expected segregation	Observed segregation	χ^2	Probability (%)
1		–	1:1	19:24	0.581	44.5766 NS[a]	1:1	20:24	0.364	54.6494 NS
2	–		1:1	Without segregation			1:1	Without segregation		
3	–	–	3:1	Without segregation			3:1	Without segregation		
4		–	1:1	20:23	0.209	64.7315 NS	1:1	22:22	0.00	100.0 NS
5	–	–	3:1	Without segregation			3:1	Without segregation		
6	–		1:1	20:23	0.209	64.7315 NS	1:1	20:24	0.364	54.6494 NS
7		–	1:1	23:20	0.209	64.7315 NS	1:1	21:23	0.091	76.3025 NS
8	–	–	3:1	32:11	0.209	64.7315 NS	3:1	32:12	0.091	76.3025 NS
9		–	1:1	23:20	0.209	64.7315 NS	1:1	23:21	0.091	76.3025 NS
10	–		1:1	27:16	2.814	9.3448 NS	1:1	25:19	0.818	36.5712 NS

[a] *NS* not significant

Table 10 Segregation analysis (χ^2 test, $P \leq 0.05$) in self progeny of two sugarcane varieties

LCP 85-384 selfed						RA 87-3 selfed				
Marker	LCP 85-384	Expected segregation	Observed segregation	χ^2	Probability (%)	RA 87-3	Expected segregation	Observed segregation	χ^2	Probability (%)
1						–	3:1	6:5	2.455	11.7185 NS
2	–	3:1	Without segregation							
3	–	3:1	Without segregation			–	3:1	Without segregation		
4						–	3:1	7:4	0.758	38.4088 NS
5	–	3:1	15:7	0.545	46.0181 NS[a]	–	3:1	Without segregation		
6	–	3:1	15:7	0.545	46.0181 NS					
7						–	3:1	7:4	0.758	38.4088 NS
8	–	3:1	15:7	0.545	46.0181 NS	–	3:1	7:4	0.758	38.4088 NS
9						–	3:1	7:4	0.758	38.4088 NS
10	–	3:1	15:7	0.545	46.0181 NS					

[a] *NS* not significant

1980) refers to the value of a marker for detecting polymorphism within a population, depending on the number of detectable alleles and the distribution of their frequency (Junjian et al. 2002). PIC values calculated for the genotypes used in this study differed when using the formula introduced by Botstein et al. (1980) and the one used by Anderson et al. (1993). However, the latest version seems to be more widely used and the high PIC level obtained by this method is a reflection of how much of diversity has been captured in cultivated sugarcane, showing the usefulness of SSR for diversity studies. As Coburn et al. (2002) found in rice, there was no direct correlation between allele numbers and PIC values (Table 3); hence, PIC differences are largely due to allele frequency variation. Studies in wheat has also shown that both PIC values and diversity obtained with SSR are higher than those obtained by AFLP (Manifesto et al. 2001).

Furthermore, based in our results, molecular markers were useful not only for genetic diversity estimation but also for genotype identification. For that

reason, we suggest that molecular markers should be included in the internationally established morphological characters for protection of new sugarcane varieties as they allow for the screening of different regions of the genome, germplasm characterization and identification at the same time.

Only eight out of the 36 genotypes, chosen because they are the most significant varieties among the evaluated genotypes, were characterized by using the characters proposed by UPOV (2005). Three of them (LCP 85-384, RA 87-3 and TUCCP 77-42) are the most widely planted cultivars in the sugarcane growing area in Tucumán and they occupy approximately 70% of the total production area (Ahmed et al. 2007). Three other genotypes (TUC 89-28, TUC 95-37 and TUC 97-8) were the varieties released in 2009, while the rest (TUC 97-7 and TUC 95-24) are advanced breeding clones at final testing stages before a possible commercial release. LCP 85-384 and TUCCP 77-42 are from USA, but the second genotype was selected in Argentina, while the rest were obtained in the local breeding program at EEAOC. One of the main aims of our breeding program is to expand the varietal spectrum, at the moment restricted to only a few genotypes, so that the genetic base of future commercial varieties will be widened.

Morphological characters proposed by UPOV (2005) were more discriminative than AFLP and SSR markers as Salem et al. (2008) found. Actually, only eight characters out of those proposed by UPOV, efficiently distinguished between varieties. However, morphological characters may be inadequate to evaluate correctly the genetic relatedness between genotypes because as they only revealed external resemblance they could overestimate genetic diversity. On the other hand, all these characters were environmentally stable as they did not vary in determinations carried out during two independent growth seasons and three different geographic locations. Morphological characterization is easier, faster and more economical than molecular and other marker systems (Wagih et al. 2004) and can be used as a valuable tool to identify genotypes.

When comparing dendrograms and correlations between genetic distances, our results revealed that the morphological data were not in agreement with the molecular data obtained (Figs. 1b, 2, 3; Table 6). Similar results have been obtained when comparing DNA markers and morphological traits in many other species (García et al. 2002; Durán et al. 2005; Martinez et al. 2005; Vollmann et al. 2005). As a consequence it has been suggested that the low correlation between DNA markers and morphological data could be due to DNA markers cover a larger proportion of the genome, including coding and noncoding regions, than the morphological markers (Semagn 2002). Other explanation for such discrepancy is that the DNA regions examined by molecular techniques are not expressed in phenotypic character proposed by UPOV, although these characters are perhaps more reflective of agronomic performances. Moreover, the SSR dendrogram obtained is topologically different from the AFLP dendogram (Fig. 1b vs 2). However, the correlation value indicates that there is a concordance between the AFLP and SSR data (Table 6). As sugarcane is a cross-pollinated, highly heterogeneous species, a high level of similarity among different marker techniques might not be seen, given that they explore different regions of the large genome. In accordance with this hypothesis, Degani et al. (2001) and Budak et al. (2004) reported correlation absence between different markers in strawberry and buffalograss, respectively.

NTSys and InfoStat softwares were analysed in order to test their robustness. These inconsistencies may be attributed to differences in rounding distance matrixes. When providing a high number of data points, both programs show the same clusters. However, when genotypes or character numbers are reduced, dendograms began to differ. Character numbers were found to be of more importance than the genotype number in order to obtain the same cluster using both programs. We found that at least 150 characters had to be included in the analysis to obtain the same cluster using both softwares. Our study indicates that NTSys is better to use when a reduced data number is available, as dendograms with fewer data were more consistent with pedigree expectations using this software (data not shown). However, InfoStat is more simple and user friendlier. Regards similarity coefficients, Jaccard and Dice coefficients assume that two genotypes are more similar due to the presence instead of the absence of a character (Balzarini et al. 2006). However, as Dice gives more weight to matching bands (Besse et al. 1998), Jaccard reduces the risk of over-estimating similarity as Dice does (Cordeiro et al. 2003). Both disregards the conjoint absence of bands in the pairwise comparison and give identical rankings among pair from inbred lines (Mohammadi

and Prasanna 2003) but might differ when analyzing heterozygous loci (Link et al. 1995). For this study, the Jaccard coefficient was used as it is deemed most appropriate when using dominant markers (Arro 2004). In addition, the large number of fragments revealed on an AFLP gel, as in the case of our study, maximises the chances of fortuitous comigration of two fragments of very similar size (Besse et al. 1998); Jaccard coefficient minimises the error resulting from scoring different bands as identical. For the aforementioned reasons, the Jaccard coefficient was chosen to calculate genotype similarity in our analysis.

A significant correlation, with low value, was observed between the coefficient of parentage (*f*) and measurements of genetic similarity ($r = 0.12$ and 0.20 for AFLP and SSR, respectively). Similar results were also observed in other studies where genetic similarity, based on molecular markers, has shown low to medium correlation with *f* based on pedigree data (Graner et al. 1994; Tinker et al. 1993; Plaschke et al. 1995; Schut et al. 1997; Barret and Kidwell 1998; Lima et al. 2002). The low to moderate correlation between genetic similarity estimates and *f* is probably due to: unequal parental contribution (Lu et al. 1994; Deren 1995); gene concentration in each generation of crossing; the male parent in the polycrosses being the same or a very close cultivar, or else the mistaken annotation of the parent during the generation of one of the two estimates (Lima et al. 2002). In addition, the pedigree records do not take into account selection and genetic drift, which play a significant role in variety development (Selvi et al. 2003). The *f*, although highly informative in a breeding program, presents inherent errors during its calculation. This is, in part, due to some genetic suppositions, which are assumed in the calculation of *f*, such as that all ancestors are not closely related (Cox et al. 1985) which is not always true when the history of sugarcane cultivars is considered (Bremer 1961a). The supposition that the genotype receives the same amount of genes from each parent is questionable; sugarcane is polyploid and highly heterozygotic, even two full sibs could inherit completely different sets of chromosomes and thus markers, from the same pair of parents, eventually resulting in a near zero genetic similarity (Selvi et al. 2003). Adding to this, the well-known fact that when using *S. officinarum* as a female parent, its meiosis is not equivalent, resulting in one parent's alleles having an advantage over the other (Bremer 1961b). Another point to be considered during the calculation of *f* is the little-known changes in the frequency of the alleles, due to the effect of genetic drift and the selection process. Both phenomena can influence the precision of *f*. As, for example, the transmission of alleles, especially those that control qualitative characteristics with high heritability, is clearly influenced by the intensity of selection in a breeding program. This fact results in a bias in the contribution of the parent, stressing the favourable alleles for the character in the resulting progeny (Cox et al. 1985; Souza and Sorrells 1989). So, genetic similarity estimates obtained from molecular markers will provide more accurate information to plant breeders than the pedigree method, allowing them to make more-efficiently reliable crossings with maximum-allele variation on a short-term basis or to strategically plan the breeding program on a more long-term basis as Barret and Kidwell (1998) and Lima et al. (2002) found. The correlation between genetic similarity coefficient obtained from morphological trait and *f* showed a low value that was not statistically significant. Schut et al. (1997) have obtained similar results in barley and this may be caused by a biased and insufficient representation of the whole genome using only morphological traits.

The EEAOC breeding program constantly incorporates new germplasm, predominantly from the USA. Our genotyping results indicate that local and USA breeding programs share a common genetic background as CP, LCP and HoCP varieties all grouped together with the local genotypes, since most of them are products of crosses shared among these programs (Fig. 1). The data provided in this study support the evidence that the local sugarcane gene pool was mainly derived from introductions from the USA. No clear genetic differentiation between these programs could be detected due to the high frequency of germplasm exchange. The genetic base of modern sugarcane varieties appears to be narrow as reported by Cordeiro (2001) and D'Hont et al. (1995). This could be the reason for the present slow progress in sugarcane breeding and the little genetic diversity.

Although McIntre and Jackson (2001) suggested that unwanted selfing is not a significant problem in the Australian sugarcane breeding program; self-pollination can occur in sugarcane. However, in the local breeding program at EEAOC, the implementation of both an effective emasculation method and the SSR technique allows to assess the fidelity of

sugarcane crosses. Results from hybridity determination experiments indicate that treatment is successful in complete emasculating of both the RA 87-3 and LCP 85-384 varieties and that it does not cause a serious reduction in stigma and ovary viability, as has been previously reported (Nagai 1984). In addition, this method is simpler, faster and cheaper than other emasculation techniques (Machado et al. 1995). The SSR technique allowed us to identify true hybrid progeny and, although only a limited number of SSR markers have been studied, this survey suggests that these markers showing Mendelian inheritance in the polyploid nature of sugarcane were enough to allow for hybrid detection. Manigbas and Villegas (2004) correctly identified hybrids derived from sugarcane crosses using only one SSR. In conclusion both implemented tools routinely applied in the EEAOC breeding program, will improve the efficiency of this process.

Despite economic and social importance of sugarcane production in Argentina, no other study has been carried out to evaluate integrally local breeding materials. The present study constitutes a detailed report on sugarcane genotype characterization for identification and genetic diversity estimation by comparing data analysis softwares and by using morphological traits and different molecular techniques in order to monitor different genomic regions. It represents a starting point to determine qualitative and quantitative shifts in diversity over time of the breeding program at EEAOC.

Sugarcane should be considered as a model crop of high genetic complexity. Many times, only a small proportion of the genotypic variation is sampled, hence, the inclusion of many kinds of characters would be of great addition. Thus, when morphological traits were included in the analysis, although they only revealed external resemblance, the dendrogram obtained reflected genotype pedigree and its topology was not modified compared with those obtained with molecular markers (Table 6); however, both methods, molecular and morphological, should be used together in order to obtain a more accurate estimation of the genetic diversity. In summary, the combined genetic and morphologic variability found in this study will aid in selection of parental crosses for mapping and for the discovery of potential quantitative trait loci. The benefits of the genetic knowledge obtained will be extremely useful not only to identify parental lines for favourable combinations that may yield greater genetic gain in breeding for important traits but also to broaden the general genetic basis of the local germplasm in the future.

Acknowledgments We thank A. D′Hont (CIRAD, France) for providing us with MSCIIR primer sequences, S. Olmos (INTA) for technical support, L. Montivero and Bjorn Welin for reviewing the manuscript English version and Estación Experimental Agroindustrial Obispo Colombres (EEAOC), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and Universidad Nacional de Tucumán (UNT) for their generous financial support. APC is CONICET member; MFP and JR are CONICET fellows.

References

Ahmed MA, Cuenya MI, García MB, Díaz Romero C (2007) Dinámica y características del recambio de variedades de caña de azúcar en tucumán desde las primeras décadas del siglo xx hasta la actualidad. Avance Agroindustrial 28: 17–22

Aljanabi SM, Forget L, Dookun A (1999) An improved and rapid protocol for the isolation of polysaccharide and polyphenol-free sugarcane DNA. Plant Mol Biol Rep 17:1–8

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36(1):181–186

Angiolillo AM, Mencuccini M, Baldoni L (1999) Olive genetic diversity assessed using amplified fragment length polymorphisms. TAG 98:411–421

Arro JA (2004) Genetic diversity among sugarcane clones using target region amplification polymorphism (TRAP) markers and pedigree relationships. Master of Science Thesis, Agronomy & Environmental Management

Balzarini M, Di Rienzo J (2003) Info-gen: software para análisis estadístico de datos genéticos. Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina

Balzarini M, Arroyo A, Bruno C, Di Renzo J (2006) Análisis de datos de marcadores con Info-gen. XXXV Congreso Argentino de Genética, San Luis

Barret BA, Kidwell KK (1998) Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. Crop Sci 38: 1271–1278

Besse PG, Taylor B, Carroll N, Berding D, Burner CL (1998) Assessing genetic diversity in a sugarcane germplasm collection using an automated AFLP analysis. Genetica 104: 143–153

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32(3): 314–331

Breaux RD, Legendre BL (1983) The USDA commercial breeding program in Louisiana. Proc Inter Am Sugar Cane Sem Var Breed III:99–105

Bremer G (1961a) Problems in breeding and cytology of sugarcane. Euphytica 10:59–78

Bremer G (1961b) Cytological crossing research of sugar cane. Euphytica 10:229–243

Budak H, Shearman RC, Parmaksiz I, Gaussoin RE, Riordan TP, Dweikat I (2004) Molecular characterization of buffalograss germplasm using sequence-related amplified polymorphism markers. TAG 108(2):328–334. doi:10.1007/s00122-003-1428-4

Caetano-Anollés G, Gresshoff PM (1994) Staining nucleic acids with silver: an alternative to radioisotopic and fluorescent labeling. Promega Notes Mag 45:13

Chang YS, Lo CC (1993) Genetic relationship among major sugar cane cultivars in Taiwan. Rep Taiwan Sugar Res Inst 140:1–10

Coburn JR, Temnykh SV, Paul EM, McCouch SR (2002) Design and application of microsatellite marker panels for semiautomated genotyping of rice (*Oryza sativa* L.). Crop Sci 42:2092–2099

Cordeiro GM (2001) Molecular marker systems for sugarcane germplams analysis, Chap. 9. In: Henry RJ (ed) Plant genotyping, the DNA fingerprinting of plants. Cabi publishing, Wallingford

Cordeiro GM, Taylor GO, Henry RJ (2000) Characterisation of microsatellite markers from sugarcane (*Saccharum* sp.), a highly polyploid species. Plant Sci 155(2):161–168

Cordeiro GM, Casu R, McIntyre CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci 160(6):1115–1123

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

Cox TS, Kiang YT, Gorman MB, Rogers DM (1985) Relationships between coefficient of parentage and genetic similarity indices in the soybean. Crop Sci 25:529–532

Cuenya MI, Ostengo S, Chavanne ER, Espinosa MA, Costilla DD, Ahmed MA (2009) Relevamiento de la distribución de variedades comerciales y de la aplicación de otras tecnologías en el área de cultivo de la provincia de Tucumán: Campaña 2007–2008. Gacetilla Agroindustrial EEAOC 72

Cuenya MI, Chavanne ER, Ostengo S, García MB, Ahmed M, Costilla DD, Díaz Romero C, Espinosa MA, Delgado N, Díaz JV (2011) TUC 95-10: una nueva variedad de caña de azúcar altamente productiva. Gacetilla Agroindustrial EEAOC 75

Da Silva JA, Bressiani JA (2005) Sucrose synthase molecular marker associated with sugar content in elite sugarcane progeny. Genet Mol Biol 28(2):294–298. doi:10.1590/S1415-47572005000200020

Degani C, Rowland LJ, Saunders JA, Hokanson SC, Ogden E, Golan-Goldrish A, Galetta GL (2001) A comparison of genetic relationship measures in strawberry (*Fragaria ananassa* duch.) based on AFLP, RAPDs and pedigree data. Euphytica 117:1–12

Demey JR, Zambrano AY, Fuenmayor F, Segovia V (2003) Relación entre caracterizaciones molecular y morfológica en una colección de yuca. Interciencia 28(12):684–689

Deren CW (1995) Genetic base of U.S. mainland sugar cane. Crop Sci 35:1195–1199

D'Hont A, Rao PS, Feldman P, Grivet LN, Islam-Faridi N, Taylo P, Glaszmann JC (1995) Identification and characterization of sugarcane intergeneric hybrids, *Saccharum officinarum* × *Eriantus arudinaceus*, with molecular markers and DNA in situ hybridisation. TAG 91:320–326

D'Hont A, Seshagiri Rao P, Alleyne S, Glaszmann JC, Feldmann P (1997) Application of molecular markers in sugarcane breeding. Proceedings West Indies sugar technologists, pp 77–79

Di Rienzo JA, Casanoves F, Balzarini MG, Gonzalez L, Tablada M, Robledo CW (2010) InfoStat versión 2009. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina

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

Durán LA, Blair MW, Giraldo MC, Macchiavelli R, Prophete E, Nin JC, Beaver JS (2005) Morphological and molecular characterization of common bean landraces and cultivars from the Caribbean. Crop Sci 45:1320–1328

Gallacher DJ (1997) Optimised descriptors recommended for Australian sugarcane germplasm (*Saccharum* spp. Hybrid). Aust J Agric Res 48(6):775–779

García MG, Ontivero M, Díaz Ricci JC, Castagnaro AP (2002) Morphological traits and high resolution rapd markers for the identification of the main strawberry varieties cultivated in Argentina. Plant Breed 121:76–80

Gillet EM (1991) Genetic analysis of nuclear DNA restriction fragment patterns. Genome 34(5):693–703

Graner A, Ludwig WF, Melchinger AE (1994) Relationships among European barley germplasm. II. Comparisons of RFLP and pedigree data. Crop Sci 34:1199–1205

Heinz DJ (1987) Sugarcane improvement through breeding. Elsevier, Amsterdam 603 pp

Hill MH, Witsenboer M, Zabeau P, Vos R, Kesseli R, Michelmore R (1996) PCR-based fingerprinting using AFLPs as a tool for studying genetic relationships in *Lactuca* spp. TAG 93:1202–1210

Junjian N, Colowit PM, Mackill DJ (2002) Evaluation of genetic diversity in rice subspecies using microsatellite markers. Crop Sci 42:601–607

Kempthorne O (1957) An introduction to genetic statistics. Wiley, New York

Lima ML, Garcia AA, Oliveira KM, Matsuoka S, Arizono H, De Souza CL Jr, De Souza AP (2002) Analysis of genetic similarity detected by aflp and coefficient of parentage among genotypes of sugar cane (*Saccharum* spp.). TAG 104(1):30–38

Link W, Dixkens C, Singh M (1995) Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers. TAG 90:27–32

Lu YH, D′Hont A, Paulet F, Grivet L, Arnaud M, Glaszmann JC (1994) Molecular diversity and genome structure in modern sugarcane varieties. Euphytica 78:217–226

Machado JR, Walker DI, Bressiani JA, Goncalves A, da Silva J (1995) Emasculation of sugarcane tassels using hot water. Proc ISSCT 22:346–348

Manifesto MM, Schlatter AR, Hopp HE, Suarez EY, Dubcovsky J (2001) Quantitative evaluation of genetic diversity in wheat germplasm using molecular markers. Crop Sci 41:682–690

Manigbas NL, Villegas LC (2004) Microsatellite markers in hybridity tests to identify true hybrids of sugarcane. Phil Crop Sci 29:23–32

Martínez L, Cavagnaro P, Masuelli R, Rodriguez J (2005) Evaluation of diversity among Argentine grapevine (*Vitis*

vinifera L.) varieties using morphological data and AFLP markers. Elect J Biotechnol 6:37–45
Maughan PJ, Saghai MA, Maroof GR, Buss GR, Huestis GM (1996) Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. TAG 93:392–401
McIntyre CL, Jackson PA (2001) Low level of selfing found in a sample of crosses in Australian sugarcane breeding programs. Euphytica 117:245–249
Mohammadi SA, Prasanna BM (2003) Analysis of genetic diversity in crop plants—salient statistical tools and considerations. Crop Sci 43:1235–1248
Mueller UG, Wolfenbarger LL (1999) AFLP genotyping and fingerprinting. Trends Ecol Evol 14:384–394
Nagai C (1984) Emasculation of sugarcane tassels by hot water. Am Rept HSPA 3-4
Paul S, Wachira FN, Powell W, Waugh R (1997) Diversity and genetic differentiation among populations of India and Kenyan tea [*Camellia sinensis* (L) O. Kuntze] revealed by AFLP-markers. TAG 94:255–263
Plaschke J, Ganal MW, Roder MS (1995) Detection of genetic diversity in closely related bread wheat using microsatellite markers. TAG 91:1001–1007
Powell W, Machray G, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1:215–222
Queme JL, Molina L, Melgar M (2005) Analysis of genetic similarity among 48 sugarcane varieties using microsatellite DNA sequences. Proc ISSCT 25:592–595
Rohlf FJ (1993) Ntsys-pc. Version 2.0. Applied biostatistics inc., Exeter Software, New York
Salem KFM, El-Zanaty AM, Esmail RM (2008) Assessing wheat (*Triticum aestivum* L.) genetic diversity using morphological characters and microsatallite markers. World J Agric Sci 4(5):538–544
SAS Institute (2003) SAS/STAT release 9.1. SAS Inst., Cary, NC
Schut JW, Qi X, Stam P (1997) Association between relationship measures based on AFLP markers, pedigree data and morphological traits in barley. Theor Appl Genet 95:1161–1168
Selvi A, Nair NV, Balasundaram N, Mohapatra T (2003) Evaluation of maize microsatellite markers for genetic diversity analysis and fingerprinting in sugarcane. Genome 46:394–403
Semagn K (2002) Genetic relationships among ten endod types as revealed by a combination of morphological, RAPD and AFLP markers. Hereditas 137:149–156
Smith S, Luk S, Sobral B, Muhawish S, Peleman J, Zabeau M (1994) Associations among inbred lines of maize using rflp and DNA amplification technologies (AFLP and AP-PCR), and correlations with pedigree, f1 yield and heterosis. Maize Genet Newslett 68:45
Sneath PHA, Sokal RR (1973) Numerical taxonomy. W. Freeman, San francisco 573 pp
Souza E, Sorrells ME (1989) Pedigree analysis of North American oat cultivars released from 1951 to 1985. Crop Sci 29:595–601
Tinker NA, Fortin MG, Mather DE (1993) Random amplified polymorphic DNA and pedigree relationships in spring barley. Theor Appl Genet 85:976–984
Union for the Protection of New Varieties of Plants (UPOV) (2005) Directrices para la ejecución del examen de la distinción, la homogeneidad y la estabilidad tg/186/1. www.Upov.Int/es/publications/tg-rom/tg186/tg_186_1.Pdf
Vettore AL, da Silva FR, Kemper EL, Arruda P (2001) The Libraries that made SUCEST. Genet Mol Biol 24:1–7
Vollmann J, Grausgruber H, Stift G, Dryzhyruk V, Lelley T (2005) Genetic diversity in camelina germplasm as revealed by seed quality characteristics and RAPD polymorphism. Plant Breed 124:446–453
Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23(21):4407–4414
Wagih ME, Musa Y, Ala A (2004) Fundamental botanical and agronomical characterization of sugarcane cultivars for clonal identification and monitoring genetic variations. Sugar Tech 6(3):127–140
Zhang MQ, Zheng XF, Yu AL, Xu JS, Zheng H (2004) Molecular marker application in sugarcane. Sugarcane Biotechnol 6(4):251–254