Agro-morphological description, genetic diversity and population structure of sugarcane varieties from sub-tropical India

Archana Siraree^{1,2,3} · Nandita Banerjee¹ · Sanjeev Kumar¹ · M. S. Khan¹ · P. K. Singh¹ · Sanjeev Kumar¹ · Swati Sharma³ · R. K. Singh⁴ · Jyotsnendra Singh¹

Received: 19 June 2018 / Accepted: 16 October 2018 / Published online: 2 November 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Genetic diversity in 92 sugarcane varieties of sub-tropical India was assessed using 30 morphological descriptors and 643 simple sequence repeat (SSR) marker loci. Out of the 30 morphological descriptors, 14 were found polymorphic, and significant variability was recorded for plant height, cane diameter and number of millable canes. Grouping traits like plant growth habit, leaf blade curvature and leaf sheath adherence were found to be predominantly monomorphic. There were a few pairs of varieties (e.g., CoP 9702 and CoP 9302, CoP 9301 and CoSe 01424, UP 05 and Co 1336, CoS 96258 and CoH 110) that showed similar DUS profiles except differing for a few descriptors. The STRUCTURE profile suggest that all the 92 sugarcane varieties had admixtures and no sub-group had a pure unblemished structure profile. An average Nei's genetic distance of 0.49 was found to be a better measure of diversity, whereas, the average band informativeness (Ib_{av}) value of all the 80 SSR primers was 0.434. Although, the mean Ib_{av} values for EST-SSR and genomic-SSR primers were same (0.43), the range of Ib_{av} of EST-SSR (0.04–0.85) was more compared to genomic-SSR (0.12–0.63) primers. The segregation of the varieties based on morphological traits was not in accordance with their geographical distribution or maturity groups, but principal component analysis was able to group the sugarcane varieties that had similar pedigree together. Results indicate that the SSRs have a potential use in the DNA fingerprinting of varieties to prevent any malpractice like unauthorised re-registration of a previously registered sugarcane variety under PPV&FR Act. The marker profiles could also be utilised for variety identification and release, since at present, it has been made mandatory to include it in addition to the morphological descriptors.

Keywords Sugarcane · Simple sequence repeats · Genetic diversity · DUS descriptors

Introduction

Sugarcane (*Saccharum* species hybrid) is one of the most productive cultivated plant system in the tropical and sub-tropical regions of more than 100 countries across the globe (daCosta et al. 2011). Apart from providing raw material for economically important white sugar, it is cultivated as a multiproduct crop providing green top as fodder, jaggery, bagasse, press mud, and paper pulp, etc. (Allen et al. 1997; Pinto et al. 2010). Recently, it has gained attention due to its potential of producing green fuel (ethanol), and in fact, in Brazil and USA, a major part of the sugarcane produced is now being utilized for such purpose (Dias et al. 2012). During the year 2015–2016, India with a total sugarcane production of 352.1 MT from cultivated area of 5 Mha ranked second only to Brazil (Anonymous 2016). Sugarcane genetic resources are classified in to traditional cultivars

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1481-y>) contains supplementary material, which is available to authorized users.

✉ Jyotsnendra Singh
jyotsnendra1@rediffmail.com

¹ Division of Crop Improvement, ICAR-Indian Institute of Sugarcane Research, Raibareli Road, Lucknow, UP 226 002, India

² Division of Seed Production, Genda Singh Sugarcane Breeding and Research Institute, Seorahi, Kushinagar, UP 274 407, India

³ Department of Biosciences, Integral University, Lucknow, UP 226 026, India

⁴ Division of Crop Science, Indian Council of Agricultural Research, Krishi Bhawan, New Delhi 110 001, India

(*Saccharum officinarum*, *S. barberi*, *S. sinense* and *S. edule*), wild relatives (*S. spontaneum* and *S. robustum*), and modern cultivars [interspecific hybrids between *S. officinarum* ($2n=80$; $x=10$) and *S. spontaneum* ($2n=40-128$; $x=8$)]. Modern cultivated varieties of sugarcane possess a complex highly polyploid genome and exhibit a high level of heterozygosity with varied number (100–130) of chromosomes (D'Hont et al. 1998; Casu et al. 2005). These complexities make varietal development a long process of 12–15 years (Hogarth 1987; Casu et al. 2005). The plant breeders are in relentless search for novel genetic combinations of superior traits, like quality, yield, and tolerance to various biotic and abiotic stresses. Therefore, it is imperative to assess the extent of genetic variability available in the population and their response towards selection. So far, the morphological features of sugarcane varieties has been the basis of their characterization. However, the continuous use of promising and high yielding interspecific hybrids in crossing, followed by selection based on few desirable traits (cane and sugar yield), and strictly vegetative means of commercial propagation has narrowed the genetic base of modern sugarcane cultivars (Walker 1987). Moreover, varietal identification in sugarcane, solely based on morphological descriptors has become a cumbersome process, as there are many varieties with overlapping traits in the absence of discrete classes for some characters. Further, being an open-pollinated crop, progenies resulting from a cross might be a mixture of self and hybrids, and therefore, selection of superior hybrids among such progenies based on visual observations may not be reliable (Heinz and Tew 1987; Tai 1989). Previous studies do emphasize about use of molecular markers in sugarcane breeding programmes so as to make varietal identification and selection process more effective and unambiguous (Pan 2006; Maccheroni et al. 2009; Oliveira et al. 2009).

In the present scenario of plant variety protection, newly bred variety has to undergo registration process so that it can be commercially used by other stakeholders, and for which DUS (Distinctiveness, Uniformity, and Stability) test is a pre-requisite. The present system of varietal registration in India allows only morphological descriptors to accomplish DUS test which poses a limitation as the number of existing varieties is quite large making the traditional method of DUS testing time-consuming and resource expensive (Pourabed et al. 2015). In addition, the subjective nature of field tests often creates a chance of human error and discretion (Cooke 1995). In view of several studies advocating the use of molecular markers for DUS test and diversity studies, UPOV constituted a working group on biochemical and molecular techniques (UPOV-BMT 2002). Molecular markers would be beneficial for DUS test in terms of markers directly linked to the descriptors, developing molecular profile of a reference variety, and as an alternative to morphological DUS descriptors (Noli et al. 2008). Microsatellites or simple

sequence repeats (SSRs) have widely been recognized as powerful and informative genetic markers in both animal and plant systems (Jarne and Lagoda 1996), including sugarcane (Xu et al. 1999; Aitken et al. 2005; Cai et al. 2005; Andru et al. 2011; Singh et al. 2011; Nayak et al. 2014; Zawedde et al. 2014). Molecular techniques like RFLP and RAPD have also been reported as methods to identify sugarcane varieties but with major limitation to process more number of samples in the case of RFLPs, whereas, reproducibility has been a key constraint with RAPD markers (Breyne et al. 1997). Sugarcane varieties have previously been characterized on the basis of distinctness, uniformity and stability (DUS) descriptors as per the guidelines of Protection of Plant Varieties and Farmer's Right Act (Anonymous 2009). Genetic diversity has been analysed in sugarcane varieties using RAPD (Silva et al. 2008; Nawaz et al. 2010; Ramakrishnan et al. 2015), RFLP (Creste et al. 2010), AFLP (Aitken et al. 2006; Besse et al. 1998), SRAP (Alwala et al. 2008; Chang et al. 2012), SCoT (Que et al. 2014), CISP (Khan et al. 2011), and EST-SSR (Silva et al. 2012; Singh et al. 2013.; Khan et al. 2018). In this study, an attempt was made to assess the efficiency of SSR markers in genetic diversity analysis, population stratification in 92 sugarcane varieties and exploit them as an additional tool for determining the distinctiveness of sugarcane varieties.

Materials and methods

Phenotyping

A panel of 92 sugarcane varieties was maintained at ICAR-Indian Institute of Sugarcane Research, Lucknow (Supplementary Table S1), and all the varieties were phenotyped for 30 (27 DUS and 3 additional traits) descriptors for three consecutive years (2011–2012, 2012–2013, 2013–2014) in a randomized block design (RBD) with three replications. Out of all the traits evaluated, 22 were qualitative and 8 were quantitative (Table 1). Three-bud sett planting was done in a six-meter row spaced at 90 cm, and observations were recorded at full maturity stage (~ 300 days after sett planting) on 30 randomly selected stools per replication as per the guidelines prescribed for DUS test of sugarcane (Anonymous 2009). The interpretation of 27 morphological characters and their different states of expression was performed visually to assess distinctiveness. The phenotypic data of 92 varieties was converted to the respective DUS test scores according to PPV&FRA guidelines (Anonymous 2009) and used for further analysis. The same DUS test scores of the 92 varieties for all the 27 DUS descriptors were used for calculating squared Euclidian distance matrix using SPSS ver. 20.0 (IBM Corp. Released 2011). Further, neighbour-joining

Table 1 The morphological DUS descriptors and some quantitative traits used for phenotyping of 92 sugarcane varieties

Sl. no.	DUS descriptor	Sl. no.	DUS descriptor
<i>Qualitative traits</i>		16	Node: bud cushion
1	Plant: growth habit	17	Node: bud tip in relation to growth ring
2	Leaf sheath: hairiness	18	Node: prominence of growth ring
3	Leaf sheath: shape of ligule	19	Node: width of root band
4	Leaf sheath: shape of inner auricle	20	Internode: cross-section
5	Leaf blade: curvature	21	Internode: pithiness
6	Plant: adherence of leaf sheath	22	Dewlap colour
7	Internode: colour (not exposed to sun)	<i>Quantitative traits</i>	
8	Internode: colour (exposed to sun)	23	Node: size of bud
9	Internode: shape	24	Plant: number of millable canes (NMC) per stool
10	Internode: zig-zag alignment	25	Plant: cane height
11	Internode: growth crack (split)	26	Leaf blade: width
12	Internode: rind surface appearance	27	Internode: diameter
13	Internode: waxiness	28	Fibre percent
14	Node: shape of bud	29	Sucrose percent
15	Node: bud groove	30	Purity percent

(NJ) dendrogram was constructed from distance matrix using the web server T-REX (Boc and Makarenkov 2012).

Genotyping using SSR markers

Genomic DNA was isolated from the young leaves of each variety using Cetyl trimethyl ammonium bromide (CTAB) method as described by Doyle and Doyle (1990). The quality and quantity of genomic DNA was checked, and DNA was diluted to a concentration of 25 ng/μl. Genotyping was done using 80 SSR primers (Supplementary Table S2) that included 16 genomic-SSRs (Govindaraj et al. 2005; Parida et al. 2009), and 64 EST-SSRs (Pinto et al. 2004; Oliveira et al. 2009; Singh et al. 2013). The DNA amplification was carried out in 15 μl reaction volume consisting of 1 × PCR assay buffer, 200 mM of each dNTPs (Fermentas, USA), 12 ng (1.8 pmol) each of forward and reverse primers (Operon Biotechnologies, GmbH, Germany), 0.5 unit of *Taq* DNA polymerase (Fermentas, USA) and 25 ng genomic DNA using a thermal cycler (MyCycler, Biorad, USA). The cycling profile was kept as follows: initial denaturation at 94 °C for 5 min, followed by 33 cycles of 94 °C for 1 min, 55–62 °C (adjusted as per T_m value of the primer) for 1 min, 72 °C for 2 min, and finally, a primer extension cycle of 7 min at 72 °C. The amplification products were resolved on 12% polyacrylamide gel with 1 × Tris-borate EDTA buffer, stained with 0.5 μg/ml ethidium bromide (Sigma-Aldrich, USA) and visualized using a gel documentation system (G:Box, Syngene, UK). The clear and unambiguous amplified products were scored as dominant markers with 1 (present) or 0 (absent) since it is difficult to identify and differentiate between amplicons from homeologous chromosomes.

Estimation of population structure and grouping pattern

Genetic structure of the 92 sugarcane varieties were determined by following Bayesian clustering analysis using the programme STRUCTURE ver. 2.3.4 (Pritchard et al. 2000). The number of groups (K) within the population was estimated using posterior probability ad hoc statistics according to Evanno et al. (2005). In addition, the grouping pattern of the 92 sugarcane varieties was estimated by principal component analysis (PCA) in software XLSTAT (XLSTAT 2017) on the basis of correlation coefficient between two genotypes using EIGEN procedure. The missing data was estimated using the nearest neighbour method in XLSTAT.

Assessment of genetic diversity

For estimating the genetic relationship among the sugarcane varieties, all polymorphic and monomorphic bands were scored either as present (1) or absent (0) for each SSR loci and the data were converted in to a binary matrix which was then used for the estimation of similarity matrix. Nei's genetic distance matrix (Nei 1972) was calculated from the genotypic data for all the varieties using software TASSEL (Bradbury et al. 2007), and neighbour-joining (NJ) dendrogram was constructed from distance matrix using the web server T-REX (Boc and Makarenkov 2012). A correlation between the genotypic and phenotypic dendrograms was assessed using Mantel's test in the software PASSaGE ver. 2.0 (Rosenberg and Anderson 2011). The amplification data was used to calculate the average band informativeness ($I_{b_{av}}$; Milbourne et al. 1997) to estimate the polymorphism

detection power of each EST- and g-SSR primer using the following formula:

$$Ib_{av} = 1/n \sum (1 - (2|0.5 - p_i|)),$$

where p_i denotes the proportion of genotypes containing the i th band and n is the total number of bands amplified by the primer.

Results

Agro-morphological diversity and clustering pattern of varieties

Consistent scores for 22 qualitative traits were observed for all the 92 sugarcane varieties over the 3 years of study, indicating stability of traits over the years. A less degree of variation was recorded for traits like plant growth habit, leaf blade curvature, and leaf sheath adherence, whereas, rest of the qualitative traits showed a high degree of variation. The mean phenotypic data of seven quantitative traits, viz., number of millable canes (NMC) per stool, leaf blade width, internode diameter, plant height, fibre content, sucrose content and purity per cent revealed a high degree of polymorphism over the replications and years (Table 2). The bud size was measured as a quantitative trait, but it did not show any significant variation. The number of millable cane per stool (NMC) ranged from 3.5 (CoS 90265) to 9.33 (Co 1336). Leaf blade width was the maximum (5.83 cm) for variety CoPant 91269, while it was the minimum (2.06 cm) for CoS 90265. The internode diameter ranged from 2.08 (UP 05) to 3.41 (CoJ 85), plant height ranged between 1.33 (CoS 797) and 2.48 (CoSe 92423). Fibre content ranged from 8.97% (CoS 99259) to 20.18% (CoP 9302), while sucrose content ranged between 11.15% (CoS 770) and 18.06% (CoS 95255). The purity was recorded in the range of 80.3% (CoJ 89) to 89.46% (CoS 95255). The DUS descriptors are generally used to test the distinctiveness, uniformity and stability of varieties, which is pre-requisite for their registration; moreover, in this

study, such descriptor data were also used to understand phylogenetic relationship among varieties. The average Euclidian distance was 239 between the varieties, and the corresponding tree plotted for 27 phenotypic DUS traits grouped the 92 varieties in to nine (I to IX) major clusters (Fig. 1), where each of the major cluster could be further divided in to 2–3 sub-clusters. The mean values for the seven quantitative traits of the group of varieties of these nine major clusters is presented in Table 3. The group of varieties in individual clusters also displayed a high degree of variability as it was recorded for the individual variety. The cluster I comprised of 17 varieties, which could be further sub-divided in to four sub-clusters (Ia, Ib, Ic, Id). Sub-cluster Ia had seven varieties (CoS 8432, CoLk 8001, CoS 96269, CoH 56, CoS 687, CoS 90269, CoS 99259), sub-cluster Ib had four varieties (CoS 95222, CoSe 92423, CoS 94257, CoH 35), sub-cluster Ic had two varieties (CoS 95270, BO 137), while, sub-cluster Id had four varieties (CoS 8315, CoB 94164, CoS 94270, CoS 02258). The cluster II had nine varieties, five (BO 128, CoH 119, CoS 97261, CoJ 85, CoSe 95422) in sub-cluster IIa, and four (BO 99, CoLk 8102, UP 0097, CoLk 94184) in sub-cluster IIb. Cluster three had only three (BO 91, BO 110, Co 419) varieties. Cluster IV has nine varieties, with two sub-clusters, viz., sub-cluster IVa with four varieties (CoS 8207, CoSe 96436, CoS 541, CoS 8436) and IVb with five varieties (BO 136, CoS 02264, CoPant 97222, BO 138, CoPant 84211). Cluster V contained 21 varieties, which could be further sub-divided in to 3 sub-clusters, viz., Va with 11 varieties (CoH 99, CoJ 88, CoJ 89, CoPant 96219, CoS 01256, CoP 9206, CoSe 01235, CoSe 98231, CoSe 00235, CoS 90265, CoS 770), Vb with two varieties (CoS 96275, CoS 109), and Vc with eight (CoPant 84212, CoS 03279, CoJ 64, CoS 88230, CoS 95255, CoS 245, CoSe 96436, CoS 91269) varieties. The major cluster VI has a total of 11 varieties dispersed in to three sub-clusters; VIa with four (BO 129, CoS 767, CoS 797, CoS 443) varieties, VIb with Co 89003, and VIc contained six (CoP 9702, CoP 9302, CoS 514, UP 9530, CoS 97264, CoS 91230) varieties. The major cluster VII had 13 varieties with six (CoS 97258, CoS 510, CoP 9301,

Table 2 The minimum, maximum, mean and standard error values for the seven quantitative traits of 92 sugarcane varieties (data pooled over 3 years 2011–2012, 2012–2013 and 2013–2014)

Trait	Minimum ^a	Maximum ^a	Mean ± SE
Plant height (m)	1.33 (CoS 797)	2.48 (Co 87263)	2.17 ± 0.28
Fibre per cent	8.97 (CoS 99259)	20.18 (CoP 9302)	14.32 ± 2.42
Sucrose per cent	11.15 (CoS 770)	18.06 (CoS 95255)	14.71 ± 1.51
Number of millable canes	3.53 (CoS 90265)	9.33 (Co 1336)	5.96 ± 1.68
Leaf blade width (cm)	2.06 (CoS 90265)	5.83 (CoPant 96219)	4.08 ± 0.73
Internode diameter (cm)	2.08 (UP 05)	3.41 (CoJ 85)	2.5 ± 0.29
Purity per cent	80.3 (CoJ 89)	89.46 (CoS 95255)	84.47 ± 3.43

^aName in parenthesis refers to the variety recorded such minimum/maximum value. S.E.: Standard error

Fig. 1 Dendrogram of 92 sugarcane varieties from subtropical India based on 27 morphological descriptors using squared Euclidean distance

Table 3 Mean values for the quantitative traits per cluster; the 92 varieties were grouped in to 9 major clusters as per squared Euclidean distance based on 27 morphological descriptor data pooled over the 3 years (2011–2012, 2012–2013, 2013–2014)

Cluster	Genotypes in the cluster	Trait						
		NMC	LBW	ID	PH	FP	SP	PP
I	CoS 8432, CoLk 8001, CoS 96269, CoH 56, CoS 687, CoS 90269, CoS 99259, CoS 95222, CoSe 92423, CoS 94257, CoH 35, CoS 95270, BO 137, CoS 8315, CoB 94164, CoS 94270, CoS 02258	6.44	4.21	2.55	2.20	13.74	14.34	83.80
II	BO 128, CoH 119, CoS 97261, CoJ 85, CoSe 95422, BO 99, CoLk 8102, UP 0097, CoLk 94184	6.09	4.33	2.62	2.26	13.94	14.70	84.54
III	BO 91, BO 110, Co 419	6.89	3.58	2.47	2.23	14.12	14.19	82.38
IV	CoS 8207, CoSe 96436, CoS 541, CoS 8436, BO 136, CoS 02264, CoPant 97222, BO 138, CoPant 84211	5.55	4.21	2.52	2.06	15.03	14.90	84.15
V	CoH 99, CoJ 88, CoJ 89, CoPant 96219, CoS 01256, CoP 9206, CoSe 01235, CoSe 98231, CoSe 00235, CoS 90265, CoS 770, CoS 96275, CoS 109, CoPant 84212, CoS 03279, CoJ 64, CoS 88230, CoS 95255, CoS 245, CoSe 95436, CoS 91269	5.68	4.14	2.53	2.22	14.12	14.90	84.22
VI	BO 129, CoS 767, CoS 797, CoS 443, Co 89003, CoP 9702, CoP 9302, CoS 514, UP 9530, CoS 97264, CoS 91230	6.00	3.67	2.40	2.02	15.16	14.18	84.99
VII	CoS 97258, CoS 510, CoP 9301, CoSe 01424, CoS 96258, CoH 110, CoPant 90223, Co 6811, Co 89029, BO 109, BO 130, BO 120, Co 1148	6.00	4.19	2.50	2.15	14.58	15.04	85.51
VIII	Co 87263, CoS 96268	4.75	4.20	2.56	2.34	12.91	16.96	85.95
IX	Co 1158, CoH 92, Co 6425, Co 1336, UP 05, CoJ 83, Co 453	5.88	3.77	2.52	2.20	14.73	14.62	84.88

NMC number of millable canes, LBW leaf blade width, ID internode diameter, PH plant height, FP fibre per cent, SP sucrose per cent, PP purity per cent

CoSe 01424, CoS 96258, CoH 110) in sub-cluster VIIa, and seven (CoPant 90223, Co 6811, Co 89029, BO 109, BO 130, BO 120, Co 1148) varieties in sub-cluster VIIb. Cluster VIII contained only two varieties Co 87263 and CoS 96268. The Cluster IX contained seven varieties with five (Co 1158, CoH 92, Co 6425, Co 1336, UP 05) in sub-cluster IXa, while CoJ 83 and Co 453 in sub-cluster IXb.

Molecular marker analysis

A total 643 dominant loci across the 92 sugarcane varieties were generated by 80 polymorphic SSR primer pairs with an average of 8.41 loci per primer that ranged from 2 (primers IISR_75, IISR_259) to 17 (primer IISR_48). The average $I_{b_{av}}$ value of all the 80 SSR primers was 0.44. On an average, the genomic-SSR primers produced more (9.25) amplicons compared to that of EST-SSRs (8.2). But the $I_{b_{av}}$ values for both type of primers was similar (0.43). For EST-SSR primers, that produced amplification products in size range of 300–1200 bp, the highest (0.85) $I_{b_{av}}$ value was recorded for primer IISR_17b, whereas, it was the lowest (0.04) for primer IISR_91. Similarly, in the case of genomic-SSRs, the highest (0.63), and the lowest (0.12) $I_{b_{av}}$ value was recorded for primers SEGMS113a and SEGMS14, respectively (Table 4). The amplification profiles of 92 varieties with EST-SSR primer IISR_298b and genomic-SSR primer SEGMS1016 is presented as Supplementary Figs. S1 and S2.

Population structure

Based on the second order rate of change of likelihood distribution mean ($L''(K)$), the optimum value of $K=7$ was obtained, where the 92 varieties could be classified in to 7 sub-clusters (Fig. 2). However, it was observed that all the 92 sugarcane varieties had admixtures and no sub-group had a pure unblemished structure profile. In addition, PCA was also carried out for identification of innate sub-groupings within the varieties. The first two principal components (PC1, PC2) accounted for ~10% of the total variation (Fig. 3). The PCA was able to group the varieties with similar parentage together. For example, the variety Co 1158 which has been frequently used as a female parent in breeding of sub-tropical sugarcane varieties, clustered with CoS 95255, CoS 97264, CoS 770, CoS 02258, and CoLk 8102 which had Co 1148 as their female parent. The varieties CoS 8315, CoS 90269, CoH 110, CoS 02264 and CoS 96269 had Co 1148 as one of their parents and clustered near to Co 1148. Similarly, varieties CoSe 92423, CoSe 96436, CoSe 95436, CoB 94164 and CoSe 10424 clustered with BO 91 which is one of the parents in their pedigree. A similar pattern was observed in the STRUCTURE results where the varieties derived from Co 1158, Co 1148 and Bo 91 exhibited similar profiles with their respective parental lines.

Table 4 Size range, number of loci amplified and $I_{b_{av}}$ value of 80 SSR primers that were used in present study

Sl. no.	Primer	Size range of product (bp)	Number of loci	$I_{b_{av}}$ value	Sl. no.	Primer	Size range of product (bp)	Number of loci	$I_{b_{av}}$ value
1	IISR_298a	140–400	7	0.198	41	IISR_114	140–580	12	0.593
2	IISR_298b	130–450	8	0.336	42	IISR_317	90–400	9	0.577
3	IISR_256	150–500	14	0.622	43	IISR_337	450–1000	6	0.183
4	IISR_116	250–380	7	0.569	44	IISR_276	100–700	11	0.729
5	IISR_127	145–170	3	0.348	45	IISR_89	150–500	7	0.457
6	IISR_261	140–300	5	0.750	46	IISR_269	150–900	11	0.581
7	IISR_75	150–350	2	0.193	47	IISR_310	110–300	9	0.488
8	IISR_132	270–900	5	0.643	48	IISR_91	110–500	8	0.043
9	IISR_17a	150–250	7	0.352	49	IISR_76	150–420	6	0.049
10	IISR_299	150–340	5	0.563	50	IISR_98	35–70	3	0.489
11	IISR_199a	200–500	7	0.623	51	IISR_14a	220–800	7	0.343
12	IISR_17b	140–250	5	0.859	52	IISR_9	150–170	3	0.310
13	IISR_17c	160–600	9	0.593	53	IISR_213	900–1200	3	0.130
14	IISR_267	150–370	5	0.730	54	IISR_78	70–700	10	0.333
15	IISR_42	160–310	7	0.574	55	IISR_82a	200–580	7	0.543
16	IISR_46b	160–280	5	0.635	56	IISR_7	30–600	6	0.687
17	IISR_40	150–400	13	0.217	57	EST A51	200–700	9	0.144
18	IISR_94a	150–300	7	0.428	58	EST A66	80–1000	11	0.247
19	IISR_336	120–370	8	0.380	59	EST A26	160–600	8	0.366
20	IISR_80	120–300	8	0.640	60	IISR_77	160–600	9	0.122
21	IISR_88a	140–450	13	0.179	61	EST A48	150–800	12	0.220
22	IISR_90	360–400	3	0.257	62	EST A68	120–200	5	0.426
23	IISR_176c	90–400	13	0.235	63	IISR_48	30–1000	17	0.602
24	IISR_227	90–170	8	0.209	64	IISR_308a	100–400	9	0.455
25	IISR_288	140–260	7	0.556	65	SEGMS770	160–1000	9	0.327
26	IISR_176a	150–320	10	0.377	66	SEGMS780	125–500	7	0.393
27	IISR_111	110–380	13	0.716	67	SEGMS14	310–1000	10	0.126
28	IISR_333	130–150	3	0.237	68	SEGMS73	50–800	14	0.343
29	IISR_81a	130–500	14	0.421	69	SEGMS35	40–120	5	0.522
30	IISR_81b	70–590	14	0.330	70	SEGMS1020	50–700	12	0.566
31	EST B14	280–1200	9	0.444	71	SEGMS121	50–200	5	0.491
32	EST A71	60–750	14	0.248	72	SEGMS40	50–1000	7	0.571
33	IISR_322	70–600	16	0.423	73	SEGMS114	50–500	6	0.470
34	IISR_259	300–350	2	0.149	74	SEGMS113a	50–550	11	0.638
35	IISR_159	100–270	6	0.653	75	SEGMS118a	30–550	14	0.574
36	IISR_4	140–500	9	0.481	76	SEGMS1016	50–1000	16	0.520
37	EST A19	90–900	9	0.146	77	NKS05	160–500	7	0.455
38	IISR_306	130–650	14	0.278	78	NKS06	150–450	9	0.368
39	IISR_104	90–400	8	0.153	79	NKS40	120–210	7	0.356
40	IISR_156	160–260	5	0.570	80	SEGMS11	180–800	9	0.389

Cluster analysis based on SSR markers

The ninety-two sugarcane varieties on the basis of dendrogram generated using 643 SSR loci could be classified into four (I, II, III, IV) major clusters (Fig. 4). The Nei's genetic distance based on the genotypic data with an average distance coefficient of 0.49 was able to satisfactorily

differentiate all the varieties. The cluster I had the maximum (67) number of varieties. The dissimilarity coefficient among varieties ranged from 0.137 (CoS 99259 and CoS 03279) to 0.405 (for varieties CoLk 94184 and CoS 91269). This indicated that varieties CoS 99259 and CoS were the most similar, while, varieties CoLk 94184 and CoS 91269 were the most dissimilar. The cluster II included 21 varieties (BO

Fig. 4 Neighbour-joining (NJ) dendrogram showing genetic relationship obtained from cluster analysis of 92 sugarcane genotypes from sub-tropical India based on 643 SSR loci

among sugarcane accessions (Ming et al. 2010). For genetic improvement of complex crops like sugarcane SSR markers could be useful for structuring genetic diversity of germplasm collection according to geographical origin (Singh et al. 2010) formation of core collection (Nayak et al. 2014) and for construction of genetic map (Liu et al. 2010; Marconi et al. 2011).

A non-significant correlation ($r=0.0477$) was observed between the similarity based on SSR and average taxonomic distance based on morphological scoring. This is in agreement with the reports that morphological trait-based dissimilarity and similarity based on molecular markers usually show non-significant correlation (Riday et al. 2003; Dey et al. 2006). This could be due to the fact that conversion of data into scores tends to reduce the contrast within the genotypes resulting to a lesser congruence to the genotypic data which is more exhaustive in nature. The reason of such mismatch may also be due to the fact that most of the morphological traits are controlled by a number of genes and are highly influenced by the environment, while detection of molecular markers is not confounded by the environmental effects. As a result, DNA markers may accurately assay the degree of genetic change distinguishing two genomes, but they may not necessarily reflect the divergence in terms of changes in traits of agronomic importance. Genetic diversity among the main progenitors and their resultant sugarcane varieties was assessed using SSR markers by dosSantos et al. (2012). It was reported that varieties grouped in to different clusters, but no correlation was obtained with regard to their parentage. The low level of diversity within the panel of 92 sub-tropical sugarcane varieties could be a result of exploiting limited number of parental lines in their breeding history.

In addition to genetic diversity, population stratification is also an important aspect for marker discovery and breeding programmes. In the present study, it was observed that the Bayesian method of population stratification using STRUCTURE was able to divide the 92 sugarcane varieties in to 7 sub-groups. However, it was observed that each of the 92 varieties had admixtures in their genetic constitution. The STRUCTURE-based grouping and PCA was not able to separate genotypes according to their geographical location. However, PCA where although the first two components were explaining only 10% of the total genetic variation, was able to group the varieties that had similar pedigree together in most if not all of the cases. This suggest that in sub-tropical sugarcane varieties from India, the pedigree of the genotype contributes to the sub-clustering pattern. And, due to that the predominantly utilized parental lines Co 1148, Co 1158, Bo 91 and Co 775 grouped together with their respective progenies. Although, the grouping of the varieties was not in concordance with their geographical distribution or maturity groups, but knowledge and information of genetic structure and molecular diversity generated in this

study could be useful for selection of parents and their combination for a targeted breeding programme of sugarcane.

The molecular profiling of a plant variety can prove to be a valuable tool for protecting a plant breeders' and Farmers' rights (Cooke 1995). In recent past, SSR markers by virtue of the ability produce highly reproducible profiles have been increasingly employed for a reliable identification of plant varieties in several crop plants (McCouch et al. 1997). Further, with the help of DNA fingerprinting data, it could be possible to identify a mislabelled variety during field experiments (Pan et al. 2003). Another significant use of SSR-based fingerprints involves marker-assisted selection (MAS) of real cross progeny at seedling stage and assessment of the degree of self-pollination during crossing. Thus, it is evident that information based on SSR markers can be efficiently used by the sugarcane breeders to identify true offspring in F_1 population, and hence can improve the sugarcane crossing programs. Such SSR fingerprints can also be used to ascertain the identity of the male parent in the resulting progeny particularly from poly-crosses (Tew and Pan 2010), which would certainly help the breeders in accelerating the number of desirable crosses in the breeding programs with less number of tassels available from desirable parents with keeping most of the pedigree information intact (Pan 2010). Such findings strongly advocate the need of inclusion of highly polymorphic and reproducible SSR markers for supplementing the identification of varieties which is primarily based on their phenotypic attributes with their unambiguous molecular identity information. This would not only aid conventional DUS testing, but also put an end to the ambiguities associated with phenotypic characterization, and therefore, also rule out any possibility of re-registration of already registered variety under PPV&FR act.

Conclusion

The study was able to effectively delineate the 92 sugarcane varieties of sub-tropical India. The pedigree was found to be the most effective factor in grouping of the sugarcane varieties. Under the Protection of Plant Varieties and Farmer's Rights Act-2001, and Indian Seed Bill-2004, the DUS test of sugarcane varieties has been made mandatory so as to establish the identity of a variety for granting plant breeders' and farmers' rights. Sugarcane morphological descriptors may not alone be sufficient to resolve the identity unambiguously in case of closely related varieties and essentially derived varieties which differ for only two or three descriptors. This study suggests the inclusion of SSR markers in profiling of varieties as an additional tool for morphological descriptors since the complete replacement of conventional phenotypic evaluation of varieties by molecular data is not possible due to poor correlation between the SSR-based similarity

and average taxonomic distance based on morphological descriptors.

Acknowledgements This work was funded by a research grant from Protection of Plant Varieties and Farmers' Rights Authority, Government of India (PPV&FRA/Reg/19-5/DUS/2012) in the form of a project on DUS Testing under Central Sector Scheme for Implementation of PVP legislation, and from Department of Biotechnology (DBT), Government of India as a BioCARE Project. The authors are thankful to the Director, ICAR-Indian Institute of Sugarcane Research, Lucknow for providing infrastructure facilities.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aitken K, Jackson P, McIntyre C (2005) A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. *Theor Appl Genet* 110:789–801
- Aitken KS, Li JC, Jackson P, Piperidis G, McIntyre CL (2006) AFLP analysis of genetic diversity within *Saccharum officinarum* and comparison with sugarcane cultivars. *Aust J Agric Res* 57:1167–1184
- Allen CJ, Mackay MJ, Aylward JH, Campbell JA (1997) New technologies for sugar milling and by-product modification. In: Keating BA, Wilson JR (eds) *Intensive sugarcane production: meeting the challenges beyond 2000*. CAB International, Wallingford, pp 267–285
- Alwala S, Kimbeng CA, Veremis JC, Gravois KA (2008) Linkage mapping and genome analysis in a *Saccharum* interspecific cross using AFLP, SRAP and TRAP markers. *Euphytica* 164:37–51
- Andru S, Pan YB, Thongthawee S, Burner DM, Kimbeng CA (2011) Genetic analysis of the sugarcane (*Saccharum* spp.) cultivar “LCP 85–384”. I. Linkage mapping using AFLP, SSR, and TRAP markers. *Theor Appl Genet* 123:77–93
- Anonymous (2009) Specific DUS test guidelines for three notified crops, PPV&FRA, publication no. SG/16/2009. *Plant Variety J India* 3(9):42–54
- Anonymous (2016) Agricultural statistics at a glance-2016, <http://eands.dacnet.nic.in/>. Accessed 22 Jan 2018
- Besse P, Taylor G, Carroll B, Berding N, Burner D, McIntyre CL (1998) Assessing genetic diversity in a sugarcane germplasm collection using an automated AFLP analysis. *Genetica* 104:143–153
- Boc A, Makarenkov V (2012) T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. *Nucl Acids Res* 40:573–579
- Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635
- Breyne P, Boerjan W, Gerats T, Van Montagu M, Gysel A (1997) Application of AFLP in plant breeding, molecular biology and genetics. *Belg J Bot* 129:107–117
- Cai Q, Aitken KS, Fan YH, Piperidis G, Jackson P, McIntyre CL (2005) A preliminary assessment of the genetic relationship between *Erianthus rockii* and the *Saccharum complex* using microsatellite (SSR) and AFLP markers. *Plant Sci* 169:976–984
- Casu RE, Manners JM, Bonnett GD, Jackson PA, McIntyre CL, Dunne R, Chapman SC, Rae AL, Grof CPL (2005) Genomics approaches for the identification of genes determining important traits in sugarcane. *Field Crops Res* 92:137–147
- Chang D, Yang FY, Yan JJ, Wu YQ, Bai SQ, Liang XZ, Zhang YW, Gan YM (2012) SRAP analysis of genetic diversity of nine native populations of wild sugarcane, *Saccharum spontaneum*, from Sichuan, China. *Genet Mol Res* 11:1245–1253
- Cooke RJ (1995) Gel electrophoresis for the identification of plant varieties. *J Chromatogr* 698:281–299
- Creste S, Sansoli DM, Tardiani ACS, Silva DN, Goncalves FK, Favero TM, Medeiros CNF, Festucci CS, Carlini-Garcia LA, Landell MGA, Pinto LR (2010) Comparison of AFLP, TRAP and SSRs in the estimation of genetic relationships in sugarcane. *Sugar Tech* 12:150–154
- D'Hont A, Ison D, Alix K, Roux C, Glaszmann JC (1998) Determination of basic chromosome numbers in the genus *Saccharum* by physical mapping of ribosomal RNA genes. *Genome* 41:221–225
- daCosta MLM, Amorim LLB, Onofre AV, deMelo LJT, deOliveira MBM, deCarvalho R, Benko-Iseppon AM (2011) Assessment of genetic diversity in contrasting sugarcane varieties using inter-simple sequence repeat (ISSR) markers. *Am J Plant Sci* 2(3):425–432
- Dey SS, Singh AK, Chandel D, Behera TK (2006) Genetic diversity of bitter melon (*Momordica charantia* L.) genotypes revealed by RAPD markers and agronomic traits. *Sci Hortic* 109(1):21–28
- Dias MOS, Tassia LJ, Otávio C, Cunha MP, Jesus CDF, Rossell CEV, Filho RM, Bonomi A (2012) Integrated versus stand-alone second generation ethanol production from sugarcane bagasse and trash. *Bioresour Technol* 103(1):152–161
- dosSantos JM, Duarte Filho LSC, Soriano ML, daSilva PP, Nascimento VX, deSouza Barbosa GV, Almeida C (2012) Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. *Ind Crop Prod* 40:145–150
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12:13–14
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620
- Govindaraj P, Balasundaram N, Sharma TR, Bansal KC, Koundal KR, Singh NK (2005) Development of new STMS markers for sugarcane. Sugarcane Breeding Institute, Coimbatore, Tamilnadu, India. <http://www.nrcpb.org/content/development-new-stms-markers-sugarcane>. Accessed 25 July 2009
- Heinz DJ, Tew TL (1987) Hybridization procedures. In: Heinz DJ (ed) *Sugarcane improvement through breeding*. Elsevier, Amsterdam, pp 313–342
- Hogarth DM (1987) Genetics of sugarcane. In: Heinz DJ (ed) *Sugarcane improvement through breeding*. Elsevier, Amsterdam, pp 255–271
- IBM Corp. Released (2011) IBM SPSS statistics for windows, ver. 20.0. IBM Corp, Armonk
- Jarne P, Lagoda PJ (1996) Microsatellites, from molecules to populations and back. *Trends Ecol Evol* 11(10):424–429
- Khan MS, Yadav S, Srivastava S, Swapna M, Chandra A, Singh RK (2011) Development and utilisation of conserved-intron scanning marker in sugarcane. *Aust J Bot* 59(1):38–45
- Khan MS, Kumar S, Singh RK, Singh J, Duttamajumder SK, Kapur R (2018) Characterization of leaf transcriptome, development and utilization of unigenes-derived microsatellite markers in sugarcane (*Saccharum* sp. hybrid). *Physiol Mol Biol Plants* 24(4):665–682
- Liu X, Mao J, Lu X, Ma L, Aitken KS, Jackson PA, Cai Q, Fan Y (2010) Construction of molecular genetic linkage map of sugarcane based on SSR and AFLP markers. *Acta Agron Sin* 36(1):177–183
- Maccheroni W, Jordao H, De Gaspari R, De Moura GL, Matsuoka S (2009) Development of a dependable microsatellite-based fingerprinting system for sugarcane. *Sugar Cane Int* 27(2):47–52

- Marconi TG, Costa EA, Miranda HR, Mancini MC, Cardoso-Silva CB, Oliveira KM, Pinto LR, Mollinari M, Garcia AA, Souza AP (2011) Functional markers for gene mapping and genetic diversity studies in sugarcane. *BMC Res Notes* 4(1):264
- McCouch SR, Chen XL, Panaud O, Temnykh S, Xu YB, Cho YG, Huang N, Ishii T, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol* 35:89–99
- Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W, Waugh R (1997) Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. *Mol Breed* 3:127–136
- Ming R, Moore PH, Wu K, D-Hont A, Glaszmann JC, Tew TL, Mirkov TE, DaSilva J, Jifon J, Rai M, Schnell RJ (2010) Sugarcane improvement through breeding and biotechnology. *Plant Breed Rev* 27:15–118
- Nawaz S, Khan FA, Tabasum S, Zakria M, Saeed A, Iqbal MZ (2010) Phylogenetic relationships among *Saccharum* clones in Pakistan revealed by RAPD markers. *Genet Mol Res* 9(3):1673–1682
- Nayak SN, Song J, Villa A, Pathak B, Ayala-Silva T, Yang X, Todd J, Glynn NC, Kuhn DN, Glaz B, Gilbert RA, Comstock JC (2014) Promoting utilization of *Saccharum* spp. genetic resources through genetic diversity analysis and core collection construction. *PLoS One* 9(10):1–10
- Nei M (1972) Genetic distance between populations. *Am Nat* 106(949):283–292
- Noli E, Teriaca MS, Sanguineti Conti S (2008) Utilization of SSR and AFLP markers for the assessment of distinctness in durum wheat. *Mol Breed* 22:301–313
- Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC (2009) Characterization of new polymorphic functional markers for sugarcane. *Genome* 52:191–209
- Pan YB (2006) Highly polymorphic microsatellite DNA markers for sugarcane germplasm evaluation and variety identity testing. *Sugar Tech* 8(4):246–256
- Pan YB (2010) Databasing molecular identities of sugarcane (*Saccharum* spp.) clones constructed with microsatellite (SSR) DNA markers. *Am J Plant Sci* 1(2):87–94
- Pan YB, Miller JD, Schnell RJ, Richard EP, Wei Q (2003) Application of microsatellite and RAPD fingerprints in the Florida sugarcane variety program. *Int Sugar J March/April*:19–24
- Parida SK, Kalia SK, Kaul S, Dalal V, Hemaprabha G (2009) Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theor Appl Genet* 118:327–338
- Pinto LR, Oliveira KM, Ulian EC, Garcia AA, deSouza AP (2004) Survey in the sugarcane expressed sequence tag database (SUCEST) for simple sequence repeats. *Genome* 47:795–804
- Pinto LR, Garcia AAF, Pastina MM, Teixeira LHM, Bressiani JA, Ulian EC, Souza AP (2010) Analysis of genomic and functional RFLP derived markers associated with sucrose content, fiber and yield QTLs in a sugarcane (*Saccharum* spp.) commercial cross. *Euphytica* 172(3):313–327
- Pourabed E, Noushabadi JMR, Jamali SH, Alipour MN, Zareyan A, Sadeghi L (2015) Identification and DUS testing of rice varieties through microsatellite markers. *Int J Plant Genomics*. <https://doi.org/10.1155/2015/965073> (Article ID 965073)
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multi-locus genotype data. *Genetics* 155:945–959
- Que YX, Pan YB, Lu YH, Yang C, Yang YT, Huang N, Xu LP (2014) Genetic analysis of diversity within a Chinese local sugarcane germplasm based on start codon targeted polymorphism. *BioMed Res Int* 2014:468375. <https://doi.org/10.1155/2014/468375>
- Ramakrishnan M, Ceasar SA, Duraipandiyan V, Al-Dhabi N, Ignacimuthu S (2015) Using molecular markers to assess the genetic diversity and population structure of finger millet (*Eleusine coracana* (L.) Gaertn.) from various geographical regions. *Genet Resour Crop Evol* 63(2):1–16
- Riday H, Brummer EC, Campbell TA, Luth D, Cazcarro PM (2003) Comparisons of genetic and morphological distance with heterosis between *Medicago sativa* subsp. *sativa* and subsp. *falcata*. *Euphytica* 131(1):37–45
- Rosenberg MS, Anderson CD (2011) PASSaGE: pattern analysis, spatial statistics and geographic exegesis. Version 2. *Methods Ecol Evol* 2(3):229–232
- Silva CM, Mangolin CA, Mott AS, Machado MFPS (2008) Genetic diversity associated with *in vitro* and conventional bud propagation of *Saccharum* varieties using RAPD analysis. *Plant Breed* 127(2):160–165
- Silva DC, Filho LS, Costa D, Santos JMD, Barbosa G, deSouza V, deCícero A (2012) DNA fingerprinting based on simple sequence repeat (SSR) markers in sugarcane clones from the breeding program RIDESA. *Afr J Biotechnol* 11(21):4722–4728
- Singh RK, Mishra SK, Singh SP, Mishra N, Sharma ML (2010) Evaluation of microsatellite markers for genetic diversity analysis among sugarcane species and commercial hybrids. *Aust J Crop Sci* 4(2):115–124
- Singh RK, Khan MS, Singh R, Pandey DK, Kumar S, Lal S (2011) Analysis of genetic differentiation and phylogenetic relationships among sugarcane genotypes differing in response to red rot. *Sugar Tech* 13(2):137–144
- Singh RK, Jena SN, Khan S, Yadav S, Banarjee N, Raghuvanshi S, Bhardwaj V, Dattamajumder SK, Kapur R, Solomon S, Swapna M, Srivastava S, Tyagi AK (2013) Development, cross-species/genera transferability of novel EST-SSR markers and their utility in revealing population structure and genetic diversity in sugarcane. *Gene* 524:309–329
- Tai PYP (1989) Progress and problems of intergeneric hybridization in sugarcane breeding. *Proc Inter Am Sugarcane Sem* 1:391–395
- Tew TL, Pan YB (2010) Microsatellite (simple sequence repeat) marker-based paternity analysis of a seven-parent sugarcane polycross. *Crop Sci* 50(4):1401–1408
- UPOV-BMT (2002) BMT/36/10 progress report of the 36th session of the technical committee, the technical working parties and working group on biochemical and molecular techniques and DNA-profiling in particular, Geneva
- Walker DIT (1987) Manipulating the genetic base of sugarcane. In: *Copersucar international sugarcane breed workshop*, Copersucar, Sao Paulo, pp 321–334
- XLSTAT (2017) Data analysis and statistical solution for Microsoft Excel. Addinsoft, Paris
- Xu ML, Melchinger AE, Xia XC, Lubberstedt T (1999) High-resolution mapping of loci conferring resistance to sugarcane mosaic virus in maize using RFLP, SSR, and AFLP markers. *Mol Genet* 261:574–581
- Zawedde BM, Ghislain M, Magembe E, Amaro GB, Grumet R, Hancock J (2014) Characterization of the genetic diversity of Uganda's sweet potato (*Ipomoea batatas*) germplasm using microsatellite markers. *Genet Resour Crop Evol* 62:501–513