RESEARCH ARTICLE

Molecular diversity and genetic structure of *Saccharum* complex accessions

Carolina Medeiros, Thiago Willian Almeida Balsalobre, Monalisa Sampaio Carneiro⊙*

Departamento de Biotecnologia e Produção Vegetal e Animal, Centro de Ciências Agrárias, Universidade Federal de São Carlos, Araras, São Paulo, Brasil

* monalisa@ufscar.br

Abstract

Sugarcane is an important crop for food and energy security, providing sucrose and bioethanol from sugar content and bioelectricity from lignocellulosic bagasse. In order to evaluate the diversity and genetic structure of the Brazilian Panel of Sugarcane Genotypes (BPSG), a core collection composed by 254 accessions of the *Saccharum* complex, eight TRAP markers anchored in sucrose and lignin metabolism genes were evaluated. A total of 584 polymorphic fragments were identified and used to investigate the genetic structure of BPSG through analysis of molecular variance (AMOVA), principal components analysis (PCA), a Bayesian method using STRUCTURE software, genetic dissimilarity and phylogenetic tree. AMOVA showed a moderate genetic differentiation between ancestors and improved accessions, 0.14, and the molecular variance was higher within populations than among populations, with values of 86%, 95% and 97% when constrasting improved with ancestors, foreign with ancestors and improved with foreign, respectively. The PCA approach suggests clustering in according with evolutionary and Brazilian breeding sugarcane history, since improved accessions from older generations were positioned closer to ancestors than improved accessions from recent generations. This result was also confirmed by STRUCTURE analysis and phylogenetic tree. The Bayesian method was able to separate ancestors of the improved accessions while the phylogenetic tree showed clusters considering the family relatedness within three major clades; the first being composed mainly by ancestors and the other two mainly by improved accessions. This work can contribute to better management of the crosses considering functional regions of the sugarcane genome.

Introduction

Sugarcane, a high efficiency photosynthetic grass, is important for economy of many countries in the tropics and subtropics, playing a central role as a primary sugar-producing crop and has major potential as a bioenergy crop [1–3]. The modern sugarcane cultivars originate from the *Saccharum* complex, which gathers two wild *Saccharum* species (*S. spontaneum and S. robustusm*), four cultivated species (*S. officinarum, S. sinense, S. barberi and S. edule*) and four related interbreeding genera (*Erianthus, Miscanthus, Narenga and Sclerostachya*) [4–7]. The

Competing interests: The authors have declared that no competing interests exist.

Saccharum species present large genome and variation in the number of chromosomes [8–10]. This complexity was inherited by modern sugarcane cultivars, which present a variable level of ploidy, frequent aneuploidy, and large genome size around 10 Gb [10–12].

The first interspecific hybridizations occurred among *S. officinarum* and *S. spontaneum* species, followed by successive backcrosses with *S. officinarum* aiming to recover the sucrose genes [4,13]. According to this initial breeding strategy, naturally few accessions were used at the crosses and approximately 80% of the genome of current sugarcane cultivars came from *S. officinarum*, 10–15% from *S. spontaneum* and the remaining 5–10% being recombinant chromosomes [14,15]. Differently of *S. officinarum*, the accessions of *S. spontaneum* present low sugar content, high biomass production and resistance to some diseases [2,16]. Thus, it is an important genetic background to increase biomass production and have been used into plant breeding for energy cane purpose [13,17]. This energy cane with more fiber content and low sugar production could be an efficient source for second-generation ethanol production [18,19]. Furthermore, the higher rates of biomass and/or sucrose production can be obtained through better management of genetic resources present in germplasm banks and core collections.

The pre-breeding strategy to choose parents for crosses is an important step to increase the probability of obtaining more productive cultivars. Although morphological and agronomical characterization plays an important role in the classification and organization of germplasm accessions, errors may occur since vegetative traits are influenced by environmental effects, showing continuous variation and a high degree of plasticity, and which many times do not reflect the real genetic diversity of the *Saccharum* spp. accessions [20]. So, the molecular profile could be used to complement the morphological characterization and identify in a more reliable way better combinations between accessions for crosses according to breeding goals [18,21,22]. Molecular markers are useful tools to detect variations directly in the genome and have been used to investigate the genetic diversity of *Saccharum* spp. accessions [23,24]. However, few studies performed molecular characterization of sugarcane core collections with functional markers, most of them evaluated non-coding or repeating regions of the genome and may not be useful about traits of interest to the breeders [25,26]. Even when functional molecular markers were used, the number of *Saccharum* spp. accessions evaluated was not more than 181 [27]. Clearly, there is a need to expand the characterization of larger and more representative *Saccharum* complex collections with functional markers, bringing together both alleles under bottleneck effect and those that may be new sources of variation for target traits.

TRAP (Target Region Amplification Polymorphism) and EST-SSR (Simple Sequence Repeats from Expressed Sequence Tag) molecular markers, beside those identified through genetic mapping, could be used to screening collections into functional regions of genome [25]. TRAP markers are interesting because they search for polymorphisms around genes that may be under selection process [28,29]. Furthermore, this approach may indicate accessions for crosses according to molecular profile and, consequently, guide introgression of the new variants for traits of interest [25,29]. Sucrose and lignin are target traits to sugarcane and energy cane breeding programs; increase sugar content is one of the main goals of sugarcane breeding programs around the world [30], while decreasing lignin content may facilitate cellulose saccharification for second-generation ethanol production from both sugarcane and energy cane [19,31]. Sucrose and lignin traits are governed by some genes and metabolic pathways described in the literature [16,21,30–32], so the use of TRAP markers based on these genes may be a valuable tool to characterize *Saccharum* spp. accessions and research new allelic variants. Therefore, in this current assignment our objectives were to (i) characterize a core collection of sugarcane composed by 254 accessions of the *Saccharum* complex, and (ii) perform diversity and population structure assessments, using genotyping data obtained through

TRAP markers based on the sucrose and lignin genes. We discuss these results in the context of how functional markers are useful to report evolutionary and breeding history of sugarcane.

Materials and methods

Plant material and DNA extraction

In this study, a total of 254 accessions (S1 Table) of the Brazilian Panel of Sugarcane Genotypes (BPSG) were used. BPSG is a mini core collection from germplasm bank of the RIDESA (Interuniversity Network for the Development of Sugarcane Industry) and consisted of 81 ancestors accessions (A) (75 accessions from *Saccharum* spp. and 06 from *Erianthus* spp.), 137 hybrids from Brazilian breeding programs (BB) and 36 hybrids from Foreign breeding programs–Foreign Hybrids (FH) [33]. The BPSG accessions were chosen according to the following criteria: i) relevant Brazilian cultivars, ii) main parents for Brazilian breeding programs; iii) cultivars from countries that grow sugarcane; iv) parents used in mapping programs [34,35]; and v) representatives of the species from which the *Saccharum* complex originated. The genetic variability present into BPSG, for the most part, was a genetic basis for Brazilian sugarcane breeding programs. The stalks of the accessions were collected and total genomic DNA was extracted from a fresh meristem cylinder as proposed by Al-Janabi et al. [36]. DNA concentration was estimated by a Nanodrop One spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and then the samples were stored at −20˚C until further use.

TRAP markers, genotyping and polymorphism analysis

To compose TRAP markers four arbitrary and five fixed primers were used (S2 Table). The arbitrary primers were adapted of Li and Quiros [37], Alwala et al. [21] and Suman et al. [30]. The three fixed primers associated with sucrose metabolism genes were based on Alwala et al. [21] (*sucrose synthase* (SuSy), *sucrose phosphate synthase* (SuPS) and *starch synthase* (StSy)) while two fixed primers associated with lignin metabolism genes were based on Suman et al. [30] (*caffeic acid O-methyltransferase* (COMT) and *ferulate-5-hydroxylase* (F5H)). Thus, eight TRAP markers were performed based on high percentage of polymorphism showed by the reference studies: StSy + Arbi2, StSy + Arbi3, SuPS + Arbi2, SuPS + Arbi3, SuSy + Arbi1-A, SuSy + Arbi2 for sucrose metabolism and COMT+Arbi1-S and F5H+Arbi1-S for lignin metabolism. The PCR were performed on Mastercycler thermocycler (Eppendorf, Westbury, NY, USA) according to the protocols described by Alwala et al. [21] and Suman et al. [30] for TRAP markers related with sucrose and lignin metabolisms, respectively. After PCR, the amplified products were run on 6.5% (w/v) polyacrylamide denaturing gel for 4.0 h at 65 W and silver staining procedure was employed to detect the fragments as described by Creste et al. [38]. The fragments were scored as "1" for presence and "0" for absence, in all accessions. Only clearly distinguishable fragments were scored. For each TRAP marker, the presence of exclusive fragments was investigated. Through the binary matrix, the PIC (Polymorphism Information Content) and DP (Discriminatory Power) values were calculated according to Botstein et al. [39] and Tessier et al. [40], respectively. PIC was used as a tool to measure the information of a given marker locus for the pool of accessions, while DP was used as a measure of marker efficiency for the purpose of identification of accession, i.e., the probability that two randomly chosen individuals have different patterns [41].

Sequence annotation

The available sequences that gave rise to fixed primers of TRAP markers were used to annotation (S3 Table). To found homologies the initial sequences from Genbank were blasted against

the NCBI non-redundant database via BLASTX and against the *Sorghum bicolor* database via the Phytozome website [42]. The metabolic pathways and biochemical reactions were also verified through the InterMine repository present on the Phytozome.

Genetic structure

The genetic structure of BPSG was investigated using different methods: i) analysis of molecular variance (AMOVA); ii) Principal component analysis (PCA); iii a Bayesian model-based method using STRUCTURE software; and iv) genetic dissimilarity and phylogenetic analysis. AMOVA was performed by the GenAIEx software [43] to quantify the degree of differentiation and distribution of the genetic variability between and within of predefined cases: a) ancestors accessions (A group) and accessions from Foreign breeding programs (FH group); b) ancestors accessions (A group) and accessions from Brazilian breeding programs (BB group); and c) accessions from Foreign breeding programs (FH group) and accessions from Brazilian breeding programs (BB group). PCA was performed in the R software [44] through the FactoMineR [45] and factoextra [46] packages and their respective functions PCA and fviz_pca_ind using raw data from genotyping of TRAP markers. The analysis with STRUCTURE software [47,48], to verify the number of subpopulations (*k*) and the membership proportion (*Q*), was performed considering the 248 accessions of the *Saccharum* genus of BPSG, i.e. without accessions of *Erianthus* genus. The *k* was set from 1 to 10 (*k*-value), with 10 iterations at a 100,000 burning period and 200,000 Markov Chain Monte Carlo (MCMC) repeats. The STRUCTURE HARVESTER software was used to find the best values of *k* and *Δk* [49]. Finally, the pair-wise dissimilarity among the accessions of the *Saccharum* genus was performed in the R software according to the Jaccard coefficient (Dissimilarity = 1 –Similarity) and a phylogenetic tree was build according to the neighbor-joining (NJ) method with 1,000 bootstrapping through ggtree and ape packages [50, 51]. To verify if the number of TRAP fragments used to estimate the genetic dissimilarities between accessions was adequate in terms of accuracy, the bootstrap resample technique [52] was applied as in Manechini et al. [23]. Briefly, an exponential function was adjusted to estimate the number of markers needed to assure that the CV associated with the dissimilarity estimates were lesser or equal to 10%, a threshold considered acceptable in this research. The median of the coefficient of variation estimates were used to evaluate the accuracy of the dissimilarity values [53].

Results

TRAP markers polymorphism and population differentiation

The results regarding the total number of fragments, number of polymorphic fragments, percentage of polymorphism, PIC and PD values for each of the eight TRAP markers used in this study are summarized in S4 Table. A total of 595 fragments were obtained of which 584 were polymorphic. The number of fragments for each TRAP markers ranged from 44 (SuPS + Arbi2) to 88 (SuSy + Arbi1-A) with an average of 74.37 fragments per locus. The polymorphism percentage was high (> 90%), ranging from 94.64% (SuPS + Arbi3) to 100% (SuPS + Arbi2, COMT + Arbi1-S and F5H + Arbi1-S). The averages of PIC and PD values were 0.97 and 0.98, respectively.

Putative exclusive TRAP fragments were observed for A and BB predefined groups and represented 11.64% of the total polymorphic fragments (S5 Table). The SuSy + Arbi2 showed the largest number of putative exclusive fragments (18), all present in the representative accessions of *Erianthus* spp. This specie was the one that had more putative exclusive fragments (49), followed by *S. spontaneum* (08), *S. robustum* (06), *S. officinarum* (01) and *S. barberi* (01). In the

Table 1. Analysis of molecular variance (AMOVA) between predefined groups A, BB and FH of the Brazilian Panel of Sugarcane Genotypes (BPSG).

	Source of variation	d.f.	Sum of squares	Variance components	Porcentage of variation
BB and A	Among population	1	944.93	9.28	14%
	Within populations	209	12345.46	59.07	86%
	Total	210	13290.39	68.35	
					Genetic differentiation (Φ_{PT}): 0.14*
FH and A	Among population	1	242.44	3.38	5%
	Within populations	113	7943.11	70.29	95%
	Total	114	8185.55	73.68	
					Genetic differentiation (Φ_{PT}): 0.05*
BB and FH	Among population	1	138.34	1.48	3%
	Within populations	174	8761.99	50.36	97%
	Total	175	8900.33	51.83	
					Genetic differentiation (Φ_{PT}): 0.03*

d.f.: degrees of freedom.

*$P < 0.001$.

BB group, three putative exclusive fragments were present, representing 0.51% of the total polymorphic fragments.

Considering all predefined groups (A, BB and FH), the AMOVA results revealed that the molecular variance found by TRAP markers was higher within populations than among populations (Table 1). The genetic differentiation value (Φ_{PT}) obtained between A and BB groups was 0.14, which means that 14% of the total variation found by the TRAP markers was distributed between these two groups, while 86% was within them. The Φ_{PT} values obtained between A and FH groups ($\Phi_{PT} = 0.05$) and between BB and FH groups ($\Phi_{PT} = 0.03$) were lower than that observed between A and BB groups. In addition, Φ_{PT} values were significant for all comparisons between groups ($P < 0.001$).

Principal component analysis

Principal component analysis (PCA) was firstly performed based on 595 TRAP fragments with all 254 accessions of BPSG, which includes accessions of predefined groups A, BB and FH (Fig 1A). Considering that the panel under study presents accessions of two genera, *Saccharum* and *Erianthus*, a second PCA was performed without accessions of the genus *Erianthus* (using 546 TRAP fragments) aiming to detect some clustering among the *Saccharum* accessions (Fig 1B).

Thereby, in the Fig 1A the first two principal components, PC1 and PC2, explained 17.8% of the total variability expressed among accessions. According to PC1 it is possible to note that *Erianthus* accessions (75//09 ERIANTHUS, H. KAWANDANG, IM76-227, IN84-73, IN84-77 and IN84-83) were grouped in an isolated cluster from the others accessions. In addition, *S. spontaneum* accessions were allocated together (IN84-58, IN84-82, IN84-88, KRAKATAU and SES205A). In contrast, accessions of the FH group were distributed in non-clustered way; some FH accessions were allocated near to accessions of the A group (for example, CR72/106 and US60-31-3), while others were closer to accessions of BB group (for example, NCo-310 and EK28). The BB group showed a tendency of clustered greater than A and FH groups, and it is possible to note two subgroups within the group.

Already in the second PCA, PC1 and PC2 explained 12.7% of the total variability expressed among accessions (Fig 1B). The accessions of A group were distributed over PC1, being some
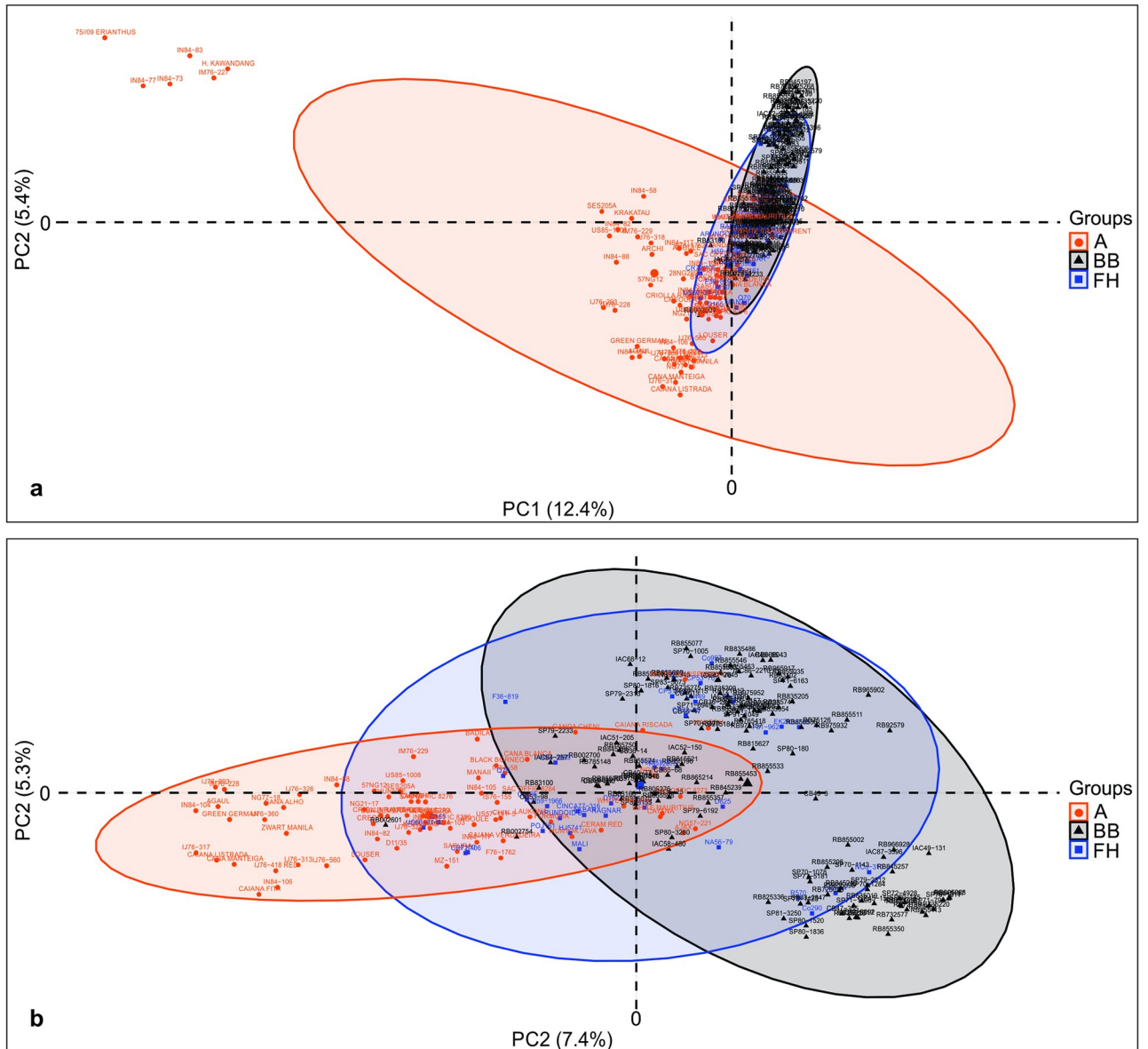
Fig 1. Principal Components Analysis (PCA) of the Brazilian Panel of Sugarcane Genotypes (BPSG) based on TRAP markers. (a) PCA performed with all 254 accessions of the BPSG. (b) PCA performed with 248 accessions of the BPSG, i.e., without accessions representatives of the genus *Erianthus*. The different colors indicates the predefined groups: ancestors accessions (A) in orange; accessions of *Saccharum* spp. hybrids from Brazilian breeding programs (BB) in black; accessions of *Saccharum* spp. hybrids from foreign breeding programs (FH) in blue. The A group in the Fig 1A was composite by ancestors accessions of the genus *Saccharum* and *Erianthus*, while in the Fig 1B, the A group was composite only by accessions of the genus *Saccharum*.

https://doi.org/10.1371/journal.pone.0233211.g001

accessions of *S. officinarum* (for example, WHITE TRANSPARENT, CAIANA RISCADA, SAC OFFIC 8272, NG21-21, NG57-221, CAYANA, WHITE MAURITIUS and AJAX) closely positioned with accessions originated from breeding programs. In addition, accessions representatives of *S. barberi* (GANDACHENI and WHITE PARARIA) and *S. sinense* (MANERIA) were also nearby of improved accessions. The accessions of FH group were distributed almost equally along PC1 and PC2, which can be observed by the blue ellipse with center near the 0–0

coordinate and also by presence of FH accessions in the four quadrants of the graph. The BB group apparently showed the division of their accessions into two clusters, one with most accessions in the second quadrant and other in the fourth quadrant of the graph. In general, this separation agrees with pedigree information, for example, the RB965917 and RB965902 accessions are full-sibs originated from cross between RB855453 and RB855536, all of them were positioned into cluster at second quadrant. Furthermore, RB845197, RB845210, RB845257, RB855036, RB855002 and RB855113 are full-sibs originated from cross between RB72454 and SP70-1143, all of them allocated into cluster at fourth quadrant. The presence of half-sibs should also contribute to this separation, for example, RB806043, RB815521, RB83102, RB855533, SP71-6163, SP716949, SP81-1763, RB815627, RB815690 and RB835054 accessions sharing the parent NA56-79 and all allocated into cluster at second quadrant. Like-wise IAC87-3396, SP83-2847, RB845197, RB845210, RB845257, RB855036, RB855002, RB855070, RB855113, RB855595 and RB855598 accessions sharing the parent SP70-1143 and all were positioned into cluster at fourth quadrant.

Structure analysis

According to the STRUCTURE analysis (without accessions of genus *Erianthus*), the best k value was two (Δk = 399.43, S1 Fig), suggesting that the 248 accessions of genus *Saccharum* could be divided into two subpopulations, P1 and P2, containing 178 and 70 accessions, respectively (Fig 2). P1 had 164 accessions belonging to BB and FH groups and only 14 accessions belonging to A group. The ancestors accessions into P1 were representatives of *S. officinarum* (AJAX, BLACK BORNEO, CAIANA RISCADA, CAYANA, CERAM RED, FORMOSA, LAUKONA, NG21-21, NG57-221, SAC OFFIC 8272, WHITE MAURITIUS and WHITE TRANSPARENT), *S. barberi* (GANDACHENI) and *S. sinense* (MANERIA). In contrast, P2 had 61 accessions belonging to A group and only nine accessions were improved accessions (AROUNDOID B, CR72/106, Q165, RB83100, RB002601 and US60-31-3, Co285, F150, HJ5741). Therefore, P1 had most of the accessions of BB and FH groups, while P2 had most of accessions of A group. Furthermore, 20 accessions showed probabilities to be part of both subpopulations (Fig 2). Among these, seven accessions were more likely to be allocated in P1 (RAGNAR, BLACK BORNEO, FORMOSA, LAUKONA, POJ161, Q70 and RB002754) and the other 13 accessions were more likely to be included in P2 (ARUNDOID B, BADILA, CAIANA VERDADEIRA, CANA BLANCA, Co285, F150, HJ5741, IS76-155, IN84-105, MANAII, Q165, RB83100 and SAC OFFIC 8284).

Genetic dissimilarity and phylogenetic analysis

The number of TRAP fragments used in this study was sufficient to estimate the pair-wise genetic dissimilarity with an acceptable level of accuracy. Considering the 546 fragments used in this analysis the CV was 8.64% (S2 Fig), under the threshold previously established of 10%. An amount around 400 fragments would be sufficient to obtain a CV average estimate around 10%.

The higher dissimilarity value was found between SES205A (*S. spontaneum*) and CAIANA FITA (*S. officinarum*) accessions (0.62), and the lower dissimilarity value was between CB40-13 and RB721012 accessions (0.10), both belonging to BB group. The average dissimilarity values within the A, BB and FH groups were 0.36, 0.25 and 0.29, respectively. Considering a sub-division of BB group according to different Brazilian breeding programs, the average dissimilarities were 0.23, 0.24, 0.26 and 0.26 within CB, IAC, RB and IAC subgroups, respectively. The highest average dissimilarities were found when A group was compared with FH group (0.34) and BB group (average of 0.34). On the other hand, smaller average dissimilarities
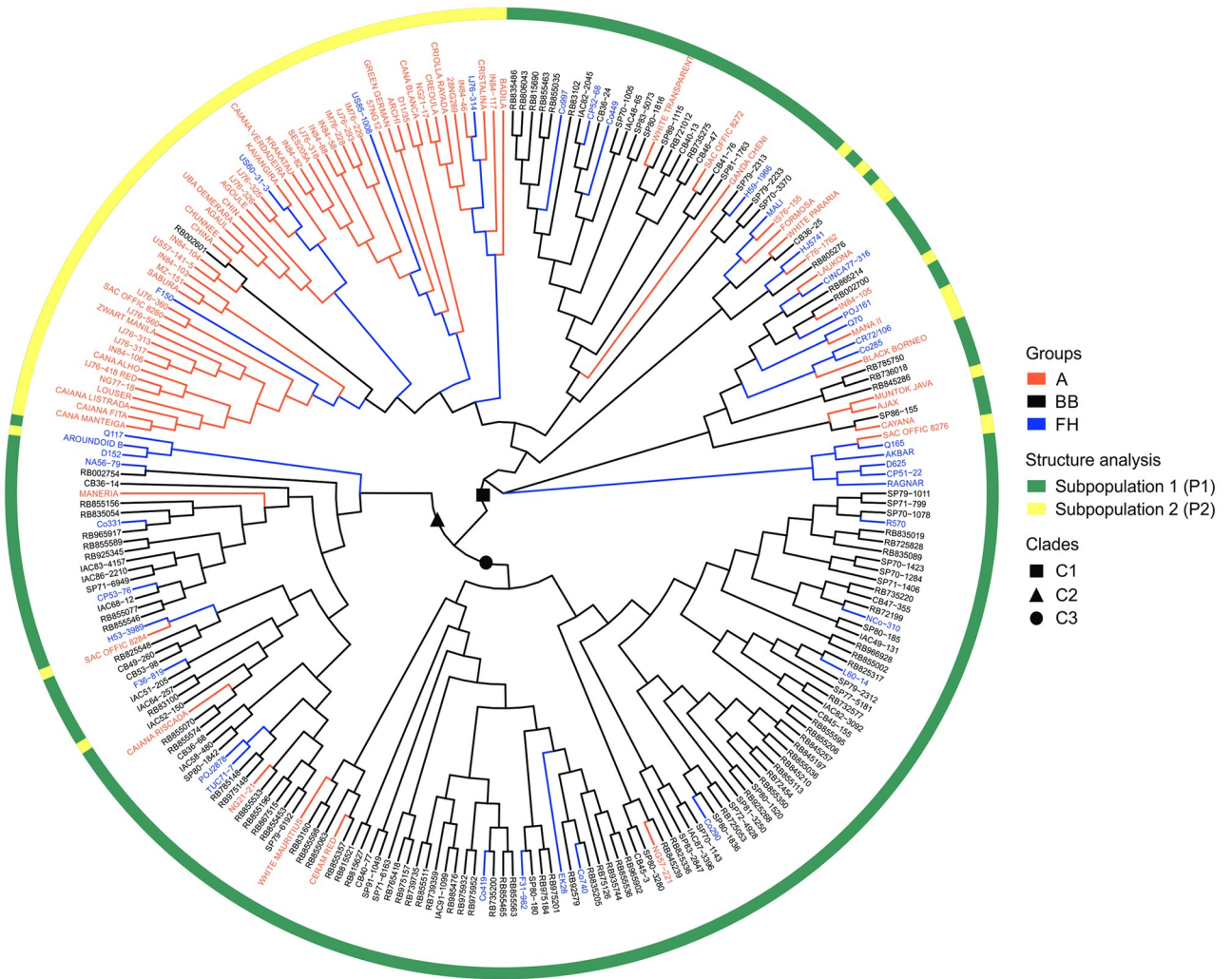
Fig 2. Phylogenetic tree estimated through neighbor-joining method for 248 accessions of the Brazilian Panel of Sugarcane Genotypes (BPSG). The names of the accessions belonging to predefined groups were write with different colors: ancestors accessions (A) in orange; accessions of *Saccharum* spp. hybrids from Brazilian breeding programs (BB) in black; and accessions of *Saccharum* spp. hybrids from foreign breeding programs (FH) in blue. The A group was composite by ancestors accessions of the genus *Saccharum*. The circumference around the phylogenetic tree represents the two subpopulations estimated by the STRUCTURE analysis and the green and yellow colors indicate accessions of the P1 and P2 subpopulations, respectively. The three major clades C1, C2 and C3 were indicated within the phylogenetic tree with square, triangle and circle in black.

https://doi.org/10.1371/journal.pone.0233211.g002

occurred between and within of the FH group and BB subgroups (CB, IAC, RB and SP), ranged from 0.23 (within CB subgroup) to 0.29 (within FH group and between FH group and SP subgroup).

The phylogenetic tree carried out with accessions of genus *Saccharum* suggests the presence of three major clades (Fig 2). The clade C1 was composed mainly by accessions of A group (68 accessions), followed by 31 BB accessions and 18 FH accessions (AKBAR, CINCA77-316, Co285, Co997, Co449, CP51-22, CP52-68, CR72/106, D625, F150, H59-1966, HJ5741, MALI, POJ161, Q70, Q165, RAGNAR and US60-31-3). On the other hand, the clades C2 and C3 were composed largely for BB accessions. Clade C2 was composed by 32 BB accessions, 10 FH accessions and four A accessions (CAIANA RISCADA, MANERIA, NG21-21 and SAC

OFFIC 8284). Finally, clade C3 had 74 BB accessions, eight FH accessions and three accessions from the A group (CERAM RED, NG57-221 and WHITE MAURITIUS).

Furthermore, the clustering of the phylogenetic tree was similar to the arrangement of accessions in the PCA (Fig 1B) and, in general, the composition of clades was also in agreement with pedigree information. Evidence of this is that some accessions that were widely used as parents by Brazilian breeding programs were grouped at same clade with their progenies of full-sibs or half-sibs. In clade C2, there are three interesting cases: a) F36-819, IAC58-480 and IAC51-205 are half-sibs with the parent POJ2878 in common; b) RB835054, RB83100, RB855533 and SP71-6949 are half-sibs with the parent NA56-79 in common; and c) RB855156, RB855196, RB855070, RB855077, RB855574, RB855589 and RB855453 are half-sibs with the parent TUC71-7 in common. In clade C3, for example, the full-sibs RB845197, RB845210, RB845239, RB845257, RB855036, RB855002, RB855113 and RB855536 were grouped together with their parents SP70-1143 and RB72454. On the other hand, although present to a lesser extent, we also noticed that some full-sib accessions were allocated in different clades and so partially diverging from pedigree information, as in the case of CB40-13 and CB40-77 accessions, being the first positioned in clade C1 and the second in clade C3. The parents POJ2878 and Co290 were positioned in clades C2 and C3, respectively.

Discussion

TRAP markers have been used in assessment of genetic diversity in plants with complex genomes such as sugarcane and wheat [21,25,29,30,54–60]. The BPSG was composed by accessions representatives of different species of the *Saccharum* complex and also by different hybrids from Brazilian and foreign breeding programs, which constitutes a broad genetic background and allelic pool to be explored and a great breeding value towards the sugarcane genetic improvement. So, the high variability and genome complexity into BPSG contributed to the large number of polymorphic fragments for each TRAP marker. The genome complexity of the modern hybrids comprises variable number of chromosomes between 100 and 130 [10,12,15,61,62], variable ploidy levels and copies of the homo(eo)logous chromosomes [12,17,63,64], gene duplication [17,64] and also genome modifications as insertions and deletions [12,34].

According to marker nature, is expected that genomic markers, such as AFLP and SSR markers, show higher polymorphism content than functional markers, since in transcribed regions the DNA sequences are more conserved [26,29]. However, PIC and DP averages values obtained in our study by functional TRAP markers were higher than related by other works in sugarcane [21,29,30,59,60] even when these values were compared with genomic markers [3,23]. Moreover, functional markers are more efficient for gene tagging than genomic markers and, consequently, facilitate the introgression of alleles that potentially control agronomic traits of interest by breeding programs [21,28,32,57]. Thereby, putative exclusive fragments for species or a specific accession could be evaluated through mapping association to further introgression process. Here, among the *Saccharum* genus, the *S. spontaneum* showed the highest number of putative exclusive fragments, all of them for TRAP makers related with sucrose metabolism (S5 Table). *S. spontaneum* is the wild species considered the most diverse species of the genus *Saccharum* due to its great ecogeographic distribution [65], show generally low sucrose levels and is used to introgress traits such as increased disease resistance, and ratooning [13]. So, our results suggest that *S. spontaneum* also could be promote variability for genes involved in sucrose metabolism and that these putative exclusive fragments probably have negative effects on the sucrose metabolism [23,24]. Furthermore, the low frequency of putative exclusive fragments in the BB group suggests that ancestor accessions did not encompass the

whole genetic pool used in prior breeding programs or that these new alleles observed in breeding accessions may have emerged over time as a result of changes in the genome as mutations and duplications [23].

The degree of the genetic differentiation (Φ_{PT}) found in the current assignment trough AMOVA are in accordance with previous studies in sugarcane [23,66,67] and also with other polyploids crops such as sweet potato, wheat and cotton [68–70]. The population differentiation depends on the balance among migration, mutation, and drift. In polyploids species, such as sugarcane, the level of diversity within populations is naturally higher when comparing with species with lower ploidy levels [71,72]. The higher differentiation among A and BB groups suggest that there was extensive use of a small number of ancestors accessions, mainly representatives of *S. officinarum* and *S. spontaneum*, in the first interspecific crosses and also that a preferential gene complexes were fixed during breeding process to develop modern Brazilian sugarcane cultivars according to yield performance interests and environmental limitations. Furthermore, mainly in the BB group, the accessions shared a larger number of parents between them [62], which contribute to increase the divergence with the A group. On the other hand, the moderate genetic differentiation among A and FH groups and the low genetic differentiation (3%) detected among BB and FH groups suggest that FH accessions have few generations from the first breeding crossings and that may be part of the genealogy of BB accessions. Indeed, the FH accessions included in the BPSG were introduced for contribute with Brazilian breeding programs [62,74].

The results of the AMOVA are interesting, since TRAP markers were partially anchored in genes under selection process (sucrose and lignin metabolism), thus not anonymous, and even were able to detect genetic differentiation within and among accessions of the compared groups. This indicates that even for these genes there is still possibility of introgression of new alleles, opening front to germplasm exchange and assisted selection with functional molecular markers, like TRAP markers, in outcrossing heterozygous species such as sugarcane. Moreover, further studies could be conducted to determine other genes under selection with potential to differentiate populations and enable better management of crosses between and within the groups for introgression of favorable alleles [13,19,28,66,73–76].

Considering the PCA approach it was possible to verify divergences between and within the predefined groups. In the first PCA (Fig 1A), into A group, the *Erianthus* accessions were clearly divergent from the *Saccharum* accessions, supporting the taxonomic evidence which assigned each of them to a separate genus [77]. Our result agrees with other studies that used AFLP [78,79], cpSSR [80], TRAP [21,30,59], SRAP [55] and SSR [3,81] markers. The introgression of alleles of the *Erianthus* genus in sugarcane breeding programs, mainly from *E. arundinaceus*, has been evaluated in recent years to increase adaptability, disease resistance, drought resistance and biomass production [82,83]. Despite this, further studies may be conducted to evaluate other regions of the genome closely related to the outstanding traits of the *Erianthus* genus.

Thereby, when we analyzed the second PCA (Fig 1B), the close position between breeding accessions and some representatives of *S. officinarum* (for example, AJAX, CAIANA RISCADA, CAYANA, NG21-21, NG57-221, SAC OFFIC 8272 and WHITE MAURITIUS) became more evident and could be explained by the fact that this specie was one of the main ancestors of modern sugarcane cultivars, which carry 80–85% of the *S. officinarum* genetic base [15]. Furthermore, the evolutionary history of sugarcane may be inferred in the clustering of the second PCA for the A group, since the *S. barberi* accessions (AGOULE, CHIN, CHUNNE, GANDA CHENI and WHITE PARARIA) were close positioned with some *S. officinarum* accessions (CAIANA RISCADA, CAIANA VERDADEIRA, CANA BLANCA, IN84-103, NG21-17, SAC OFFIC 8272, SAC OFFIC 8276, SAC OFFIC 8280 and WHITE

MAURITIUS) and some *S. spontaneum* accessions (KRAKATAU and SES205A), possibly because *S. barberi* were originated from the hybridization of *S. officinarum* with *S. spontaneum* [17,18,84–86]. The relatedness of modern sugarcane cultivars also appears to be represented in the second PCA, since some FH accessions (NA56-79, POJ2878, TUC71-7, Co290, Co331, Co413 and Co419) used as parents in crosses to obtain Brazilian cultivars were close positioned of BB accessions, for example, the FH accession TUC71-7 was near to their progenies RB855453, RB855574 and RB855196. It is interesting to note the central position of NA56-79, which was used as parent of several accessions [74] that were located into BB subgroups in the second quadrant of the graph. On the other hand, some FH accessions, for example POJ161, Co285, Q70 and US60-31-3 were found near to accessions of the A group, suggesting that this accessions could be have few generations from the crosses between the firstly ancestors (Trop-GeneDB Sugarcane: http://tropgenedb.cirad.fr/tropgene/JSP/interface.jsp?module=SUGARCANE).

When analyzed the genetic structure through STRUCTURE software, almost all ancestors were separate of the improved accessions, especially BB accessions (Fig 2). However, the STRUCTURE results should be viewed with caution, since it is based on the assumption that all loci are considered to be in Hardy-Weinberg equilibrium within each population, without any linkage disequilibrium among loci, if they are not closely linked [87]. Thus, for complex genomes such as sugarcane, these assumptions are not fulfilled, even more when are used non-neutral markers related with traits under selection during generations [24,86,88]. Nevertheless, the comparison between PCA, STRUCTURE results and also phylogenetic tree (Fig 2), showed a good way to infer the genetic structure for BPSG.

In the phylogenetic tree, which was obtained from genetic dissimilarity matrix, there was a great similarity with the clustering seen in PCA and the almost all ancestors were grouped within a cluster such as suggested by STRUCTURE analysis. In general, the family relatedness between the BPSG accessions was present in the clusters within the clades of the phylogenetic tree. In addition, the high dissimilarity value (0.62) was found between accessions representatives of *S. officinarum* (CAIANA FITA) and *S. spontaneum* (SES205A), two morphologically distinct species used in the firstly interspecific crosses of sugarcane, while the low dissimilarity value (0.10) was found between two BB accessions, CB40-13 and RB721012. Both have in their genealogies four generations and sharing at least three ancestors, since RB721012 was obtained from a polycross (RIDESA: www.ridesaufscar.com.br; TropGeneDB Sugarcane: http://tropgenedb.cirad.fr/tropgene/JSP/interface.jsp?module=SUGARCANE).

Furthermore, as expected, the highest average dissimilarities (0.34) were found when A group was compared with FH and BB groups. Otherwise, the lower average dissimilarity within the BB group (0.25) suggests that the Brazilian accessions shared approximately 75% of the genic regions assessed with TRAP markers and, therefore, a level of genetic uniformity for these loci between BB accessions. Similar results were found by Alwala et al. [21], Devarumath et al. [59] and Manechini et al. [15]. As a first approach to overcome this finding and considering that small number of initial parents contributed to modern hybrids [3], the incorporation of distinct genetic background may be useful to raise the genetic gain rate for the traits of interest, especially for those under high selection pressure. Despite this, although less frequently, some half-siblings (for example, CB40-13 and CB40-77) and full-sibs (for example, RB855589 and RB855598) were allocated to distinct clades, which is not uncommon in outcrossing heterozygous species, since they are characterized by high ploidy and may present genetic differences due to chromosomal inconsistencies during meiosis [10,17,63]. In this way, we can infer that the sugarcane genetic base did not narrow as much as some studies point out [20,32,89], since the genetic complexity mentioned above is able to promote variability even at loci that were possibly fixed by selection over decades. The results provide by AMOVA also corroborate

with these findings. The high linkage disequilibrium extend detected in sugarcane [24,86] regulates the exclusive allelic reservoir of each genotype that is transmitted to its progeny, which allowed the action of classic breeding programs to the present day. The use of molecular tools, as demonstrated in this study, can contribute to estimate genetic diversity and detected population structure in core collections, to increase the assertiveness of the crosses and efficiency of introgression of favorable alleles.

Supporting information

S1 Table. Brazilian Panel of Sugarcane Genotypes (BPSG): Accessions, pedigree information, origin and predefined groups of the 254 accessions.
(DOCX)

S2 Table. Names, sequences 5'– 3', genbank ID and the references of the fixed and arbitrary primers that compose TRAP markers.
(DOCX)

S3 Table. Functional description of the sequences that gave rise to fixed primers of TRAP markers used in this study.
(DOCX)

S4 Table. TRAP genotyping information. Total number of fragments, number of polymorphic fragments, percentage of polymorphism, polymorphism information content (PIC) value and discriminatory power (DP) value for each of the eight TRAP markers evaluated in the Brazilian Panel of Sugarcane Genotypes (BPSG).
(DOCX)

S5 Table. Putative exclusive TRAP fragments observed in the Brazilian Panel of Sugarcane Genotypes (BPSG).
(DOCX)

S1 Fig. Best k analysis showing k values from 2 to 9 (10 suppressed).
(TIF)

S2 Fig. Bootstrap analysis of TRAP genotyping. Boxplots of the coefficients of variation (CV %), associated with the estimates of genetic dissimilarities, by bootstrap analysis for subsamples with different numbers of TRAP fragments.
(TIF)

Author Contributions

Conceptualization: Monalisa Sampaio Carneiro.

Data curation: Monalisa Sampaio Carneiro.

Formal analysis: Carolina Medeiros, Thiago Willian Almeida Balsalobre, Monalisa Sampaio Carneiro.

Funding acquisition: Monalisa Sampaio Carneiro.

Investigation: Carolina Medeiros, Monalisa Sampaio Carneiro.

Methodology: Carolina Medeiros, Thiago Willian Almeida Balsalobre, Monalisa Sampaio Carneiro.

Project administration: Monalisa Sampaio Carneiro.

Supervision: Monalisa Sampaio Carneiro.

Writing – original draft: Carolina Medeiros, Thiago Willian Almeida Balsalobre.

Writing – review & editing: Carolina Medeiros, Thiago Willian Almeida Balsalobre, Monalisa Sampaio Carneiro.

References

1. Singh RK, Singh RB, Singh SP, Sharma ML. Identification of sugarcane microsatellites associated to sugar content in sugarcane and transferability to other cereal genomes. Euphytica. 2011 Jul 18; 182 (3):335–54.

2. Moore PH, Paterson AH, Tew T. Sugarcane: the crop, the plant, and domestication. In: Moore PH, Botha FC, editors. Sugarcane: Physiology, Biochemistry, and Functional Biology. Wiley-Blackwell Publishing; 2013. pp. 1–17.

3. Ali A, Pan Y-B, Wang Q-N, Wang J-D, Chen J-L, Gao S-J. Genetic diversity and population structure analysis of *Saccharum* and *Erianthus* genera using microsatellite (SSR) markers. Sci Rep. 2019 Jan 23; 9(1):1–10.

4. Daniels J, Roach BT. Taxonomy and evolution. In: Heinz DJ, editor. Sugarcane Improvement through breeding. Elsevier Science; 1987. pp. 7–84.

5. Matsuoka S, Garcia AAF, Arizono H. Melhoramento da cana-de-açúcar. In: Borém A, editor. Melhoramento de espécies cultivadas. UFV; 2005. pp. 225–274.

6. Zhang J, Zhou M, Walsh J, Zhu L, Chen Y, Ming R. Sugarcane genetics and genomics. In: Moore PH, Botha FC, editors. Sugarcane: Physiology, Biochemistry and Functional Biology. Wiley-Blackwell Publishing; 2013. pp. 623–643.

7. Todd J, Wang J, Glaz B, Sood S, Ayala-Silva T, Nayak SN, et al. Phenotypic characterization of the Miami World Collection of sugarcane (*Saccharum* spp.) and related grasses for selecting a representative core. Genet Resour Crop Ev. 2014 Jun 6; 61(8):1581–96.

8. Irvine J. E. *Saccharum* species as horticultural classes. Theor Appl Genet. 1999 Feb; 98(2):186–94.

9. Grivet L, Daniels C, Glaszmann JC, D'Hont A. A review of recent molecular genetics evidence for sugarcane evolution and domestication. Ethnobot Res App. 2004; 2:9–17.

10. Vieira MLC, Almeida CB, Oliveira CA, Tacuatiá LO, Munhoz CF, Cauz-Santos L, et al. Revisiting Meiosis in Sugarcane: Chromosomal Irregularities and the Prevalence of Bivalent Configurations. Front Genet. 2018 Jun 14; 9:213. https://doi.org/10.3389/fgene.2018.00213 PMID: 29963076

11. D'Hont A, Glaszmann JC. Sugarcane genome analysis with molecular markers, a first decade of research. Proc Int Soc Sugarcane Technol. 2001; 24:556–9.

12. Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, et al. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. Nature genet. 2018 Nov; 50(11):1565–73. https://doi.org/10.1038/s41588-018-0237-2 PMID: 30297971

13. Aitken K, Li J, Piperidis G, Qing C, Yuanhong F, Jackson P. Worldwide genetic diversity of the wild species *Saccharum spontaneum* and level of diversity captured within sugarcane breeding programs. Crop Sci. 2018 Jan 1; 58(1):218–29.

14. D'Hont A. Unravelling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. Cytogenet Genome Res. 2005; 109(1–3):27–33. https://doi.org/10.1159/000082378 PMID: 15753555

15. Garsmeur O, Droc G, Antonise R, Grimwood J, Potier B, Aitken K, et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat Commun. 2018 Jul 6; 9(1):1–10.

16. Ma P, Yuan Y, Shen Q, Jiang Q, Hua X, Zhang Q, et al. Evolution and Expression Analysis of Starch Synthase Gene Families in *Saccharum spontaneum*. Trop Plant Biol. 2019 May 14; 12(3):158–73.

17. Thirugnanasambandam PP, Hoang NV, Henry RJ. The challenge of analyzing the sugarcane genome. Front Plant Sci. 2018 May 14; 9:1–18.

18. Singh RB, Singh B, Singh RK. Evaluation of Genetic Diversity in *Saccharum* Species Clones and Commercial Varieties Employing Molecular (SSR) and Physiological Markers. Indian J Plant Genet Res. 2018 Apr 9; 31(1):17–26.

19. Llerena JPP, Figueiredo R, Brito MS, Kiyota E, Mayer JLS, Araujo P, et al. Deposition of lignin in four species of *Saccharum*. Sci Rep. 2019 Apr 10; 9(1):1–19.

20. Lima MLA, Garcia AAF, Oliveira KM, Matsuoka S, Arizono H, De Souza CL Jr, et al. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugar cane (*Saccharum*

spp.). Theor Appl Genet. 2002 Jan; 104(1):30–8. https://doi.org/10.1007/s001220200003 PMID: 12579425

21. Alwala S, Suman A, Arro JA, Veremis JC, Kimbeng CA. Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. Crop Sci. 2006 Jan 1; 46 (1):448–55.

22. Dos Santos JM, Filho LSC, Soriano ML, Silva PP, Nascimento VX, Barbosa GVS, et al. Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. Ind Crop Prod. 2012 Nov; 40:145–50.

23. Manechini JRV, Costa JB, Pereira BT, Carlini-Garcia L, Xavier MA, Landell MGA, et al. Unraveling the genetic structure of Brazilian commercial sugarcane cultivars through microsatellite markers. PLoS One. 2018 Apr 23; 13(4):e0195623. https://doi.org/10.1371/journal.pone.0195623 PMID: 29684082

24. Barreto FZ, Rosa JRBF, Balsalobre TWA, Pastina MM, Silva RR, Hoffmann HP, et al. A genome-wide association study identified loci for yield component traits in sugarcane (*Saccharum* spp.). PloS One. 2019 Jul 18; 14(7):e0219843. https://doi.org/10.1371/journal.pone.0219843 PMID: 31318931

25. Singh RB, Singh B, Singh RK. Study of genetic diversity of sugarcane (*Saccharum*) species and commercial varieties through TRAP molecular markers. Indian J Plant Physiol. 2017 Oct 19; 22(3):332–8.

26. Parthiban S, Govindaraj P, Senthilkumar S. Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. 3 Biotech. 2018 Feb 21; 8(3):1–12.

27. You Q, Pan Y-B, Gao S-W, Wang Q-N, Su Y-C, Yang Y-Q, et al. Genetic Diversity Analysis of Sugarcane Germplasm Based on Fluorescence-Labeled Simple Sequence Repeat Markers and a Capillary Electrophoresis-based Genotyping Platform. Sugar Tech. 2016 Aug 7; 18(4):380–90.

28. Tienderen PH, de Haan AA, van der Linden CG, Vosman B. Biodiversity assessment using markers for ecologically important traits. Trends Ecol Evolut. 2002 Dec; 17(12):577–82.

29. Creste S, Accoroni KAG, Pinto LR, Vencovsky R, Gimenes MA, Xavier MA, et al. Genetic variability among sugarcane genotypes based on polymorphisms in sucrose metabolism and drought tolerance genes. Euphytica. 2010 Nov 25; 172(3):435–46.

30. Suman A, Ali K, Arro J, Parco AS, Kimbeng CA, Baisakh N. Molecular diversity among members of the *Saccharum* complex assessed using TRAP markers based on lignin-related genes. Bioenergy Res. 2012 Apr 7; 5(1):197–205.

31. Kasirajan L, Hoang NV, Furtado A, Botha F, Henry RJ. Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. Sci Rep. 2018 Aug 2; 8(1):1–12.

32. Da Silva JA, Veremis J, Solís-Gracia N. *Saccharum spontaneum* gene tagging by markers developed from sugarcane expressed sequence tags. Subtrop Plant Sci. 2007; 58(2):6–14.

33. Souza CM. Molecular diversity and genetic structure of Saccharum complex accessions as revealed by TRAP markers. Repositório UFSCar. https://repositorio.ufscar.br/handle/ufscar/11771 Cited 23 March 2020.

34. Balsalobre TWA, Pereira GS, Margarido GRA, Gazaffi R, Barreto FZ, Anoni CO, et al. GBS-based single dosage markers for linkage and QTL mapping allow gene mining for yield-related traits in sugarcane. BMC genomics. 2017 Jan 11; 18(1):1–19.

35. Cardoso-Silva CB, Costa EA, Mancini MC, Balsalobre TWA, Canesin LEC, Pinto LR, et al. De novo assembly and transcriptome analysis of contrasting sugarcane varieties. PLoS One. 2014 Feb 11; 9(2): e88462. https://doi.org/10.1371/journal.pone.0088462 PMID: 24523899

36. Al-Janabi SM, Forget L, Dookun A. An improved and rapid protocol for the isolation of polysaccharide and polyphenol-free sugarcane DNA. Plant Mol Biol Report. 1999 Sep; 17(3):1–8.

37. Li G, Quiros CF. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. Theor Appl Genet. 2011 Aug; 103(2–3):455–61.

38. Creste S, Neto AT, Figueira A. Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Mol Biol Report. 2001 Sep 3; 19(4):299–306.

39. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980 May; 32(3),314–31. PMID: 6247908

40. Tessier C, David J, This P, Boursiquot JM, Charrier A. Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera L*. Theor Appl Genet. 1999 Jan; 98(1):171–7.

41. Oliveira KM, Pinto LR, Marconi TG, Mollinari M, Ulian EC, Chabregas SM, et al. Characterization of new polymorphic functional markers for sugarcane. Genome. 2009 Feb; 52(2):191–209. https://doi.org/10.1139/g08-105 PMID: 19234567

42. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 2012 Jan; 40:1178–86.

43. Peakall PE, Smouse R. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012 Dec 21; 28(1):2537–39.

44. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2016.

45. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw. 2008; 25:1–18.

46. Kassambara A, Mundt F. Factoextra: Extract and visualize the results of multivariate data analyses. R package version. 2016; 76:1–74.

47. Pritchard JK, STephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000 Jun; 155(2):945–59. PMID: 10835412

48. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005 Jul; 14(8):2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x PMID: 15969739

49. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012 Oct 13; 4(2):359–61.

50. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol. 2017; 8(1):28–36.

51. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2018 Jul 17; 35(3):526–8.

52. Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: Chapman & Hall; 1993.

53. Garcia AAF, Benchimol LL, Barbosa AMM, Geraldi IO, Júnior CLS, Souza AP. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. Genet Mol Biol. 2004; 27(4):579–88.

54. Alwala S, Kimbeng CA, Gravois KA, Bischoff KP. TRAP, a new tool for sugarcane breeding: comparison with AFLP and coefficient of parentage. J Am Soc Sugar Cane Technol. 2006 Jan; 26:62–86.

55. Suman A, Kimbeng CA, Edmé SJ, Veremis J. Sequence-related amplified polymorphism (SRAP) markers for assessing genetic relationships and diversity in sugarcane germplasm collections. Plant Genet Resour. 2008 Dec; 6(3):222–31.

56. Que Y, Chen T, Xu L, Chen R. Genetic diversity among key sugarcane clones revealed by TRAP markers. J Agric Biotechnol. 2009 Jan; 17(3):496–503.

57. Creste S, Sansoli DM, Tardiani ACS, Silva DN, Gonçalves FK, Fávero TM, et al. Comparison of AFLP, TRAP and SSRs in the estimation of genetic relationships in sugarcane. Sugar Tech. 2010 Oct 22; 12 (2),150–4.

58. Al-Doss AA, Saleh M, Moustafa KA, Elshafei AA, Barakat MN. Grain yield stability and molecular characterization of durum wheat genotypes under heat stress conditions. Afr J Agric Res. 2010 Dec; 5 (22):3065–74.

59. Devarumath RM, Kalwade SB, Bundock P, Eliott FG, Henry R. Independent target region amplification polymorphism and single-nucleotide polymorphism marker utility in genetic evaluation of sugarcane genotypes. Plant Breed. 2013 Sep 13; 132(6):736–47.

60. Farsangi FJ, Thorat AS, Devarumath RM. Assessment of the Utility of TRAP and EST-SSRs Markers for Genetic Diversity Analysis of Sugarcane Genotypes. Cytol Genet. 2018 Dec 18; 52(6):467–77.

61. Cheavegatti-Gianotto A, de Abreu HM, Arruda P, Bespalhok Filho JC, Burnquist WL, Creste S, et al. Sugarcane (Saccharum X officinarum): A Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. Trop Plant Biol. 2011 Mar; 4(1):62–89. https://doi.org/10.1007/s12042-011-9068-3 PMID: 21614128

62. de Morais LK, Aguiar MS, Silva PA, Câmara TMM, Cursi DE, Júnior ARF, et al. Breeding of sugarcane. In: Cruz VMV, Dierig DA, editors. Industrial crops: breeding for bioenergy and bioproducts. New York, NY: Springer; 2015. pp. 29–42.

63. Garcia A, Mollinari M, Marconi T, Serang OR, Silva R, Vieira MLC, et al. SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. Sci Rep. 2013 Dec 2; 3:1–10.

64. Sforça DA, Vautrin S, Cardoso-Silva CB, Mancini MC, Cruz MVR, Pereira GDS, et al. Gene Duplication in the Sugarcane Genome: A Case Study of Allele Interactions and Evolutionary Patterns in Two Genic Regions. Front Plant Sci. 2019 May 7; 10:553. https://doi.org/10.3389/fpls.2019.00553 PMID: 31134109

65. Liu P, Chandra A, Que Y, Chen P-H, Grisham MP, White WH, et al. Identification of quantitative trait loci controlling sucrose content based on an enriched genetic linkage map of sugarcane (*Saccharum* spp. hybrids) cultivar 'LCP 85–384'. Euphytica. 2016 Sep 14; 207(3):527–49.

66. Glynn NC, McCorkle K, Comstock JC. Diversity among mainland USA sugarcane cultivars examined by SSR genotyping. J Am Soc Sugar Cane Technol. 2009; 29:36–52.

67. Tazeb A, Haileselassie T, Tesfaye K. Molecular characterization of introduced sugarcane genotypes in Ethiopia using inter simple sequence repeat (ISSR) molecular markers. Afr J Biotechnol. 2017 Mar 8; 16(10):434–49.

68. Su W, Wang L, Lei J, Chai S, Liu Y, Yang Y, et al. Genome-wide assessment of population structure and genetic diversity and development of a core germplasm set for sweet potato based on specific length amplified fragment (SLAF) sequencing. PloS One, 2017 Feb 10; 12(2):e0172066. https://doi.org/10.1371/journal.pone.0172066 PMID: 28187178

69. Eltaher S, Sallam A, Balemkar V, Emara HA, Nower AA, Salem KFM, et al. Genetic diversity and population structure of F3: 6 Nebraska winter wheat genotypes using genotyping-by-sequencing. Front Genet. 2018 Mar 12; 9:76. https://doi.org/10.3389/fgene.2018.00076 PMID: 29593779

70. Seyoum M, Du XM, He SP, Jia YH, Pan Z, Sun JL. Analysis of genetic diversity and population structure in upland cotton (*Gossypium hirsutum* L.) germplasm using simple sequence repeats. J Genet. 2018 Jun; 97(2):513–22. PMID: 29932072

71. Meirmans PG, Liu S. Analysis of Molecular Variance (AMOVA) for autopolyploids. Front Ecol Evol. 2018 May 23; 6:66.

72. Meirmans PG, Liu S, van Tienderen PH. The analysis of polyploid genetic data. J Hered. 2018 Jan 27; 109(3):283–96. https://doi.org/10.1093/jhered/esy006 PMID: 29385510

73. Junior CADK, Manechini JRV, Corrêa RX, Pinto ACR, Costa JB, Favero TM, et al. Genetic structure analysis in sugarcane (Saccharum spp.) using target region amplification polymorphism (TRAP) markers based on sugar- and lignin-related genes and potential application in core collection development. Sugar Tech. 2020. https://doi.org/10.1007/s12355-019-00791-0.

74. Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann HP, Garcia AA, et al. Sugarcane improvement: how far can we go? Curr Opin Biotechnol. 2012 Apr; 23(2):265–70. https://doi.org/10.1016/j.copbio.2011.09.002 PMID: 21983270

75. Park JW, Solís-Gracia N, Trevino C, Da Silva JA. Exploitation of conserved intron scanning as a tool for molecular marker development in the *Saccharum* complex. Mol Breed. 2011 Dec 14; 30(2):987–99.

76. Thirugnanasambandam PP, Mason PJ, Hoang NV, Furtado A, Botha FC, Henry RJ. Analysis of the diversity and tissue specificity of sucrose synthase genes in the long read transcriptome of sugarcane. BMC Plant Biol. 2019 Apr 25; 19(1):160. https://doi.org/10.1186/s12870-019-1733-y PMID: 31023213

77. Daniels J, Smith P, Paton N, Williams CA. The origin of the genus *Saccharum*. Sugarcane breeding newsletter. 1975; 36:24–39.

78. Cai Q, Aitken KS, Fan YH, Piperidis G, Jackson P, McIntyre CL. A preliminary assessment of the genetic relationship between *Erianthus rockii* and the "*Saccharum complex*" using microsatellite (SSR) and AFLP markers. Plant Sci. 2005 Nov; 169(5):976–84.

79. Selvi A, Nair NV, Noyer JL, Singh NK, Balasundaram N, Bansal KC, et al. AFLP analysis of the phenetic organization and genetic diversity in the sugarcane complex, *Saccharum* and *Erianthus*. Genet Resour Crop Evol. 2006 Feb 27; 53(4):831–42.

80. Raj P, Selvi A, Prathima PT, Nair NV. Analysis of Genetic Diversity of *Saccharum* Complex Using Chloroplast Microsatellite Markers. Sugar Tech. 2016 May 19; 18(2):141–8.

81. Nayak SN, Song J, Villa A, Bhuvan P, Ayala-Silva T, Yang X, et al. Promoting utilization of *Saccharum* spp. genetic resources through genetic diversity analysis and core collection construction. PloS One. 2014 Oct 21; 9(10):e110856. https://doi.org/10.1371/journal.pone.0110856 PMID: 25333358

82. Chen JW, Lao F-y, Chen X-w, Deng H-h, Liu R, He H-y, et al. DNA marker transmission and linkage analysis in populations derived from a sugarcane (*Saccharum* spp.) x *Erianthus arundinaceus* hybrid. PloS One. 2015 Jun 8; 10(6):e0128865. https://doi.org/10.1371/journal.pone.0128865 PMID: 26053338

83. Yang S, Zeng K, Wu J, Wang Q, Li X, Deng Z, et al. Chromosome transmission in BC 4 progenies of intergeneric hybrids between *Saccharum* spp. and *Erianthus arundinaceus* (Retz.) Jeswiet. Sci Rep. 2019 Feb 21; 9(1):1–8.

84. Lu YH, D'Hont A, Walker DIT, Rao PS, Feldmann P, Glaszmann JC. Relationships among ancestral species of sugarcane revealed with RFLP using single copy maize nuclear probes. Euphytica. 1994 Jan; 78(1–2):7–18.

85. Metcalfe CJ, Oliveira SG, Gaiarsa JW, Aitken KS, Carneiro MS, Barreto FZ, et al. Using quantitative PCR with retrotransposon-based insertion polymorphisms as markers in sugarcane. J Exp Bot. 2015 Jun 19; 66(14):4239–50. https://doi.org/10.1093/jxb/erv283 PMID: 26093024

86. Yang X, Luo Z, Todd J, Sood S, Wang J. Genome-wide association study of multiple yield components in a diversity panel of polyploid sugarcane (*Saccharum* spp.). bioRxiv. 2018;387001.

87. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes. 2007 Jul 1; 7(4):574–8. https://doi.org/10.1111/j.1471-8286.2007.01758.x PMID: 18784791

88. Racedo J, Gutierrez L, Perera MF, Ostengo S, Pardo EM, Cuenya MI, et al. Genome-wide association mapping of quantitative traits in a breeding population of sugarcane. BMC Plant Biol. 2016 Jun 24; 16 (1):142. https://doi.org/10.1186/s12870-016-0829-x PMID: 27342657

89. Aitken KS, Li JC, Jackson P, Piperidis G, McIntyre CL. AFLP analysis of genetic diversity within *Saccharum officinarum* and comparison with sugarcane cultivars. Aust J Agric Res. 2006 Oct 27; 57 (11):1167–84.