PHILIPP AGRIC SCIENTIST
Vol. 104 No. 3, 223-237
September 2021

ISSN 0031-7454

Genetic Diversity and Population Structure of Sugarcane (*Saccharum officinarum* L.) Collected from Seven Philippine Regions for Use as Breeding Stocks for Crop Improvement

Jan Leandro L. Acedera[1], Jhun Laurence S. Rasco[1], Jemimah C. Banganan[1], and Antonio G. Lalusin[2]

[1]Institute of Plant Breeding, College of Agriculture and Food Science, University of the Philippines Los Baños, Los Baños, Laguna, Philippines
[2]Institute of Crop Science, College of Agriculture and Food Science, University of the Philippines Los Baños, Los Baños, Laguna, Philippines

[*]Author for correspondence; E-mail: jlacedera@up.edu.ph

Received: January 06, 2021/ Revised: May 26, 2021/ Accepted: August 08, 2021

Selection of parents for hybridization relies on the information on their genetic relationship and diversity which are essential in any breeding program. This study aimed to estimate the extent of genetic diversity and population structure of 76 sugarcane accessions from seven regions in the Philippines using 57 morphological characters and 50 microsatellite markers. The sugarcane collections exhibited moderate to high diversity with mean of H' = 0.72 for qualitative and H' = 0.75 for quantitative morphological characters, respectively. This is corroborated by the analysis of variance (ANOVA) of agronomic parameters, except for stalk length. Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis based on morphology subdivided the accessions into 31 clusters which reveal phenotypic variability among sugarcane samples. The fingerprints of the 76 accessions were also evaluated using 45 *Saccharum*-based genomic SSR and 5 EST-SSR primer pairs to measure genetic diversity and population structure. Based on UPGMA, a total of six clusters were generated with a 0.65 coefficient of dissimilarity, and the sugarcane accessions were further subdivided into five major sub-populations. Out of 50 markers used, 41 (82%) were found to be highly informative with a mean PIC value of 0.69. It was also expected that *Saccharum*-based genomic SSRs were more polymorphic (92%) compared to EST-SSRs (82%) since the latter preferably amplify in more conserved and expressed sequences in the genome. Out of 2,850 pairwise combinations based on the Jaccard coefficient index, large diverse parental combinations (genetic dissimilarity= 0.51 - 0.70) were observed, indicating substantial diversity from the existing breeding pool of IPB-UPLB for genetic improvement. Cluster analysis based on UPGMA, STRUCTURE analysis, and Principal Coordinate Analysis (PCoA) were predominantly consistent. However, no association was observed between the geographical origin and genetic distance of the genotypes based on molecular data. The results showed that accessions were grouped into five sub-populations and genetic differentiation within sub-population was higher (85%) as compared to between sub-population (15%) based on the analysis of molecular variance (AMOVA), suggesting an active exchange of genetic pool across provinces and regions of both Luzon and Visayas islands. The findings from this study will be useful for future breeding efforts by exploiting genetic variation existing in the current breeding population.

Keywords: sugarcane, SSR, genetic diversity, population structure, germplasm

Abbreviations: SSR - simple sequence repeats, EST - expressed sequenced tag, NSIC - National Seed Industry Council, Philsurin - Philippine Sugar Research Institute, PIC - polymorphic information content

INTRODUCTION

Sugarcane is a widely cultivated cash crop in the tropical and subtropical regions and is the main source of commercial sugar contributing to 75% of global production (Sharma et al. 2014). Interest in its by-products and co-products has been increasing and has considerable effects on the sugar industry (Meyer et al. 2011). The Philippine sugarcane production from October to December 2019 was down by 14.8% compared to 4th quarter of 2018 (7-03 million metric tons - 5.99 million metric tons) where the largest producer was recorded in

Western Visayas (4.9 million metric tons) accounting for 81.8 % of the total production followed by Northern Mindanao and Central Visayas at 8.2% and 5.9% share, respectively (Philippine Statistics Authority 2019). At present, the Philippines has a low sugar yield at 5.1 tons of sugar per hectare and sugar recovery per ton of milled cane which is comparatively lower than other countries such as Brazil, Australia, and Thailand with an average of 38.6%. Factors such as cost of production from inputs, interest rates and labor, low yield and low market price that leads to small income of farmers labor shortage especially during harvesting, and adverse effect of climate change (El Niño, La Niña, and typhoons) limit productivity of sugarcane industry in the country (Mendoza 2020).

Commercially available modern sugarcane varieties are highly complex polyploids and heterozygous in nature with a chromosome number range of 100-130 (D'Hont et al. 1996). It was known to be derived from a few interspecific hybridizations between *S. Saccharum officinarum* (2n = 80) and *Saccharum spontaneum* (2n = 40-128) (Aitken et al. 2006). This hybridization event led to high linkage disequilibrium and narrow genetic base of most breeding stocks being used in breeding programs around the world (Raboin et al. 2008; Alwala et al. 2009). This adds to the breeding difficulty of the crop in increasing sugar content and improving cane yield aside from long breeding and selection cycle that gain a lot of concerns to all sugarcane breeders. Therefore, the choice of parents and germplasm diversity are the key components of a successful breeding program (Kawar et al. 2009; Sharma et al. 2014). Plant performance such as yield, adaptability, and disease resistance is not only the main focus in parental selection but also the genetic diversity of both parents (Hapsoro et al. 2015). Genetically diverse parents may be selected based on the geographical distribution of the genotypes, agronomic characters, pedigree, and diversity analysis data generated from molecular markers (Melchinger 1998). Utilization of introduced germplasm and the knowledge of genetic remoteness among them are crucial especially on locally collected germplasm, as they can provide adapted genes for better crop improvement (Malik et al. 2010; Gashaw et al. 2016). Therefore, a sufficient understanding of the genetic diversity is essential in creating new crosses and selection of best individual clones as parental materials (Alwala et al. 2006; dos Santos et al. 2012). Traditional ways to investigate genetic variation rely on agronomic and morphological characters (Skinner et al. 1972). However, some of the characters are greatly influenced by various environmental factors creating continuous variation and degree of plasticity leading to the false positive measurement of genetic diversity (Lima et al. 2002). This makes the use of genetic markers very important since they are stable and can efficiently gauge the genetic diversity of the crop. Several molecular techniques were used in studying genetic diversity in sugarcane within and between members of the *Saccharum* complex. Among the PCR-based markers, microsatellites or simple sequence repeats (SSRs) have gained considerable importance in the field of plant genetics and breeding (Kalia et al. 2011) and have been extensively used in different genetic diversity studies in plants due to its various attributes. This includes high variability, wide genomic distribution, co-dominant inheritance, high reproducibility, large multi-allelic nature, and chromosome-specific location (Singh et al. 2010).

Currently, there are only few studies on intensive assessment of genetic diversity of locally collected germplasm in the Philippines using *Saccharum*-based and Expressed Sequenced Tag (EST) microsatellites and morphological characters. The efficient utilization of this introduced germplasm and a better estimate of the available genetic diversity among accessions are requisite to the current sugarcane breeding program of the Institute of Plant Breeding-University of the Philippines Los Baños (IPB-UPLB) with the collaboration of Sugar Regulatory Administration (SRA). The main objectives of the study were to (1) estimate the morphological diversity among sugarcane accessions based on 46 qualitative and 11 quantitative traits, (2) evaluate the genetic diversity and population structure of this collection by using 50 sets of SSR primer pairs and (3) to identify the group of accessions as parental for future breeding efforts.

MATERIALS AND METHODS

Plant Materials

A total of 76 sugarcane accessions – 65 collected from seven Philippine regions, and ten Sugar Regulatory Administration (SRA) commercial checks and one variety developed by the Philippine Sugar Research Institute (Philsurin) (Figure 1) were planted at the experimental field of the Institute of Plant Breeding (IPB), College of Agriculture and Food Science (CAFS), in Brgy. Tranca Bay, Laguna, Philippines from May 2017 to May 2018. The collected sugarcane accessions per region were as follows: CAR-1, Region I – 3, Region II – 4, Region III – 14, Region Iva – 17, Region V – 9 and Region VIII – 27.

Experimental Design and Layout

The experiment was laid out in Randomized Complete Block Design (RCBD) with subsampling. Cane points

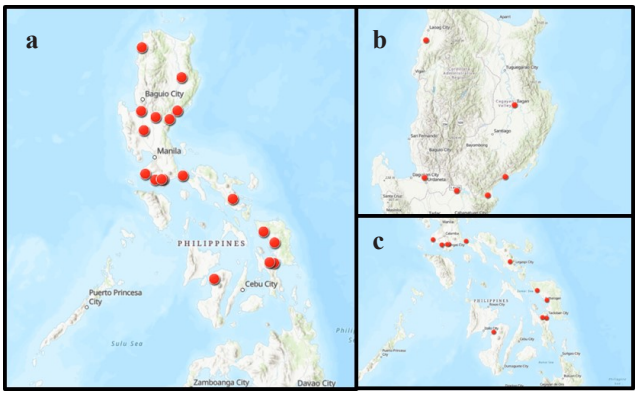

Fig. 1. Philippine map showing the location of the origin of the collected sugarcane (*Saccharum officinarum* L.) accessions. (a) Phillippines (b) Luzon island (c) Visayas islands.

were cut into two budded setts and planted to two rows and five hills in three replicates following the National Seed Industry Council (NSIC) guideline on sugarcane production. Standard practices on field maintenance activities i.e. weeding, fertilizer application, and irrigation were performed.

Genomic DNA Extraction, Quantification, and Amplification

Young leaf tissues of the 76 sugarcane accessions at approximately 4 months after planting were collected and extracted at the Molecular Plant Breeding Laboratory - Institute of Crop Science (ICropS), College of Agriculture and Food Science (CAFS), UPLB using the modified CTAB extraction protocol of Doyle and Doyle (1987). The DNA samples were quantified using NanoDrop (NanoDrop 1000 Spectrophotometer™, Thermo Fisher Scientific Inc., Wilmington, USA) and the quality was checked using agarose gel electrophoresis set-up in a 1% UltraPure Agarose in 1X Tris-borate EDTA (TBE) running buffer, at a constant power of 100V for approximately 20 minutes.

Forty-five *Saccharum*-based genomic SSR markers (Sanguillosa et al. 2019) and five (5) developed EST-primers were utilized in the study (Table 1). A 15 µl PCR reaction was prepared containing 100 ng/µl of template DNA, 1X PCR buffer (10 mM Tris-HCL, *pH* 8.3, 50 mM KCl), 2 mM $MgCl_2$, 0.15 nM dNTPs, and 0.2 µM primers (both forward and reverse) and 0.05 U of Taq DNA polymerase. Amplification was done using optimized PCR cycle conditions: initial denaturation for 5 min. at 95° C, followed by 35 cycles of 10 sec. denaturation at 94°C, 10 sec. annealing at 55°C and 30-sec elongation at 72°C, with final elongation at 72°C for 5 min. soaking temperature is set at 4°C. Amplified fragments were resolved on 8% polyacrylamide gel which was run at 100V for 4 hours in 1X TBE using ClearPAGETM (C.B.S. Scientific Co., Del Mar, CA). All gels were stained with GelRed (Biotium, CA) and the gel image was captured using GenoSens 1000 Gel documentation system (Clinx Science Instruments Co., Ltd).

Genotyping

The polyploid nature of sugarcane which made scoring the alleles difficult, each band resulting from the polyacrylamide gel electrophoresis was considered as a locus with two alternative alleles and scored as 1 if present and 0 if absent. The size of the amplified bands on each run was estimated using Invitrogen™ 100bp DNA ladder (Thermo Fisher Scientific Inc., Wilmington, USA). The gel was then scored using gen analyzer software (GenoSens 1000 Gel documentation system).

Data Analysis

Morphological Characterization

Accessions were characterized morphologically using standardized sugarcane descriptor list developed by PHILSURIN and IPB (2002) sugarcane descriptors described by Moore (1987). Data were gathered at three (3) plant stages in 5 months, 7-8 months, and 10-12 months after planting following the standard genebank characterization protocol. The agronomic components (quantitative traits) were analyzed and interpreted using ProcANOVA (Analysis of Variance) of SASv9.0 (SAS Institute Inc. 2002). To further characterize the phenotypic diversity of all accessions, phenotypic frequency distribution of each qualitative and quantitative characters were used to compute the Shannon Weaver Diversity Index (H') using the formula:

$$H' = -\sum_{i=1}^{n} p_i (log_2 p_i)$$

where p_i is the proportion of the total number of entries belonging to the i^{th} class or phenotype (I= 1, 2, …, n). Each H' value was divided by its maximum value (log⊚n) and normalized to keep the values between 0 and 1 (Gashaw et al. 2016). The arbitrary rating scale adapted from Canepoint (Altoveros et al. 2006): maximum diversity (H' = 1.00), high diversity (H' = 0.67 - 0.99), moderate diversity (H' = 0.34 - 0.66), and poor or low diversity (H' = 0.00 - 0.33), was used to classify H' value on each phenotypic characters measured. The coefficient of dissimilarity between each pair of accessions was calculated using distance method correlation in R package PVCLUST. The generated genetic distance matrix was then used for cluster analysis with the Unweighted Pair Group Method based on Arithmetic Averages (UPGMA) using Rstudio v1.1.423 (R Core Team 2017).

Genetic Diversity

The data were treated as a dominant marker to create a binary matrix that serves as an input file for the analysis. The Polymorphism Information Content (PIC) was taken as a parameter to evaluate the discriminatory power of the 50 microsatellites set to recognize polymorphism and the ability to differentiate sugarcane genotypes into clusters. PIC value was calculated using the formula:

$$PIC_i = 1 - \sum_{j=1}^{n} p_i^2$$

where p_i is the frequency of allele i raised to the second power for the j^{th} marker (Nei 1973). For each SSR primer pair, the number of alleles per locus and the number of polymorphic alleles were also explored.

The genetic dissimilarity among all possible genotype combination was estimated using R function *dist* based on Jaccard's coefficient (Jaccard 1901) from the R package CLUSTER (Maechler et al. 2012). Jaccard coefficient has the advantage of disregarding the shared absence of bands in the paired comparison, thereby reducing the risk of similarity overestimation between samples (Clifford and Stephenson 1975). The dendrogram was constructed according to the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) with bootstrap analysis (nboot = 1000 permutations) using package PVCLUST of RStudio v1.1.423 (R Core Team 2017).

Population Differentiation and Structure

Analysis of the population structure of 76 sugarcane genotypes was estimated under the Markov Chain Monte Carlo (MCMC) algorithm with the Bayesian model-based clustering method implemented in STRUCTURE software v2.3.4 (Pritchard et al. 2000). For each analysis, given no prior population information, K value was set from 2 to 8 with five iterations for each value of K. For each run, the length of the burnin period was set to 50,000 with 200,000 MCMC iterations. The final K value was then estimated using both Delta (Δ) K calculated using the online program STRUCTURE Harvester (Earl 2012) and the maximum value of L(K) against the given K value. The gene flow (N_m) or the effective number of migrants entering a population per generation was also calculated using the formula $N_m = 0.25 \times (1 - F_{ST})/F_{ST}$ as described by Zhou et al. (2015). Other genetic statistics for measuring genetic variation within sub-population/cluster such as the number of private/specific alleles, percentage of polymorphic loci, Nei's diversity (*h*), and Shannon's index diversity (*I*) were employed in GenAlEx v6.3 (Peakall and Smouse 2006).

The non-parametric analysis of molecular variance (AMOVA) was performed using GenALex v6.3 to detect genetic variance within and among populations based on STRUCTURE analysis (K = 5). The genetic differentiation between populations was determined using PHI_{PT} (Φ_{PT}) value (an analogue of Fst) that allows the within-population variance to be suppressed and calculate population differentiation using binary data. To further analyze the genetic relationship of the accessions and compare clustering patterns obtained in UPGMA cluster analysis and population structure, principal coordinate analysis (PCoA) of the binary genetic distance matrix used for AMOVA was carried out using the covariance distance matrix method in GenALEx v6.3. The first three PCoA values were used to create a three-dimensional scatter plot using PAleontological STatistics (PAST v4.02) (Hammer et al. 2001) to further validate the sub-population grouping.

RESULTS AND DISCUSSION

Morphological Characterization

Morphological Diversity in Qualitative and Quantitative Characters

Phenotyping is an important work in doing plant genetic conservation to maximize the utilization of the germplasm collection, especially in breeding programs. Seventy-six accessions presently conserved at IPB were assessed using 46 qualitative and 11 quantitative morpho-agronomic characters following the Philsurin - IPB descriptor. The Shannon-Weaver diversity index is a statistical parameter that was used to determine species diversity in this population with an overall mean qualitative diversity index of 0.72 (Table 2a). Out of the 46 qualitative characters, 32 or 70% were highly diverse and most of this characters were stool and leaf blade-related. Most of the genotypes observed in the population have erect to intermediate plant habit, more compact with 5 - 8 number of tillers, less waxy stalks, and trichomes, and with intermediate trashiness. Moderately diverse traits were observed for 8 (17%) descriptors with indices ranging between 0.43 – 0.66. Two of these traits, split/ growth cracks (H' = 0.63) and corky cracks (H' = 0.60) are important characters to consider during selection since presence of these characters allows infection and infestation of diseases and pests (i.e. spittlebug) that leads to low level of sucrose content and percent soluble solids (Brix) (Garcia et al. 2010). This is the reason why farmers and sugarcane breeders avoid selecting genotypes with these traits. Moreover, six of the 46 characters had a low diversity with an average of 0.15 that includes stool stripes on the cane, leaf scar, dewlap (waxiness and hair),

Table 2a. Shannon-Weaver Diversity Index (H') on 46 qualitative characters.

TRAIT		H'	TRAIT		H'
Stool	Plant Habit	0.88	Leaf-blade	Leaf Margin Pubescence	0.69
	Tillering Habit	0.97		Presence of Genetic Freckles	0.66
	Tillering Density	0.79		Waxiness	0.72
	Tops	0.99		Primary Color	0.93
	Leaf Carriage	0.98		Secondary Color	0.96
	Color	0.81		Presence of trichomes	0.81
	Texture	0.35		Trichome Quality	0.88
	Erectness	0.85	Auricle	Outer Shape	0.87
	Trashiness	0.95		Inner Shape	0.84
	Waxiness	0.79	Dewlap	Waxiness	0.31
	Predominant Internode Color	0.92		Primary Color	0.83
	Secondary Internode Color	0.94		Secondary Color	0.92
	Stripes on Cane	0.18		Shape	0.80
	Splits/growth cracks	0.63		Margin Undulation	0.88
	Corky Cracks	0.60		Hair	0.18
	Corky Patch	0.97	Ligule	Shape	0.43
	Internode Shape	0.65		Hairness	0.95
	Alignment	0.60	Bud	Shape	0.93
	Node Swelling	0.47		Prominence	0.69
	Growth Ring	0.81		Bud Germ Pore Posi-tion	0.13
	No. of Rows of Root Primordia	0.77		Bud Hair	0
	Leaf Scar	0.10		Bud Tip Position	0.99
	Root Band	0.91		Bud Base Position	1.00
				Mean	0.72

Diversity level adapted from Altoveros et al. (2006) : Maximum (H '= 1.00), high (H' = 0.67 - 0.99), moderate (H' = 0.34 - 0.66), low (H '= 0.00 - 0.33).

and bud (germ pore position and hair). Two qualitative traits were observed to be of extreme indices- 1.0 and 0.0. The bud base position having a diversity index of 1.0 has an equal frequency of genotypes having bud base position (a) at leaf scar and position (b) above a leaf scar. On the contrary, all genotypes exhibited no prominent bud hair (H' = 0). Like most of any crops, presence of hair or trichomes is one of the resistance mechanisms of sugarcane against pathogens such *U. scitaminea* that causes smut in sugarcane as studied and discussed by Da Gloria et al. (1995). Hair serves as barrier in the basal region of the bud thus hindering immediate penetration of spores and other pathogens to the host plant. This suggests that the introduction of clones with wide phenotypic distribution on those poorly diverse characters is something to look at to improve the average genetic diversity of the collection. However, increasing its variability by introducing clones to your breeding population does not guarantee a genetic gain in the breeding program unless it is positively and directly correlated to yield and sugar-related components.

Sugarcane accessions showed diverse quantitative phenotypes in terms of cane yield, percent Brix, millable canes, plant height, stalk length, internode (length and diameter), leaf blade and bud (length and width) having an average diversity index of 0.75 (Table 2b). The majority of these characters have a moderate to high diversity index. Stool-related characters, the number of millable canes, and stalk length all had a moderate diversity index of H' = 0.64 and H' = 0.65, respectively. Cane yield/plot, leaf blade (length and width), and bud length showed the highest diversity index which ranged from 0.77 to 0.91. The Analysis of Variance (ANOVA) of the mean square of selected quantitative characters resulted in highly significant variations ($P < 0.01$) but was found non-significant for stalk length where P = 0.079 (Table 3). The non-significant variation observed in stalk length is consistent with the result from Muhammad et al. (2014) and Mohammed et al. (2019), who assessed sixteen sugarcane varieties in Pakistan and Nigeria. At the IPB, UP Los Baños condition, the majority of the quantitative characters also performed differently. Both the number of millable cane and stalk length were classified as less variable in the population. Future collection travel and selection of these characters must be considered, as they are important parameters for improving crop biomass.

Overall, it was observed that most of the characters assessed across all accessions were moderate to highly variable based on Canepoint classification. This result

Table 2b. Shannon-Weaver Diversity Index (H') on 11 quantitative characters.

TRAIT		H'
Stool	Cane weight/ plot (kg)	0.77
	% Brix (12 mon)	0.74
	No. of millable Cane	0.64
	Plant Height (m)	0.72
	Stalk length (cm)	0.65
	Internode Diameter (cm)	0.73
	Internode Length (cm)	0.70
Leaf Blade	Length (cm)	0.81
	Width (cm)	0.86
Bud	Length (cm)	0.91
	Width (cm)	0.70
	Mean	0.75

Diversity level adapted from Altoveros et al. (2006) : maximum (H' = 1.00), high (H' = 0.67 - 0.99), moderate (H' = 0.34 - 0.66), low (H' = 0.00 - 0.33).

Table 3. Analysis of variance of different agronomic components in 76 sugarcane (*Saccharum officinarum* L.) accessions.

TRAITS	SOURCE OF VARIATION					
	BLOCK	GENO-TYPES	(EE)[1]	(SE)[2]	CV [3](%)	MEAN
Plant Height (m)	3.81**	1.27**	0.41	0.11	7.89	4.02
No. of Millable Cane	16.92ns	12.20**	6.73	4.73	12.94	6.30
Stalk Length (cm)	55832.50**	6652.53ns	5045.98	1995.77	21.63	147.42
Internode Length (cm)	144.19**	28.54**	12.05	4.66	14.33	10.96
Internode Diameter (cm)	17.00**	6.43**	3.28	1.03	9.48	9.14
% Brix (TOP)	288.16**	45.86**	7.64	3.48	5.15	22.87
% Brix (MIDDLE)	188.30**	40.79**	9.60	4.06	5.91	22.99
% Brix (BOTTOM)	117.19**	35.80**	8.38	3.85	5.23	23.48
Total % Brix	189.73**	38.25**	6.44	1.39	5.60	23.15
Cane Yield per plot (kg)	10151.07*	6107.46**	2918.04	1189.58	25.25	95.05

*, **significant at 0.05 and 0.01 probability levels, respectively.

ns no significance, [1]Experimental Error, [2]Sampling Error, [3]Coeifficent Variation.

would provide future strategic research plan to sugarcane breeders emphasizing on conservation and exploitation of existing genetic variability of the collection. Also, it provides insights on the geographic distribution of the morphological characteristics of the sugarcane accessions throughout the regions of origin of collection.

Cluster Analysis

The association among the 76 sugarcane accessions as revealed in the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) cluster analysis is shown in Figure 3. The analysis divided the accessions into 31 clusters using the 57 morphological markers at 0.8 coefficient of similarity. It was observed that nine clusters had three to eight accessions and more than half (69%) of the clusters were composed of one to two accessions. This observation already reveals the genetic diversity of the collected accessions. Among the single-accession cluster, most of those accessions were collected from regions I, III, IVA, VI, and VIII particularly the provinces of Leyte, Negros Occidental, Aurora, Batangas, Samar, Romblon, and Quezon. The biggest cluster – cluster I was composed of eight accessions coming from five geographically distant regions. The same clustering pattern was observed in clusters II, III, XI, XIV, XVII, and XXVIII having three accessions collected from all sampling regions. Two next large clusters were groups of Luzon regions, cluster V and XX with seven and five accessions, respectively. This showed that there was no distinctive clustering pattern according to their place of collection. Some accessions collected from different regions tend to group together and likewise, accessions from Luzon were sometimes clustered to accessions from the Visayas. This indicates the likelihood of having a shared pool of conserved alleles among collected accessions and also disclosing the possibility of the close similarity between their evolutionary courses. Similar findings were observed from the work of Ftwi et al. (2016) and Gashaw et al. (2016) from their cluster analysis using the phenotypic data set. Out of the 13 clusters composed of paired-accessions, four clusters were collected in the same provinces or regions. FIC-34 and FIC-51 of Cluster IV, for example, were both collected at Brgy. Putlan, Carrangalan, Nueva Ecija. Similarly with FIC-41 and FIC-76 (cluster XVIII) which were both commercial checks developed by Sugar Regulatory Administration having a trait to survive under drought stress condition. Cluster VII (FIC-39 and FIC-49) and cluster XXIX (FIC-14 and FIC-69) also contained paired-accession from region VIII and IVA, respectively. Having samples collected from the same province or regions, the presence of duplicates was not observed thereby explaining that the collected accessions from the same geographic regions or province had different genetic backgrounds. Thus, genetic variation between local sugarcane clones within the region suggests some degree of genetic diversity due to variability in farmer's field condition and the possibility of human-mediated gene flow apart from artificial selection.

Genetic Diversity

Descriptive Statistics of 50 Simple Sequence Repeats (SSR) Markers

Microsatellite markers have been successfully used in assessing the genetic relationships of individuals within a population. One consideration in selecting a genetic marker is its polymorphic information content (PIC) which is a measure of marker's informativeness. A total of 259 alleles were amplified using the 50 SSR markers across accessions with a mean of 5.40 alleles per genotype (Table 1). The number of alleles amplified per locus by one primer pair varied from one (Sach28 and Sach33) to ten (Sach1, Sach7, and Sach16) with fragment size ranging from 65 – 610 bp, indicating a high level of heterozygosity and polyploidy nature of sugarcane that led to amplification of more number alleles per locus. Among the 50 SSR markers, only 41 (82%) were highly informative with a mean PIC value of 0.69. The highest

Table 1. Microsatellite primers used in diversity analysis of sugarcane (*Saccharum officinarum* L.) accessions.

	Code	SSR Motif	Orientation	Primers (5' to 3')	AT[2]	$PR_{(bp)}$[3]	NB[4]	PIC[5]
Genomic SSRs	Sach 1	$(TCGA)_3$	Forward	ATCCAAATATCCAATCAATCC	54.6	130-300	10(9)	0.91
			Reverse	TCTCGTGTCATGTGTACTTGT				
	Sach 2	$(CT)_{19}$	Forward	GTACAACTTTTGCTTCCTGTG	55.3	105-215	5(5)	0.91
			Reverse	GGTACTCAGACCCTTGAATTT				
	Sach 3	$(CAAA)_3$	Forward	TCGTAATCTCGGTCTACTTTTT	54.7	135-235	3(2)	0.37
			Reverse	AAATTTTGGGGTGATCTAAAC				
	Sach 4	$(TTC)_4$	Forward	ATGATTGATGAGACGAGATGA	55.3	155-195	4(4)	0.76
			Reverse	AAGCACATAAACTGAACCGTA				
	Sach 6	$(AATA)_4$	Forward	GTAAAATGTTTTTCCCCCTTA	55.5	105-220	4(3)	0.81
			Reverse	TGTGTCAAAAGATTTGATGTG				
	Sach 7	$(AGAAA)_3$	Forward	CTTTCCCTGCAAGATTTTC	55.5	190-340	10(10)	0.94
			Reverse	ATGATCACAGTCTGAGAAAGG				
	Sach 8	$(CAGG)_3$	Forward	TCGTAGACTGCGTACAACCA	54.7	145-195	4(4)	0.79
			Reverse	CGCGTAGACATAGCAACTAAC				
	Sach 9	$(AAG)_4$	Forward	GAATTTCTACTCCACCCATCT	54.0	145-160	3(3)	0.72
			Reverse	TTAGCTTTCTTTTGTCCCTTT				
	Sach 12	$(CCAG)_3$	Forward	ATATAAAAACCACTCCCGAAC	57.0	160-195	5(5)	0.72
			Reverse	GTTACCGGTGGTGATGAG				
	Sach 14	$(ACTG)_3$	Forward	GGTGCTGATTTGTTATGAGAG	52.1	125-250	9(9)	0.94
			Reverse	GAACATACTCCAGCGTTCAT				
	Sach 15	$(CTC)_4$	Forward	ATATGTCAACCGAGTGTTTTG	55.0	125-240	7(7)	0.89
			Reverse	CTGAGTTGGAGAAGAAAGAGG				
	Sach 16	$(AGGC)_3$	Forward	CCTCTTTCTTCTCCAACTCAG	50.8	210-495	10(10)	0.97
			Reverse	GCTGCCATCTTCTTTAACTC				
	Sach 18	$(CACG)_4$	Forward	TCGTTAGAACTTGCTTTTTGT	55.0	110-360	9(9)	0.86
			Reverse	GAGATAGGCTCAATCTTGAAA				
	Sach 19	$(CATG)_3$	Forward	AGCCTATCTCCTATGGCTATG	54.7	65-90	3(3)	0.59
			Reverse	TCGTAGACTGCGTACATGAG				
	Sach 20	$(AATA)_3$	Forward	TAAAATGTTGTTCCAACCATC	52.1	110-250	6(4)	0.81
			Reverse	TCTGCTCATTTTGTTTGTTCT				
	Sach 21	$(GAGGAC)_3$	Forward	GGAAGTAGAGGTGGTTCTTGT	55.0	80-240	9(9)	0.91
			Reverse	CTGTTGTCGCTATCGTAATG				
	Sach 22	$(GCA)_5$	Forward	GGCGACTAACTCTATCAACAA	56.4	100-280	6(6)	0.86
			Reverse	TAGCAGCAAAGAGATGAATGT				
	Sach 23	$(TATT)_3$	Forward	ATTTCACTGAACACGTATGCT	55.4	140-360	6(5)	0.68
			Reverse	TTGTTTAGCTTCAAAATGGAC				
	Sach25	$(ATT)_5$	Forward	TATGAAAAGGCAGGCATACTA	55.3	160-320	7(7)	0.80
			Reverse	GTCAACTTGTCAGAACCATGT				
	Sach 26	$(CATG)_3$	Forward	GCACGTAAATCAAGTGAATAAA	56.0	180-360	4(3)	0.97
			Reverse	CCACACAACTGATGAAGAGAT				
	Sach 27	$(TTCTTT)_3$	Forward	TAGAAAGACACTTGGAGATGC	54.8	160-230	5(3)	0.70
			Reverse	GGTAACGGGTTTGGAATATAA				
	Sach 28	$(TTC)_4$	Forward	TTGGCAGACACCTTCTTG	55.2	120	1(0)	0
			Reverse	GGTCAGAGCAATTTTCAACC				
	Sach 29	$(TGCA)_3$	Forward	CACATGAGTTCATCCTTGAAT	55.3	140-380	7(7)	0.86
			Reverse	ATGCACCGTACCACACATA				
	Sach 32	$(AT)_6$	Forward	AGACTGCGTACAACGTCTATC	54.6	140-380	8(8)	0.92
			Reverse	GGTCAGATTTAGGGTCAAAAA				
	Sach 33	$(CTG)_4$	Forward	ATGTATGTCATTGTCGTGGTT	56.5	140	1(0)	0
			Reverse	TGCTCGTAGACTGCGTACAA				
	Sach 36	$(TTGT)_3$	Forward	CCTTACTTCATGCTTCTTTGA	55.1	200-430	6(6)	0.67
			Reverse	TACATGCAGAGAGGAAGAGAC				
	Sach 38	$(AAG)_4$	Forward	GAATTTCTACTCCACCCATCT	55.0	80-180	6(6)	0.55
			Reverse	TTAGCTTTCTTTTGTCCCTTT				
	Sach 39	$(TGT)_4$	Forward	TTGGGTCTGCATGCTACT	55.5	80-200	4(4)	0.70
			Reverse	ATGCGTAGACCTAGCAACAG				
	Sach 40	$(TTTC)_4$	Forward	CAAGCTCTTCCTCTCTGTTTT	55.5	150-260	5(5)	0.83
			Reverse	CATGCAGACCCAACAGTAG				
	Sach 41	$(GCCG)_3$	Forward	ATAATTAGCGGAGCACACTTT	56.0	140-180	4(3)	0.62
			Reverse	GTGGTGAGCCTCCTCTTC				
	Sach 42	$(TGGA)_3$	Forward	GAGGTTTGACTCATGGATACC	55.5	110-150	2(1)	0.29
			Reverse	GAAATTTTTGCTCGTTTCC				
	Sach 43	$(AACA)_3$	Forward	GCTTTGAATACCAGCACATAG	55.0	90-240	6(6)	0.87
			Reverse	ATGTACGTTGTGCTCTTCCTA				

Table 1. Continuation...

	Code	SSR Motif	Orientation	Primers (5' to 3')	AT[2]	PR$_{(bp)}$[3]	NB[4]	PIC[5]
	Sach 44	$(CCA)_5$	Forward	ACATTTTTCAGCTTTGTTCAC	55.5	160-220	3(2)	0.20
			Reverse	ACCTTTCAGACCATCTGTTTC				
	Sach 45	$(TGGA)_3$	Forward	CACGCACATCATGTTCATAC	55.5	90-170	5(4)	0.76
			Reverse	GTACCTGGCCATCTCCTAC				
	Sach 46	$(TTTC)_3$	Forward	GTCTCAATGCTCTGCTCTG	54.5	160-200	2(2)	0.49
			Reverse	TCCCATCGTTGTACAGATTAC				
	Sach 47	$(TTCT)_3$	Forward	CTGCATTTCTTTCTTTCTTTG	55.0	160-170	3(2)	0.41
			Reverse	TCCCATCGTTGTACAGATTAC				
	Sach 48	$(CTGCTC)_3$	Forward	GCTCGTAGACTGCGTACAA	55.0	130-210	6(6)	0.82
			Reverse	CATCTGCAGAGGATCTCG				
	Sach 49	$(ATA)_4$	Forward	CCCCTAAAAACTTAGGTCTGA	55.0	150-190	4(4)	0.71
			Reverse	CGTTAATTTGTTAGTGCCTGT				
	Sach 50	$(ATGA)_3$	Forward	TTTCTCGATGAAGGAATATGA	55.0	240-540	7(7)	0.90
			Reverse	GCCTCTTTTCCTTGTTTATGT				
	Sach 52	$(TTGC)_3$	Forward	GTCACCTTTCTGACTTCACTG	54.5	150-210	4(4)	0.70
			Reverse	TTCTAGCAGAATCCTCAAAAG				
	Sach 53	$(CGAA)_3$	Forward	TACCAAATACAAACGAAATGC	55.0	130-170	4(3)	0.71
			Reverse	TAAGCTGAAACAAGCTGAAAC				
	Sach 54	$(GTGCGC)_4$	Forward	CAAGTACAGCGTCGTCAGT	55.0	140-170	3(2)	0.46
			Reverse	GGACAAACAAACAGAGAACAG				
	Sach 55	$(AAAAAG)_3$	Forward	CAAGCTATGGTCTTTCTTTGA	55.5	120-200	8(8)	0.92
			Reverse	CTAGAACTTGATGCCAAGTCA				
	Sach 56	$(AAAAAG)_3$	Forward	AGGTAACGGGTTTGGAATA	54.5	340-400	2(2)	0.48
			Reverse	AGACACTTGGAGATGCTCTTA				
	Sach 57	$(GAAG)_3$	Forward	TGAAAATGGGAAAACAGACTA	55.0	140-175	5(5)	0.79
			Reverse	AGTAGTGCTGGTGAACAGGTA				
EST SSRS	SOD [6]		Forward	GCTGTTGGGGATTGATGTCT	60.0	100-200	4(3)	0.62
			Reverse	ATGGTCAAACAGCCGAAAAC				
	I3GPS[7]		Forward	GGATCGCGTCCTCAAGATAG	60.5	430-610	5(5)	0.82
			Reverse	TCGCTCACAACCTGTATTCG				
	LEAP [8]		Forward	CCGTCTCGTGAACCATCC	59.5	410-560	6(6)	0.56
			Reverse	GCAACAACACTGCCACAAAC				
	WRKY-TF[9]		Forward	GCACCACGAGAAGAAGAAGG	60.0	180-300	4(1)	0.74
			Reverse	GCCCACCTCTCAGTCTCTCA				
	Aqua[10]		Forward	GCATGATCTTCGTCCTCGTC	61.0	140-340	7(7)	0.87
			Reverse	CGCCGTACCTGTCGAAGTAG				

[1] repeat type , [2] annealing temperature, [3] product range, [4] no. of bands (No. of polymorphic band), [5] polymorphic information content, [6]super oxide dismutase, [7]indole-3-glycerol phosphate synthase, [8]late embryogenesis abundant protein, [9]WRKY transcription factor, [10]aquaporin.

PIC value (0.97) was observed on SSR markers Sach15 and Sach25, whereas the lowest PIC value (0.20) was observed on primer set Sach55. A PIC value of 0.7 and 92% polymorphism was observed between genomic SSRs while PIC value and percent polymorphism for EST-SSRs observed were 0.72 and 82%, respectively. It is expected that genomic SSRs are more polymorphic compared to EST-SSRs since the latter amplified more in regions of the sugarcane genome that are conserved (Parthiban et al. 2018).

Genetic Relatedness Among Accessions and Phylogenetic Analysis

Based on the estimated genetic dissimilarity (GD) (Jaccard's coefficient index) values as shown in the pie chart, a total of 2850 pairwise combinations were created from 76 accessions (Figure 2). Analysis of SSR data showed that the pairwise genetic dissimilarity values ranged from 0.301 (between FIC-14 and FIC-15) to 0.688 (between FIC-05 and FIC-27) with a mean of 0.564. Five hundred forty four combinations were highly diverse with a GD index between 0.61 – 0.69, for example, FIC-05 (Cluster II) and FIC-27 (Cluster VI), were both genetically diverse and geographically apart from each other. High level of dissimilarity was also recorded between FIC-10 and FIC-63 (0.685), FIC-10 and FIC-65 (0.684), and FIC-05 and FIC-57 (0.684) (data not shown). GD values ranging from 4.7 to 53.8% were reported by Zeni Neto et al. (2019) using 36 SSR markers in 82 sugarcane accessions that are parental materials of the Sugarcane Breeding Program of

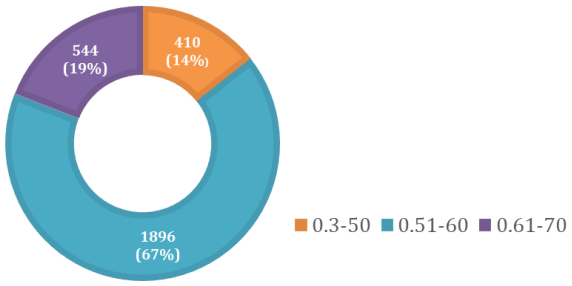

Fig. 2. A pie chart depicting the number of cross-combination across each set of genetic distance values.

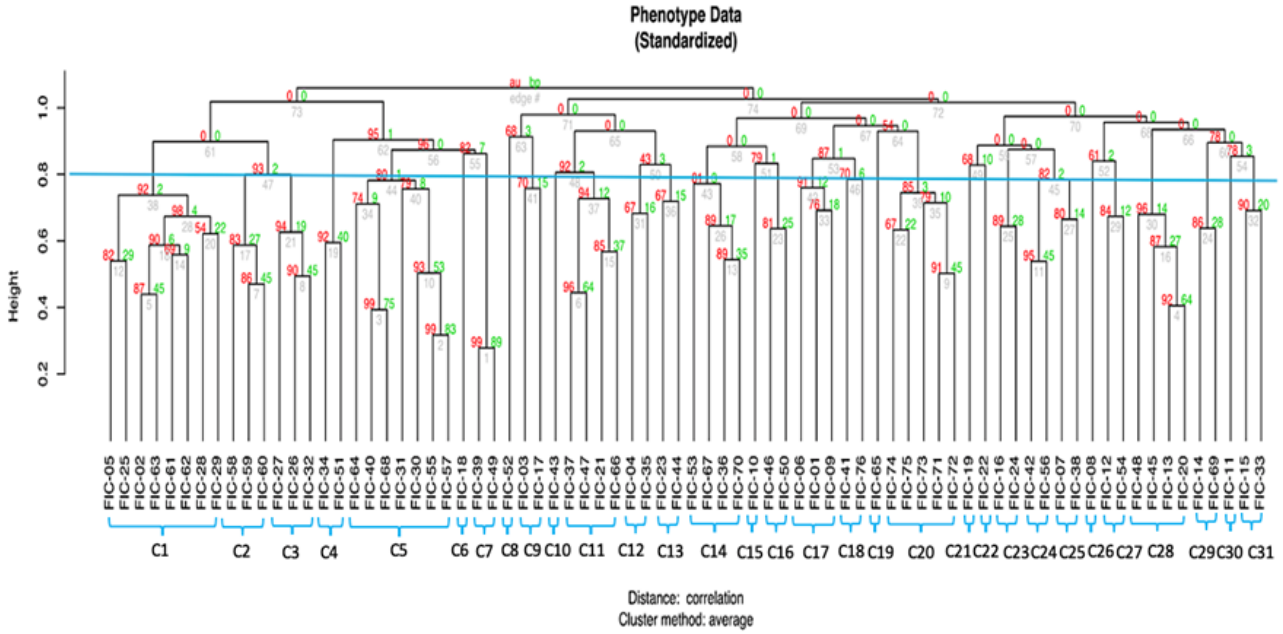

Fig. 3. Phylogenetic tree of 76 sugarcane (*Saccharum officinarum* L.) accessions and varieties using 57 morphological markers.

Brazil. A large chunk of combinations (1896) exhibited moderate diversity was in the range of GD of 0.51 – 0.60, which reflects a substantial level of diversity among experimental material studied. On the contrary, 14% (410) of the total combinations were with GD between 0.30 – 5, and these groups of highly similar combinations of accessions need to be avoided during artificial hybridization. According to previous research, using these highly similar combinations for breeding resulted in less genetic gain (Hemaprabha et al. 2013; Padmanabhan and Hemaprabha 2018). The result of the study opens up a selection of highly diverse parental materials from the present-day breeding pool for guided hybridization. This will serve as an advantageous step to the current breeding program of IPB-UPLB for varietal development of the crop. Strategic utilization of these large diverse combinations in the breeding program will establish a good transgressive population with a wide array of good agronomic characters such as biomass and sugar-related traits and expedite the breeding process in planning crosses for hybridization.

At 0.65 coefficient of dissimilarity, a total of six clusters were generated based on Jaccard's coefficient and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) as shown in Figure 4. Out of which, two singletons, FIC-04 and FIC-39 were observed as distinct from each other and the rest of the clusters based on the molecular marker data. The FIC-04 of cluster I was considered distinct based on its phenotype due to its good tillering density (> 8) with compact tillering habit while FIC-39 of cluster IV had a characteristic of recline growing habit. Cluster III formed the largest cluster consisting of 33 locally collected accession from 13 provinces and five SRA check varieties. Cluster VI included 17 accessions where the majority were collected from regions III and VIII particularly in the provinces of Nueva Ecija and Leyte. The third-largest group cluster V contained 12 accessions followed by cluster II with 8 accessions coming from six regions of Luzon and Visayas. It was also observed that the commercial check varieties were dispersed from the four main clusters excluding the single-accession clusters. There was no distinctive trend of clustering genotypes according to their place of collection, similarly from the result of cluster analysis using morphological data where most of the samples from different regions tend to cluster frequently than samples coming from the same region. These findings are molecularly supported with the observation in sugarcane accession where materials from the same country were often clustered in different groups and likewise, materials from different countries were often in the same cluster, largely because the genetic exchange of gene pool between countries is usually observed (Tena et al. 2014; Singh et al. 2020).

Genetic Structure and Population Differentiation

The structure harvester computed the best value of K at 5 (ΔK = 41.068), suggesting that the entire accessions set can be divided into five sub-populations viz. SP1, SP2, SP3, SP4, and SP5 (Figure 5). Based on the membership proportion for each accession with probability ≥ 50% were placed into the corresponding population and

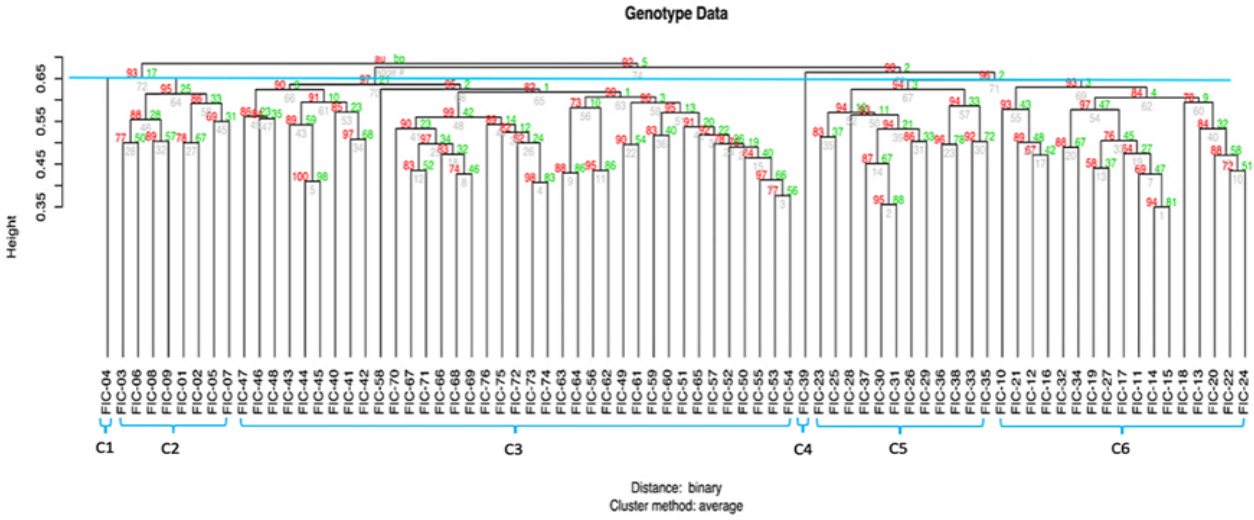

Fig. 4. Dendrogram (UPGMA) of 76 sugarcane (*Saccharum officinarum* L.) accessions using Jaccard's coefficient dissimilarity of 0.65.

others classified as admixture indicated by bar colors (Figure 6). Both SP1 and SP5 were the biggest population comprised of 17 accessions and followed by SP2, SP4, and SP2 with 16, 14, and 12 accessions respectively. The 13 accessions across five sub-populations except SP3 were considered as an admixture with a membership proportion of less than 50% including three SRA commercial checks (FIC-22, FIC-41, and FIC-45). An admixture of sugarcane samples between two-sub-populations was normally observed aside from the fact that different sugarcane breeders have a certain specific trait of interest that contributes to the structure of the breeding population (Luo et al. 2019; Wu et al. 2019). It was found out that there was high genetic variability among sub-population as revealed by the percentage of polymorphic bands (60%), where the highest mean percentage of polymorphic bands was recorded in SP2 (66.93%), and SP4 (58.76%) as the least. The number of private bands among sub-populations ranged from 7 (SP4) to 19 (SP2). The mean genetic diversity among sub-population was estimated using Shannon's information index (*I*) and Nei's gene diversity (*h*) was recorded as 0.282 and 0.183, respectively. Among the sub-population, the highest degree of variability was exhibited by sub-population, SP2 (*h* = 0.197 and *I* = 0.306), and the lowest was recorded in a sub-population, S4 (*h* = 0.182 and *I* = 0.280) (Table 6). The studied population was genetically less diverse compared with the study on exotic sugarcane cultivars (*h* = 0.214 and *I* = 0.321) in Ethiopia by Tena et al. (2014), mainly because of the difference in the number of accessions evaluated and most of their accessions were coming from 15 different sources in 14 countries.

Analysis of molecular variance (AMOVA) analysis was also used to assess the level of genetic variation among five sub-populations and within them based on defined grouping from structure analysis (Table 4). AMOVA with Φ_{PT} calculation showed genetic variance among and within sub-population was statistically significant with a value of Φ_{PT} = 0.152 (*P* value < 0.001) and the variation within a population (85%) was

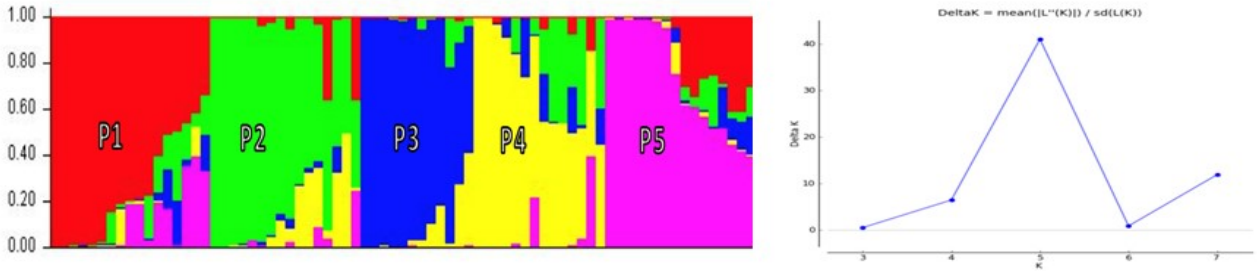

Fig. 5. Population structure of 76 collected sugarcane (*Saccharum officinarum* L.) accessions with a value of K at 5 based on 50 SSR markers (left) and the maximum adhoc measure ΔK determined by STRUCTURE HARVESTER with a sharp peak at K = 3.

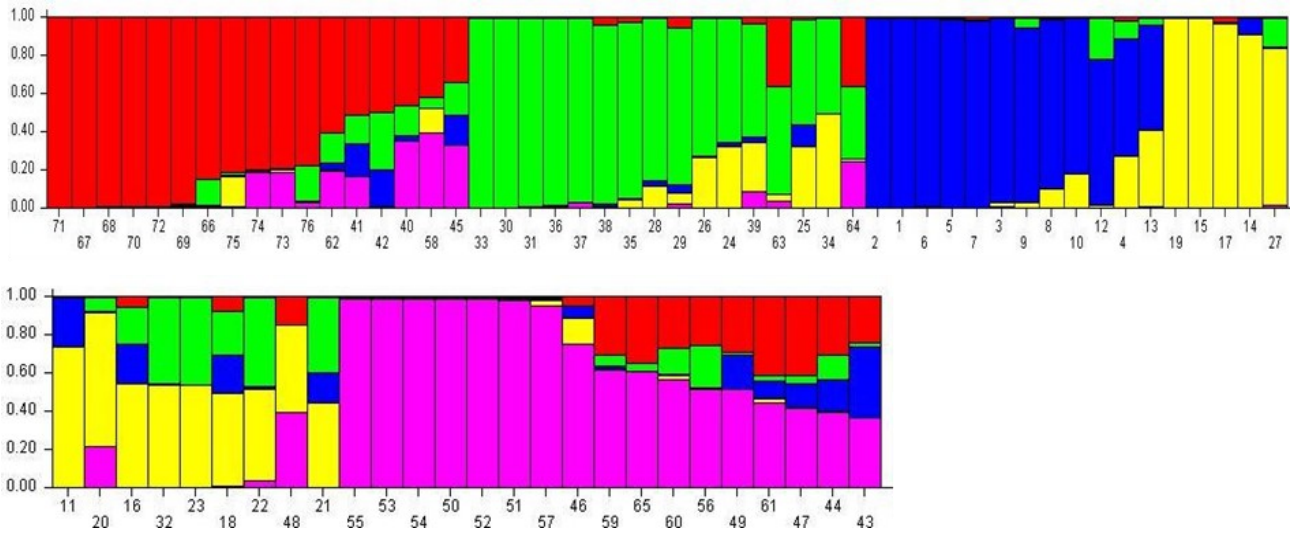

Fig. 6. Membership proportion of each sugarcane (*Saccharum officinarum* L.) accession based on STRUCTURE analysis at K value = 5.

significantly higher than that among sub-population (15%). Φ^{PT} (analogous to fixation index, F_{ST}) measures the genetic variation within the population showed SP4 (Φ_{PT} = 0.134) to be more variable, while among the population, SP1 and SP4 (Φ_{PT} = 0.216) exhibited more variability (Table 5). According to Peakall and Smouse (2006), the F_{ST} value of zero indicates no differentiation between the sub-populations, otherwise when the value is one it indicates complete differentiation. Also, the F_{ST} value greater than (> 0.15) is considered significant in differentiating populations (Wright 1965), where the present study exhibits a considerable amount of genetic differentiation among studied populations. Gene flow (N_m) is the migration of genetic variation from one population to another. The rate of gene flow among the population ranged from 0.24 (SP4) – 0.57 (SP2) with an average of 0.39 (Table 6). It was reported that gene flow value less than one suggest limited genetic exchange between sub-populations (Deng and Li 2007). The average N_m value from the five sub-populations supports the previous result from the AMOVA where there is high genetic differentiation within sub-population (85%) as compared within sub-population (15%). These results corroborate from the previous studies that reported genetic differentiation in the sugarcane population (Wu et al. 2019; Zeni Neto et al. 2020; Singh et al. 2020). Despite that sugarcane is clonally propagated there is still substantial genetic differentiation within sub-population among studied sugarcane collection. According to Zheng et al. (2017), perennial crop species that are vegetatively propagated and outcrossing are generally highly heterozygous and maintain high levels of genetic variation within populations. To estimate the genetic relationship among sub-population principal coordinates analysis (PCoA) was carried out, which was based on the covariance matrix with data standardization using GenALEx software (Figure 7). Two and three-dimensional scatter plots revealed that the first three principal coordinates accounted for total variations of 17.98%. According to PC-1 explaining 8.36% of the variability expressed among accessions, samples were somehow grouped into clusters isolated from the others except for P1 and P5. Looking at PC-2 accounting for 4.97% variation, similar observation were revealed showing accessions from P3, P4, and P2 can be distinctively separated as a unique group while P1 and P5 were both concentrated in the middle between 3rd and 4th quadrant. To further characterize the relationship of the populations, three-dimensional PCoA was employed. The PC-3 accounted for 4.64% of the total variation, allows the accessions from the later population to be discretely separated as shown in Figure 7. PCoA results showed that the genetic distance between samples does not correlate with the geographical location where the accessions being collected. FIC-14 and FIC-15 for example were collected from geographically distant provinces of Luzon. Present findings were analogous with the recent studies on sugarcane diversity (Singh et al. 2020). Also, the analysis was comparable in terms of their genetic relationship and structure with the UPGMA clustering using molecular marker data set, AMOVA, and STRUCTURE analysis.

CONCLUSION

Sugarcane despite being asexually propagated exhibited moderate to highly variable morphological characteristics. The UPGMA cluster analysis, in which sugarcane accessions were morphologically subdivided into 31 clusters, backed up this variability. When compared to molecular data, clustering based on

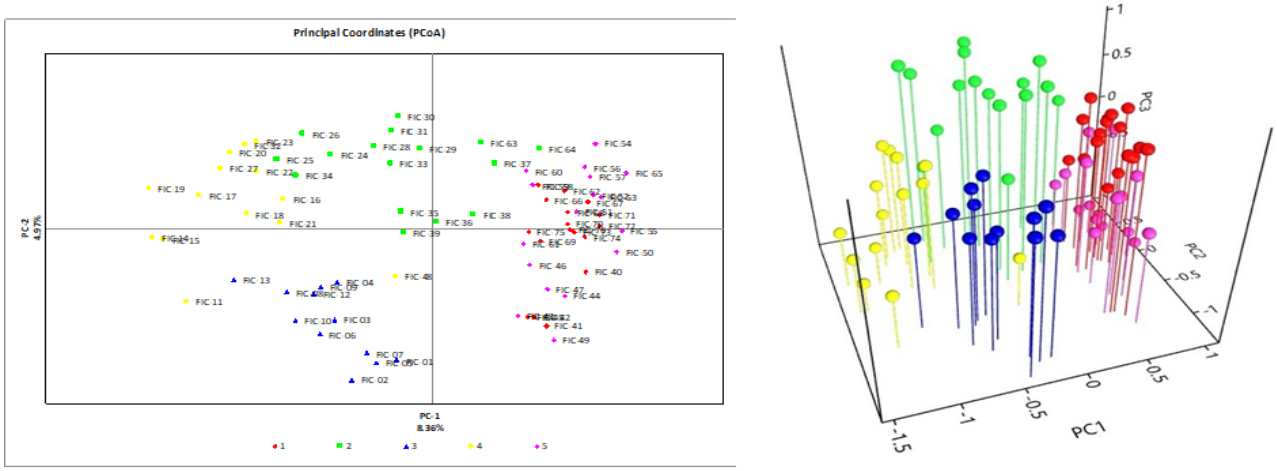

Fig. 7. Two (left) and three (right) - dimensional principal coordinates analysis (PCoA) scatter plots of 76 sugarcane (*Saccharum officinarum* L.) accessions based on 50 SSR markers. Red, green, blue, yellow, and pink colors represent the population P1, P2, P3, P4, and P5, respectively.

morphological parameters gives an indication of how phenotypically diverse the experimental samples analyzed are, but it is ineffective in grouping the samples based on their origin. Fingerprints of the 76 accessions using 45 genomic-SSR and 5 EST-SSR primer pairs were subjected to different bioinformatics tools to identify the genetic relationship, diversity, phylogeny, population structure, and principal coordinate analysis. The results showed a large portion of moderate to highly diverse (GD = 0.51 - .70) pairwise combination based on the Jaccard coefficient index which can be successfully exploited through molecular selection- and hybridization -guided approach. The clustering based on UPGMA, structure, and PCoA was predominantly consistent with minor differences due to observed admixtures. However, no association was observed between geographical and genetic distance. Clustering revealed that accessions from the different regions or provinces were generally, but not totally, clustered into the same cluster and more so with accessions from the same region or province normally falls into different clusters. Current results showed that accessions were clustered into five sub-populations and genetic differentiation within sub-population was high than that among sub-population. Among the sub-populations, SP2 (comprising of FIC-24, FIC-25, FIC-26, FIC-28, FIC-29, FIC-30, FIC-31, FIC-33, FIC-34, FIC-35, FIC-36, FIC-37, FIC-38, FIC-39, FIC-63, and FIC-64) would be the best group for selecting candidate parents for hybridization having moderately high genetic variation (h = 1.82, I = 28, F_{ST} = 0.51) with minimal gene flow (N_m = 0.23) and also intercrossing these with the accessions from SP1 and SP5 will also broaden the genetic pool of IPB-UPLB sugarcane germplasm. In summary, the results of the study serve as a baseline for future breeding efforts and expedite the breeding processes in selecting parents for crosses while maintaining genetic diversity in the breeding population.

Table 4. Analysis of molecular variance (AMOVA) between five sub-populations of sugarcane (*Saccharum officinarum* L.) accessions based on defined grouping from STRUCTURE analysis.

Source	df	SS	MS	Est. Var.	%	*P*-Value
Among Pops	4	1096.167	274.042	13.220	15%	<0.001
Within Pops	71	5248.518	73.923	73.923	85%	<0.001
Total	75	6344.684		87.142	100%	

Table 5. Pairwise (PHI_{PT}) genetic distance between sub-population of sugarcane (*Saccharum officinarum* L.) accessions based on defined grouping from STRUCTURE analysis.

Sub-Pop	SP1	SP2	SP3	SP4	SP5
SP1	0				
SP2	0.122	0			
SP3	0.183	0.150	0		
SP4	0.216	0.116	0.138	0	
SP5	0.093	0.129	0.191	0.201	0
Mean	0.123	0.103	0.132	0.134	0.123

Table 6. Summary of diversity estimates for each of the five sub-populations of sugarcane (*Saccharum officinarum* L.) accessions based on defined grouping from STRUCTURE analysis.

Sub-Pop	NGS	PPB (%)	NPB	Ave. *h*	Ave. *I*	Ave. F_{st}	Ave. N_m
SP1	17	58.96	14	0.181	0.278	0.4099	0.3599
SP2	16	66.93	19	0.197	0.306	0.3058	0.5675
SP3	12	50.60	17	0.168	0.254	0.3844	0.4004
SP4	14	58.76	7	0.182	0.280	0.5104	0.2398
SP5	17	64.74	19	0.186	0.291	0.3992	0.3763

[1]number of genotypes, [2]percentage of polymorphic bands, [3]number of private bands, [4]mean of Nei's diversity (*h*), [5]mean of Shannon's information index (*I*), [6]mean of F_{st} value, [7]mean of gene flow (N_m).

REFERENCES CITED

AITKEN KS, JACKSON PA, MCINTYRE CL. 2006. Quantitative trait loci identified for sugar-related traits in a sugarcane (*Saccharum* spp.) cultivar× *Saccharum officinarum* population. Theoretical and Applied Genetics 112 (7): 306-1317.

ALTOVEROIS NC, HEULGAS VC, GUECO LS, SINOHIN VGO, ASORA FG, QUILLOY RB, MUYCO RR, DE OCAMPO M HAUTEA DM. 2002. Canepoint: A catalogue of Philippine Sugarcane Genetic Resources. Netherlands: common fund for commodities.

ALWALA S, KIMBENG CA, VEREMIS JC, GRAVOIS KA. 2009. Identification of molecular markers associated with sugar-related traits in a Saccharum interspecific cross. Euphytica 167 (1): 127.

ALWALA S, SUMAN A, ARRO JA, VEREMIS JC, KIMBENG CA. 2006. Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. Crop Science 46 (1): 448-455.

CLIFFORD HT, STEPHENSON W. 1975. An introduction to numerical classification. London, UK: Academic Press.

D'HONT A, GRIVET L, FELDMANN P, GLASZMANN JC, RAO S, BERDING N. 1996. Characterization of the double genome structure of modern sugarcane cultivars (*Saccharum* spp.) by molecular cytogenetics. Molecular and General Genetics MGG 250 (4): 405-413.

DA GLORIA B, ALBERNAS M, AMORIM L. 1995. Structural characteristics of buds of sugarcane cultivars with different levels for resistance to smut/ Strukturelle Charakteristika der Knospen von Zuckerrohrsorten mit unterschiedlichen Resistenzgraden gegen Zuckerrohrbrand. Zeitschrift Für Pflanzenkrankheiten Und Pflanzenschutz/ Journal of Plant Diseases and Protection, 102(5): 502-508. Retrieved July 23, 2020, from www.jstor.org/ stable/43386903.

DENG HH, LI QW. 2007. Utilization of CP72-1210 in sugarcane breeding program in mainland China. Guangdong. Agric. Sci. 11:18–21.

DOS SANTOS JM, DUARTE FILHO LSC, SORIANO ML, DA SILVA PP, NASCIMENTO VX, DE SOUZA BARBOSA GV, TODARO AR, NETO CER, ALMEIDA C. 2012. Genetic diversity of the main progenitors of sugarcane from the RIDESA germplasm bank using SSR markers. Industrial Crops and Products. Industrial Crops and Products 40: 145-150.

DOYLE J, DOYLE JL 1987. Genomic plant DNA preparation from fresh tissue- CTAB method. Phytochem Bull 19 (11): 11-15.

EARL DA. 2012. Structure Harvester: a website and program for visualizing structure output and implementing the evanno method. Conserv. Genet. Resour. 4: 359–361.

FTWI M, MEKBIB F, ABRAHA E. 2016. Multivariate analysis of sugar yield contributing traits in Sugarcane (*Saccharum officinarum* L), in Ethiopia. African Journal of Plant Science 10 (8): 145-156.

GARCIA DB, RAVANELI GC, MADALENO LL, MUTTON MA, MUTTON MJR. 2010. Damages of spittlebug on sugarcane quality and fermentation process. Scientia Agricola, 67 (5): 555-561.

GASHAW ET, MEKBIB F, AYANA A. 2016. Genetic diversity among sugarcane genotypes based on qualitative traits. Advances in Agriculture, 2016.

HAMMER Ø, HARPER DAT, RYAN PD. 2001. Past: paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4 (1): 9.

HAPSORO D, WARGANEGARA HA, UTOMO SD, SRIYANI N, YUSNITA. 2015. Genetic diversity among sugarcane (*Saccharum officinarum* L.) Genotypes as shown by randomly amplified polymorphic DNA (RAPD). Agrivita 37 (3): 247.

HEMAPRABHA, G, PRIJI PJ, PADMANABHAN TS. 2013. Molecular fingerprinting of recently notified sugarcane (*Saccharum* L.) varieties using STMS markers. Journal of Sugarcane Research, 3 (2): 107-117.

PHILIPPINE SUGAR RESEARCH INSTITUTE – PHILSURIN AND INSTITUTE OF PLANT BREEDING. 2002. Standardization of descriptors for morphological, cytological, biochemical, and molecular characterization of sugarcane germplasm. Unpublished workshop output, National Plant Genetic Resources Laboratory, Institute of Plant breeding, University of the Philippines, Los Baños.

JACCARD P. 1901. Tude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin de la Société Vaudoise des Sciences Naturelles, 37 (142): 547-579.

JAIN SK, QUALSET CO, BHATT GM, WU KK. 1975. Geographical pattern of phenotypic diversity in a world collection of durum wheats. Crop Science 15: 700-704.

KALIA RK, RAI MK, KALIA S, SINGH R, DHAWAN AK. 2011. Microsatellite markers: An overview of the recent progress in plants. Euphytica 177: 309-334.

KAWAR PG, DEVARUMATH RM, NERKAR Y. 2009. Use of RAPD markers for assessment of genetic diversity in sugarcane cultivars. Indian Journal of Biotechnology 8: 67-71.

LIMA MLA, GARCIA AAF, OLIVEIRA KM, MATSUOKA S, ARIZONO H, DE SOUZA JR. CL, DE SOUZA AP. 2002. Analysis of genetic similarity detected by AFLP and coefficient of parentage among genotypes of sugarcane (*Saccharum* spp.). Theoretical and Applied Genetics 104 (1): 30-38.

LUO Z, BROCK J, DYER JM, KUTCHAN TM, AUGUSTIN M, SCHACHTMAN DP, ABDEL-HALEEM H. 2019. Genetic diversity and population structure of a Camelina sativa spring panel. Frontiers in plant science, 10: 184.

MAECHLER M, ROUSSEEUW P, STRUYF A, HUBERT M, HORNIK K. 2012. Cluster: cluster analysis basics and extensions. R package version, 1(2): 56.

MALIK SR, BAKHSH A, ASIF MA, IQBAL UMER, IQBAL SM. 2010. Assessment of genetic variability and interrelationship among some agronomic traits in chickpea. International Journal of Agriculture and Biology 12 (1): 81-85.

MELCHINGER AE, KUNTZE L, GUMBER RK, LÜBBERSTEDT T, FUCHS E. 1998. Genetic basis of resistance to sugarcane mosaic virus in European maize germplasm. Theoretical and Applied Genetics 96 (8): 1151-1161.

MENDOZA T. 2020. Multiple threats to *PH* sugar industry. Inquirer. net. https://opinion.inquirer.net/119717/multiple-threats-to-ph-sugar-industry.

MEYER J, REIN P, TURNER P, MATHIAS K, MC GREGOR C. 2011. Good management practices manual for the cane sugar industry (Final). PGBI Sugar and Bio Energy, Johannesburg, South Africa.

MOHAMMED AK, ISHAQ MN, GANA A, AGBOIRE S. 2019. Evaluation of sugarcane hybrid clones for cane and sugar yield in Nigeria. African Journal of Agricultural Research, 14 (1): 34-39.

MOORE PH. 1987. Anatomy and morphology. Chapter 3. In: DJ Heinz, ed. Sugarcane improvement through breeding. Elsevier Amsterdam. pp. 85-142.

MUHAMMAD K, HIDAYAT R, RABBANI MA, FARHA T, AMANULLAH K. 2014. Qualitative and quantitative assessment of newly selected sugarcane varieties. Sarhad Journal of Agriculture 30(2): 187-191.

NEI M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences, 70 (12): 3321-3323.

PADMANABHAN TS, HEMAPRABHA G. 2018. Genetic diversity and population structure among 133 elite genotypes of sugarcane (*Saccharum* spp.) for use as parents in sugarcane varietal improvement. 3 Biotech, 8 (8): 339.

PARTHIBAN S, GOVINDARAJ P, SENTHILKUMAR S. 2018. Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. Biotech 8: 144

PEAKALL R, SMOUSE PE. 2006. GenAlEx tutorials: introduction to population genetic analysis. Australian National University, Canberra.

PHILIPPINE STATISTIC AUTHORITY. 2019. Major non-food and industrial crops quarterly bulletin, October-December 2019. Philippine Statistic Authority. https://psa.gov.ph/non-food/sugarcane.

PRITCHARD JK, STEPHENS M, DONNELLY P. 2000. Inference of population structure using multilocus genotype data. Genetics 155 (2): 945–959.

R CORE TEAM. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

RABOIN LM, PAUQUET J, BUTTERFIELD M, D'HONT A, GLASZMANN JC. 2008. Analysis of genome-wide linkage disequilibrium in the highly polyploid sugarcane. Theor. Appl. Genet. 116: 701–714.

SANGUILLOSA JB, RASCO JLS, LAURENA AC, LALUSIN AG. 2019. Enriched mini-genomic library for the development and characterization of simple sequence repeat (SSR) markers in sugarcane (*Saccharum* species hybrids). Philippine Agricultural Scientist 102 (2): 95-106.

SAS INSTITUTE INC. 2002. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

SHARMA MD, DOBHAL U, SINGH P, KUMAR S, GAUR AK, SINGH SP, JEENA AS, KOSHY EP, KUMAR S. 2014. Assessment of genetic diversity among sugarcane cultivars using novel microsatellite markers. African Journal of Biotechnology 13 (12): 1444-1451.

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 RB, JUGRAN AK, SINGH RK, SRIVASTAVA RK. 2020. Assessing genetic diversity and population structure of sugarcane cultivars, progenitor species, and genera using microsatellite (SSR) markers. Gene, 144800.

SKINNER JC. 1972. Description of sugarcane clones. III. Botanical description. Proceedings of the International Society of Sugar Cane Technology 14: 124-127.

TENA E, MEKBIB F, AYANA A. 2014. Analysis of genetic diversity and population structure among exotic sugarcane (*Saccharum* spp.) cultivars in Ethiopia using simple sequence repeats (SSR) molecular markers. African Journal of Biotechnology, 13 (46).

WRIGHT S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution, 395-420.

WU J, WANG Q, XIE J, PAN YB, ZHOU F, GUO Y, QIU Y. 2019. SSR Marker-assisted management of parental germplasm in sugarcane (*Saccharum* spp. hybrids) breeding programs. Agronomy, 9 (8): 449.

ZENI NETO H, BORSUK LGDM, SANTOS LRFD, ANGELI HS, BERTON GS, SOUSA LLD. 2020. Genetic diversity and population structure of sugarcane (*Saccharum* spp.) accessions by means of microsatellites markers. Acta Scientiarum. Agronomy, 42.

ZHENG Y, XU S, LIU J, ZHAO Y, LIU J. 2017. Genetic diversity and population structure of Chinese natural bermudagrass [*Cynodon dactylon* (L.) Pers.] germplasm based on SRAP markers. PLoS One: 12 (5).

ZHOU R, WU Z, JIANG FL, LIANG M. 2015. Comparison of gSSR and EST-SSR markers for analyzing genetic variability among tomato cultivars (*Solanum lycopersicum* L.). Genet Mol Res 14: 13184–13194.