International Journal of Agriculture, Environment and Bioresearch

Vol. 3, No. 04;2018

ISSN: 2456-8643

ASSESSMENT OF GENETIC DIVERSITY OF FIRST PRIORITY PARENTALS OF THE SUGAR REGULATORY ADMINISTRATION

John Moises G. Relles[1], Antonio C. Laurena[2], Ma. Genaleen Q. Diaz[3], Antonio G. Lalusin[2]

[1]Sugar Regulatory Administration – La Granja Agricultural Research and Extension Center La Granja, La Carlota City 6130, Negros Occidental, Philippines

[2]Institute of Plant Breeding, University of the Philippines Los Baños College, Los Baños 4031, Laguna, Philippines

[3]Institute of Biological Science, University of the Philippines Los Baños College, Los Baños 4031, Laguna, Philippines

ABSTRACT

Genomic DNA of 160 first priority parentals from the Sugar Regulatory Administration were extracted and subjected to genetic diversity analysis using 14 polymorphic SSR markers, two of which were developed from a previous study. Statistical indices of the genetic diversity are: H_E = 0.998926 and PIC = 0.84, showing high genetic diversity and efficacy of the markers used. Three different clusters (I, II,III) were generated at dissimilarity coefficient 0.68. Cluster I had 10 varieties, mostly VMC and other non-Phil varieties. Sixteen varieties, mostly Phil 88- to Phil 90- varieties were grouped in Cluster II. Cluster III had 134 varieties, mostly Phil 58- to Phil 2005- series. Principal Component Analysis (PCA) also showed that alleles from the 160 SRA first priority parentals were widely distributed and grouped accordingly based on their initial clustering in the dendrogram.

Keywords: Genetic diversity, Sugar Regulatory Administration

1. INTRODUCTION

The Sugarcane Industry Roadmap 2020 of the Sugar Regulatory Administration (SRA) is formulated to serve as guide in the identification and implementation of appropriate programs and interventions to prepare the industry for year 2015 and beyond when tariff of imported sugar will be reduced to 5%, and the full integration of the ASEAN Economic Community (AEC) takes effect. It proposes a multi-product sugarcane industry with bioethanol and power as major products other than sugar, and bioplastics, biowater, biofertilizer, and special sugars as sources of revenue. The sugar industry contributes about Php 70 billion to the Philippine economy from the production of raw and refined sugar, molasses, and bioethanol.

The SRA's target outputs are: (1) increase in sugarcane area from 422,384 hectares to 465,000 hectares; (2) increase in farm productivity from 57 tons cane per hectare to 75 tons cane per hectare (57 TC ha^{-1} to 75 TC ha^{-1}); and (3) increase in sugar yield from 1.80 bags per ton cane to 2.1 bags per ton cane (1.80 LKg TC^{-1} to 2.1 LKg TC^{-1}).

These can be achieved by improving farm productivity and sugar yield. Around 90% of the total sugarcane farms in the country are small farms, and they only produce an average of 50 TC ha^{-1}.

The agency, through its programs and interventions, has provided two experimental stations that will cater to the research, development and extension needs of the sugarcane industry in coordination with the Philippine Sugar Research Institute Foundation, Inc. (PHILSURIN), which is the private sector research arm of the sugar industry and University of the Philippines Los Baños. SRA-Luzon Agricultural Research and Extension Center (LAREC) in Floridablanca, Pampanga conducts research and development projects on sugarcane nutrition, while SRA-La Granja Agricultural Research and Extension Center (LGAREC) in La Carlota City, Negros Occidental conducts breeding for variety improvement and maintains the germplasm. LGAREC serves as the only center for sugarcane breeding. The station focuses on the development of varieties that will surpass the performance of control varieties.

Efficiency of commercial farms can be improved by evaluating the genetic diversity of the SRA First Priority Parentals. In the Philippines, the current basis for the evaluation of genetic diversity in sugarcane germplasm is only based on pedigree records and phenotypic traits. The screening and evaluation of available first priority parentals from SRA with simple sequence repeat markers (SSRs) could both optimize and facilitate the development of varieties resistant to sugarcane smut.

This study assessed the genetic diversity of the SRA first priority parentals. Specifically, the study determined the variability and utility in genome analysis of the developed SSRs.

2. MATERIALS AND METHODS

The study was conducted from June 2015 to May 2017 at several locations namely: the Sugar Regulatory Administration – La Granja Agricultural Research and Extension Center, La Carlota City, Negros Occidental; the Biochemistry Laboratory and the Molecular Plant Breeding Laboratory, College, Laguna under the Institute of Plant Breeding, College of Agriculture and Food Sciences, University of the Philippines Los Baños.

DNA Source

For the genetic diversity, a total of 160 SRA First Priority Parentals from La Granja Agricultural Research and Extension Center (SRA-LGAREC) were extracted and analyzed. DNA from all varieties was isolated using a protocol described by Hulbert and Bennetzen [1] with slight modifications. Seedlings at the three-leaf stage were used as tissue sources, and 0.5 g of tissue was homogenized using liquid N2. After homogenization, 10 mL of 2X CTAB extraction buffer (1.4 M NaCl, 100 mMTris pH 8.0, 2% CTAB, 20 mM EDTA pH 8.0, 0.5% Na bisulfite, 1% 2-mercaptoethanol) was added, mixed gently, and incubated for 1 h at 65°C. An equal volume of chloroform:isoamyl alcohol (24:1 v/v) was added, mixed, and the supernatant was stored. The chloroform:isoamyl step was repeated, and the DNA was precipitated by adding an equal volume of 100% isopropanol and incubated for 2h at room temperature. The DNA pellet was washed once using 70% v/v ethanol, air dried, and resuspended in 200 µL of 1X TE [2]. RNase (Qiagen, Valencia, CA) was added to the DNA solution producing a final concentration of 25 ng gDNA µl-1 and the mixture was incubated overnight at 4°C. A single extraction using

Vol. 3, No. 04;2018

ISSN: 2456-8643

phenol/chloroform/isoamyl alcohol (25:24:1 v/v/v) was used to clean up the DNA, followed by an ethanol/sodium acetate precipitation [2].

Assessment of Genetic Diversity

DNA profiling was performed using two primers developed from CV-Phil 97-3933 (SGS P20 and SGS P141) and 12 SSRs compiled by UPLB Institute of Plant Breeding (IPB) specific for sugarcane (Table 1). The selection of these SSR markers by UPLB IPB is due to their robustness in fingerprinting sugarcane cultivars. Amplification by Polymerase Chain Reaction (PCR) was carried out in a total volume of 15 µL containing 50 ng genomic DNA, 2 mM $MgCl_2$, 1X Vibuffer A, 0.2 mM dNTP, 0.5 unit of Taq DNA polymerase and 0.4 µM each of forward and reverse primers. Reactions were performed in a SpeedCycler^2 PCR (Fast Protocol) with the following cycling pattern: 95°C for 3 min followed by 35 cycles of 95°C for 10 sec, T^a for 10 sec, and 72°C for 20 sec, then finished it at 72°C for 5 min and was put on hold at 12°C.

Table 1. Microsatellite Primers compiled by UPLB IPB specific for sugarcane, utilized for the assessment of genetic diversity of the Sugar Regulatory Administration's 160 first priority parentals.

SSR Code		Primers (5' → 3')	T^a	Size (bp)
ESTA15	Forward Reverse	TGA AGC AGC TAG CGG TCC AC GTT CTC GCG GTT GAT GTC CA	57 °C	136
ESTA16	Forward Reverse	AGT CCC TCT GGC CCT CAC AC GAG GCT CTG TGA TGG GTT CG	52 °C	165
ESTA58	Forward Reverse	GTC GTG CCC AAC ATC AAT GTG GGT CAA CTC CTC TTA CAT	54 °C	255
ESTB94	Forward Reverse	GAG GCA GCC AGG CAG GTC AC GGT GGC AGT GTT CAG GCA GAT G	61 °C	236
ESTB118	Forward Reverse	CTT GGC TAG GGT TTC TTG AGT CGT CAT GGC TTT TGG CTT GCT TCT	56 °C	110
ESTB130	Forward Reverse	GCC CAG GTA ATT ATC CAG ACT C GCT GTT GCT CAC TGG TTC C	53 °C	124
ESTC33	Forward Reverse	GCT CTC GCG CAT CCA TCT GAA ATC AAA CCC GCG GCC CTG ACG AC	66 °C	129
ESTC45	Forward Reverse	GCC GCC GTC GCT GGA TTG GAT GGA TCC CCG CCT ACC CTA CAC	61 °C	139
ESTC66	Forward Reverse	AGT ACA GGC TGC TCT CAA TCA A TCT GTC ATC TGT GTT CGT TCT G	55 °C	152
ESTC119	Forward Reverse	GGA ATT AAG CTT TGC CGA CAC CAC GGC AGC ACC TCC CCT TCA CC	64 °C	159
SOMS118	Forward Reverse	GAG GAA GCC AAG AAG GTG TAG AGC GAG GAG CGA AGG	55 °C	82-1018
SOMS120	Forward Reverse	GCA TCT ATC GGT CTT CTG G ATC CAA TCC TTC ATC TTC TTC	55 °C	84-1155
SGS P20	Forward Reverse	AAA ATT CAT GAG AGC ACG TC TTT GAT CAA TAG TTC CCC TCT	46 °C	199
SGS P141	Forward Reverse	CTG GCA GGA TAT GAA ATA TGA CCT TTT ACA GGC GAA ATT TAT	48 °C	149

Vol. 3, No. 04;2018

ISSN: 2456-8643

Molecular Marker Scoring

Polymerase chain reaction (PCR) was done for each marker and each polyacrylamide gel electrophoresis run were subjected to molecular marker scoring. Only clear and unambiguous bands were scored. Bands were scored one (1) for the presence and zero (0) for the absence of a DNA fragment. The presence or absence of the amplified band in all genotypes indicates similarity, whereas presence in one and absence on another indicates dissimilarity. The polymorphism and robustness of the SSR markers were analyzed using polymorphism information content (PIC) and Nei's [3] genetic diversity (H_E). The genetic diversity of the 160 SRA first priority materials was evaluated by construction of a dendrogram and a principal component analysis (PCA) biplot**.**

Polymorphism Information Content (PIC)

Polymorphism information content (PIC) is a measure of the allelic diversity at a locus. It can be determined using the formula:

$$PIC = 1 - \sum_{i=1}^{l} p_i^2 - \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} 2p_i^2 p_j^2$$

wherein P_i and P_j represents the population frequency of the *i*th and *j*th allele. PIC is the probability that the marker genotype of a given offspring will allow deduction, in the absence of crossing over, of which the two marker alleles of the affected parents it received [4]. PIC is also defined as the modification of the heterozygosity measure that subtracts and additional probability from the H value that an individual linkage analysis does not contribute information to the study [5].

Genetic Diversity Analysis

The genetic diversity (HE) was computed using the unbiased genetic diversity index equation by Nei [3]:

$$H_E = 1 - \sum_{i=1}^{n} p_i^2$$

where p_i is the frequency of the *i*th allele and n is the number of alleles at a single SSR locus.

Principal Component Analysis (PCA) was also computed using R statistical program. This provides an alternative view of the genetic distances among accessions compared to the dendrogram. Statistical analyses were accomplished with the aid of the R Statistical Language and Environment [6].

Vol. 3, No. 04;2018

ISSN: 2456-8643

3.RESULTS AND DISCUSSION

Fingerprinting of SRA First Priority Parentals

A total of 160 varieties were analyzed. The fourteen SSR primer pairs screened showed polymorphism. A band is considered polymorphic if it is differentially present across the cultivars tested. Among the 14 polymorphic loci, a total of 147 alleles were detected with mean value of 10.5 alleles per locus. The number of amplified DNA per primer ranged from 5 to 22 alleles. The largest allele size (1000 bp) was observed from primer ESTA 16, whereas the smallest allele (100 bp) was detected in primer ESTB 130. Primer ESTA 16 also had the most number of fragments at 22 alleles, while primer SOMS 118 had the least number of fragments at 5 alleles.

Table 2. Polymorphism Information Content (PIC) of the SSR markers and the genetic diversity (HE) of the 160 Sugar Regulatory Administration's first priority parental for each marker.

SSR	PIC	H_E
ESTA 15	0.87	0.99942
ESTA 16	0.90	0.99973
ESTA 58	0.88	0.99629
ESTB 94	0.88	0.99956
ESTB 118	0.76	0.99935
ESTB 130	0.85	0.99968
ESTC 33	0.81	0.99937
ESTC 45	0.86	0.99999
ESTC 66	0.79	0.99998
ESTC 119	0.86	0.99998
SOMS 118	0.78	0.99994
SOMS 120	0.88	0.99338
SGS P20	0.91	0.99929
SGS P141	0.79	0.99942
mean	0.84	0.998926

Polymorphism information content (PIC) refers to the value of markers in detecting polymorphism within given genotype. PIC value of the 14 SSR markers used in the study ranged from 0.76 to 0.91 (Table 2). All markers were highly informative, including the two developed markers. These are deemed useful for the study of polymorphism and genetic diversity.

Genetic Diversity and Cluster Analysis

Statistical indices are: H_E = 0.998926 and PIC = 0.84. The calculated H_E implies that there is 99.89% probability that two randomly sampled individuals possess different alleles in at least one locus. The computed PIC value indicates that there is 84% probability that a randomly sampled individual possesses different alleles in a locus. The overall level of polymorphism can

be highly affectedwith the nature of germplasm used, nature and type of SSR loci repeats, and protocol used

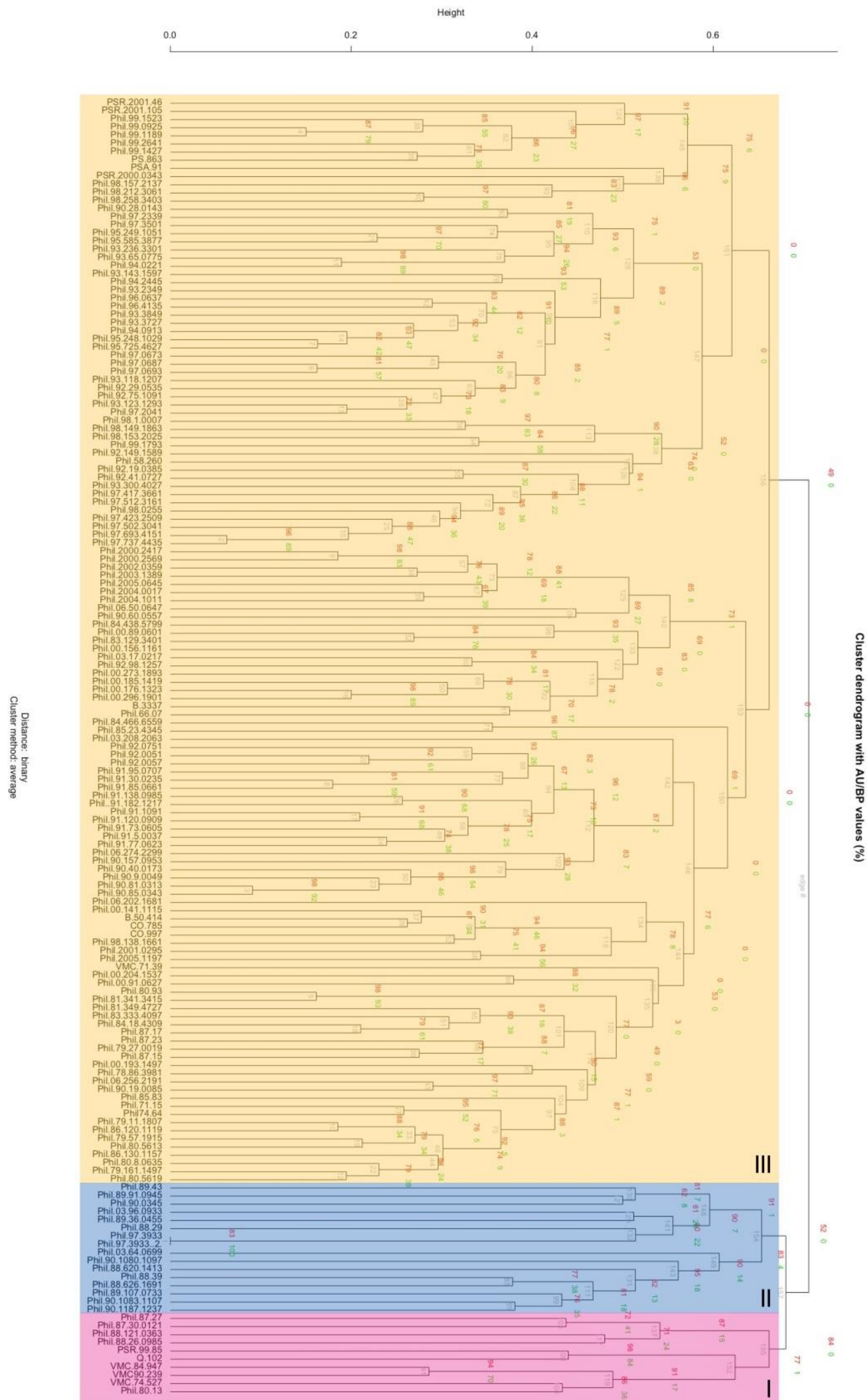

Figure 1. Dendrogram of the Sugar Regulatory Administration's first priority sugarcane parentals showing Clusters I (10 varieties), II (16 varieties) & III (134 varieties).

for allele detection. High genetic diversity can be attributed to sugarcane being allopolyploid, and to the high heterogeneity of the first priority parentals.

All the 14 SSR primers utilized were able to distinguish most of the varieties. Using R, a language and environment for statistical computing and graphics, the dendrogram reveals three clusters (I, II, III) at dissimilarity coefficient of 0.68 (Fig. 4). Three different clusters (I, II, III) were generated at coefficient 0.68 (bootstrap=1000). Cluster III had the most number of varieties (134), mostly Phil-series varieties, ranging from Phil 58- to Phil 2005-series. Non-Phil-series varieties other than VMC were also found at clusterIII. Cluster I with 10 varieties, had the VMC varieties, non-Phil varieties PSR and Q, and few Phil 87- and Phil 88-varieties. Cluster II, with 16 varieties, mostly had the Phil 88- to Phil 90-varieties.

In Cluster III, it can be observed that closely related Phil-series cluster together at coefficient 0.60. The upper part of Cluster III shows a sub cluster of mostly Phil 98- and Phil 99-varieties. A second sub cluster can be observed, comprising mainly of Phil 92- to Phil 97- varieties. The third sub cluster is largely composed of Phil 2000- and later varieties. A fourth sub cluster shows only Phil 85-23-4345 and Phil 84-466-6559 together. The lower part of Cluster III is predominantly composed of varieties preceding Phil 92-series. It is a mixture of Phil-70s, Phil-80s and early Phil-90-varieties.

Also at coefficient 0.60, Cluster II can also be divided into three sub clusters. Two of its sub clusters are predominantly Phil 88- to Phil 90- series, while another sub cluster only has Phil 03-64-0699.

Cluster I can similarly be separated into three sub clusters at coefficient 0.60. One sub cluster is predominantly composed of VMC varieties. Varieties Q 102 and PSR 99-85 constitutes another sub cluster, while another sub cluster is mainly Phil 87- and Phil 88- varieties.

Twenty-six varieties had their parentage available and 10 had previous molecular characterization [7]. These data were analyzed with respect to the clustering of the varieties in the dendrogram. It was observed that varieties having a similar parent are found in the same cluster, like Phil 92-0057 (parentage: Phil 79-001 x CO 467)& Phil 92-0751 (parentage Phil 79-001 x Phil 64-2227), Phil 66-07(parentage: Phil 56-60 x CO 440) & Phil 74-64 (parentage: CO 440 xPhil 54-60), and Phil 93-2349 (parentage: Q102 x Phil 84-77) & Phil 97-3501 (parentage: Q102 x Phil 84-438-5799). A variety, whose parents came from the 160 first priority parentals, was also seen to be in the same cluster as either one of its parents. It was evident in Phil 94-0913 (parent on same cluster:Phil 81-120-1119), Phil 97-3501 (parent on same cluster: Phil 84-438-5799), Phil 98-0255 (parent on same cluster: Phil 92-0751), Phil 99-1793 (parent on same cluster: Phil 93-236-3301), and VMC 71-39 (parent on same cluster: Phil 58-260). Using the previous molecular characterization available from primers mSSCIR74, SMC334BS, SMC119CG, SMC278CS, and SMC336BS, it can be noticed that all varieties screened for these primers (PSA 91, Phil 58-260, Phil 83-129-3401, CO 785, Phil 84-18-4309, Phil 87-15, Phil 80-8-0635, Phil 79-161-1497) were in the same cluster, only with the exception of two VMC varieties, VMC 90-239 and VMC 74-527.

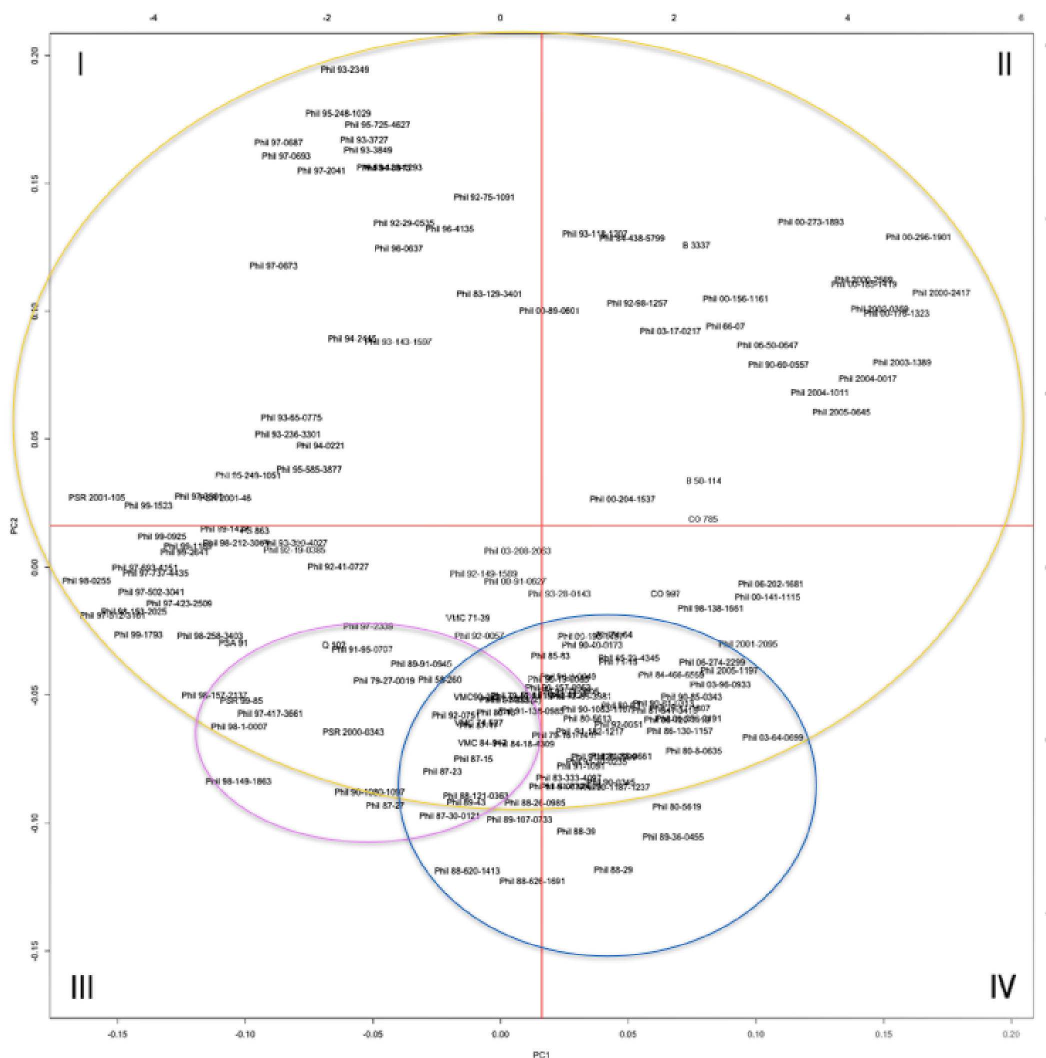

Figure 2. Principal Component Analysis (PCA) showing an alternative view of the genetic distances among varieties. It was observed that alleles from the 160 first priority sugarcane parentals were widely distributed in the four quadrants. Varieties from Cluster III were found all over the four quadrants, while varieties from Clusters II and I, were mostly concentrated at Quadrants IV and III, respectively.

Based on the Principal Component Analysis (PCA), it was observed that alleles from the 160 SRA first priority parentals were widely distributed in the four quadrants (Figure 2). The wide distribution can be attributed to sugarcane being allopolyploid, and to the high heterogeneity of the first priority parentals. Varieties from Cluster III are found all over the four quadrants, while varieties from Cluster I were concentrated on Quadrant III, and varieties from Cluster II were mostly found in Quadrant IV.

4. CONCLUSION

Two microsatellite markers developed from the genomic library (SGS P20 and SGS P141) were utilized for DNA fingerprinting of 160 first priority parentals from Sugar Regulatory Administration, together with the 12 SSR primers from UPLB-IPB. PIC value of the 14 SSR markers used in the study ranged from 0.76 to 0.91. All markers were highly informative, including the two developed markers. These are deemed useful for the study of polymorphism and genetic diversity. PIC values suggest high probability that randomly sampled individual possess different alleles in a locus. All the 14 SSR primers utilized were able to distinguish most of the varieties. Three different clusters (I, II, III) were generated at coefficient 0.68. Cluster III had the most number of varieties (134), followed by Cluster II (16), then Cluster I (10). Principal Component Analysis (PCA) also showed that alleles from the 160 SRA first priority parentals were widely distributed, and grouped accordingly based on their clustering in the dendrogram.

5.RECOMMENDATIONS

The 25 other microsatellite primers developed from Phil 97-3933 should be synthesized and used in DNA fingerprinting and genetic diversity of the SRA first priority parentals. Other SSR primers that the UPLB-IPB has compiled should also be utilized in the genetic diversity analysis.

In addition, traits of interest (e.g. sugarcane smut resistance) of the 160 first priority parentals should be identified so that it can be linked to the microsatellite markers developed, and be used for marker assisted selection.

6. Acknowledgment

This study was done in collaboration with the Sugar Regulatory Administration La Granja Agricultural Research and Extension Center (SRA-LGAREC) and the University of the Philippines Los Baños Institute of Plant Breeding (UPLB-IPB) as part of the Sugarcane Genomics Project of the Philippine Genome Center - Agriculture Program.

REFERENCES

[1] Hulbert, S.H., and J.L. Bennetzen. 1991. Recombination at the Rp1 locus of maize. Mol. Gen. Genet. 226:377–382.

[2] Sambrook, J. 2001. Molecular cloning: A laboratory manual. In J. Sambrook and D.W. Russel (ed.) 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

[3] Nei, M. 1973. Analysis of Gene Diversity in Subdivided Populations. Proceedings of the National Academy of Sciences of the United States of America, 70 (12 Pt 1-2), 3321–3323.

[4] Guo, X., and Elston, R.C. 1999. Linkage informative content of polymorphic genetic markers. Hum Hered 49:112–118.

Vol. 3, No. 04;2018

ISSN: 2456-8643

[5] Speer, M.J. 1999. Genetic linkage: concepts and methods. In: Albers MJ (ed) Genetics of cerebrovascular disease. Blackwell, Oxford, pp 25–26.

[6] 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/.

[7] Altoveros, N.C., et al. 2006. Canepoint: A catalogue of Philippine Sugarcane Genetic Resources. Common Fund for Commodities.