American Journal of Experimental Agriculture
3(1): 102-109, 2013

SCIENCEDOMAIN *international*
www.sciencedomain.org

Genetic Divergence in Sugarcane Genotypes

Mohammad Tahir[1*], Hidayatur Rahman[2], Rahmani Gul[3], Amjad Ali[4] and Muhammad Khalid[4]

[1]Sugar Crops Research Institute, Mardan, KPK, Pakistan.
[2]Department of Plant Breeding and Genetics, KPK A.U. Peshawar, Pakistan.
[3]Agriculture Research Institute (North), Mingora, Swat, KPK, Pakistan.
[4]Research Officers, Sugar Crops Research Institute, Mardan, KPK, Pakistan.

Authors' contributions

This work was carried out in collaboration between all authors. Author MT designed, and laid out the experiment; compiled the study results, followed by statistical analyses; wrote the first draft. Author HR critically reviewed the first draft. Author RG helped in relevant literature search. Authors MK and AA helped a lot during field work and compilation of the data. All authors read and approved the final manuscript.

Research Article

Received 5th October 2012
Accepted 1st December 2012
Published 20th December 2012

ABSTRACT

To assess genetic divergence of sugarcane germplasm, an experiment comprising 25 sugarcane genotypes was conducted at Sugar Crops Research Institute (SCRI), Mardan, Khyber Pakhtunkhwa, Pakistan, in quadruple lattice design during 2008-09. Among the 14 parameters evaluated, majority exhibited significant differences while some showed non-significant mean squares. The initial correlation matrix revealed medium to high correlations. Principal Component Analysis (PCA) showed that there were two principal components accounting for 88% of the total variation in the tested breeding material. The new components were named "Vigor", and "Quality". Principal Component Regression (PCR) indicated that these two accounted for 93.64% and 7.36% of variation in the yield, thus signifying the role of the "Vigor" Component. Cluster analysis using Ward's method on the newly created variables using principal components revealed that there were 3 clusters at a linkage distance of 4.5. Cluster I and III had 11, and cluster II had 3 genotypes. Cluster I showed high mean values for Vigor Component while Cluster II for Quality Component and Cluster III showed genotypes with high mean yield. There was no correspondence of the clustering with the geographic location of the genotypes. It could be concluded from these analyses that there are two main components i.e. vigor, and quality accounting for

**Corresponding author: Email: tasran@gmail.com;*

maximum variation in yield. The genotypes in cluster I and II could be utilized as source for future selection or hybridization program for the improvement of these characters in sugarcane.

Keywords: Sugarcane; genetic divergence; PCA; cluster analysis; Mardan.

1. INTRODUCTION

An essential first step in any varietal development program is to come up with sugarcane germplasm which has sufficient genetic variability. Accurate assessment of genetic diversity is very important in crop breeding as it helps in the selection of desirable genotypes, identifying diverse parental combination for further improvement through selection in the segregating populations, and introgressing desirable genes from diverse germplasm into the available genetic base [1]. Therefore, genetically diverse germplasm is needed in breeding programs to enhance the productivity and diversity of cultivars. Utilization of introduced germplasm and the knowledge of genetic remoteness among them are vital for their manipulation in crop improvement program [2].

Multivariate statistical analysis techniques like Principal Component Analysis (PCA) and Cluster Analysis techniques could be used for evaluating genetic divergence among sugarcane genotypes. These analyses have been used successfully to study genetic diversity in chickpea [2], sugarcane [3,4,5], groundnut [6], and other crops. Ram and Hemaprabha [7] studied 30 hybrid clones involving *Saccharum barberi*, *S. officinarum*, and Co hybrid to evaluate their seven parents to find out the nature and pattern of genetic divergence. The clones were grouped in 15 clusters and grouping of progeny clones was independent of parent cross combination. They concluded that hybridization among clones from diverse clusters may help in isolating progenies with higher sugar yield and its traits. PCA based on 16 morphological traits [8] resulted in 4 principal components which accounted for 76% of the total variation. The first principal component accounted for 29% variation which was mainly attributed to variation in juice quality, yield and stalk diameter traits. Cluster analysis of the 81 cultivars resulted in two major and eight minor clusters. Tai and Miller [9] evaluated sugarcane germplasm from field plots of 4 *saccharum* species and 4 commercial cultivars by means of analysis of sugar composition. Cluster analysis indicated heterogeneity within and among these species. They concluded that information on sugar composition should assist breeders in selecting superior clones for the relevant breeding programs. Ninety-four genotypes of *S. spontaneum* were studied by Zhang et al. [10] for principal component and cluster analysis based on seven quantitative traits of *S. spontaneum*. The three principal components obtained, provided 82.47% cumulative variance. Based on these seven traits, the 94 *spontaneum* genotypes were grouped into 4 clusters.

The present study was conducted to ascertain the role of the most important characters in causing variation in yield and to evaluate the magnitude of genetic diversity in sugarcane genotypes as well for further improvement of the crop.

2. MATERIALS AND METHODS

Plant crop of 25 genotypes was grown (Fig. 2 for the names of the genotypes) at the experimental fields of Sugar Crops Research Institute, Mardan, Khyber Pakhtunkhwa,

Pakistan, using 5 x 5 quadruple lattice design in 4 replications and with one repetition during 2008-09. The genotypes belonged to Canal Point (CP), Houma (HO), Sao Paulo (SP), Natal (N), and others (procured inland), and consisted of 9, 5, 2, 4, and 5 genotypes, respectively. Data were recorded on Germination, Tillering, Plant height (growth), Cane Yield, Millable Canes, POL%, Recovery and Sugar Yield. Crop Growth Rate (CGR) was calculated as subtracting germination, tillering and growth data of 2nd date from the first and then dividing by 30 as the interval in data taking was 30 days.

The data were subjected to statistical analysis using statistical software MSTATC. Principal Component Analysis was carried out using computer software package, SPSS Version 17.0, as it gives a scree plot for numbers of components to retain, and to portray them in a rotated space. Kaiser's MSA (Measure of sampling advocacy) and eigen vectors were worked out in "SAS version 9.1 Portable". Yield character was excluded from PCA as it had to be included in the principal component regression (PCR). SAS is a handy package for finding out MSA values, and eigen vectors using MSA method of "Factor" procedure. The extracted components were rotated using Varimax rotation [2]. A character was said to load well on a principal component if its correlation with the component was greater than 0.40 [11]. The most significant principal components (eigen vectors) were taken and transposed. The raw data were transposed as well. The matrix product of these principal components and the transposed raw data were new variables to be used in subsequent analysis. These new variables were analyzed using SAS Version 9.1 Portable's PROC PLS Procedure for PCR. Cluster analysis was carried out using STATISTICA 7.0 using the newly created variables after standardizing [2, 6]. STATISTICA software was used for cluster analysis for the ease of clustering the genotypes using scree plot and determining the linkage distance at which to make clusters.

3. RESULTS AND DISCUSSION

The initial correlation matrix is given in Table 1 which showed low to high correlation among the characters studied in this investigation. Table 2 shows values of respective Kaiser's Measures of Sampling Advocacy (MSA), for the characters under study. According to Karl [12] MSA value of above 0.5 was sufficient for a variable to be retained for principal component analysis. MSA values for CGR1, CGR2, Growth1 and CGR3 were very low and hence were excluded from further analyses. Growth2 loaded positively on both components 1 and 2, therefore, it was dropped as well. Millable canes were also omitted because of low MSA value. Yield was omitted because it had to go into the Principal Components Regression (PCR) analysis. Overall MSA Value after the two analyses went up to 0.595 which was satisfactory.

Initial eigenvalues presented in Table 3 were above 1 for components 1, and 2. Component 1 accounted for 46.83%, and component 2 for 41.32% of the total variance. The cumulative variance explained by the two components was 88% (above the range of 70 – 80% and hence was satisfactory). Muyco [8] on the contrary found 4 principal components giving rise to 76% variation in the data, with the first component comprising juice quality, yield and stalk diameter traits. Similarly, Deepak et al [13] recorded quality traits (Brix, Pol %) loading on the first two principal components. The rotated component eigenvectors matrix (Table 3) showed that the characters like Tiller 1, Germ2, Germ1, and Tiller2 loaded well on component 1, and Recovery, POL%, and Purity on Component 2. The characters were said to load well on a component if their values were above 0.40 [11]. The components were grouped together in rotated space (Fig. 1). Based on the characters loading on the principal components they could be named as "Vigor", and "Quality", components.

Table 1. Correlation matrix of plant and yield characters of 25 sugarcane genotypes

	Germ1[a]	Germ2[a]	CGR1[b]	Tiller1	Tiller2	CGR2	Growth1	Growth2	CGR3	MCanes[c]	POL%	Purity	Recovery	Yield
Germ1	1.000	.889**	-.152	.764**	.508**	-.507**	.281	.447*	.340	.023*	.085	.025	.068	.165
Germ2	.889	1.000	.317	.819**	.671**	-.296	.219	.343*	.259	-.139	.014	-.056	.006	.123
CGR1	-.152	.317	1.000	.184	.397*	.411	-.111	-.185	-.146	-.350	-.143	-.171	-.126	-.078
Tiller1	.764	.819	.184	1.000	.868**	-.269	.224	.297	.197	-.089	.024	-.054	.004	.245
Tiller2	.508	.671	.397	.868	1.000	.245	.188	.102	-.021	-.146	-.066	-.156	-.083	.135
CGR2	-.507	-.296	.411	-.269	.245	1.000	-.073	-.384*	-.425*	-.109	-.175	-.198	-.168	-.216
Growth1	.281	.219	-.111	.224	.188	-.073	1.000	.605**	-.026	-.008	.061	.142	.084	.167
Growth2	.447	.343	-.185	.297	.102	-.384	.605	1.000	.780**	.181	-.025	.044	.002	.271
CGR3	.340	.259	-.146	.197	-.021	-.425	-.026	.780	1.000	.234	-.080	-.056	-.064	.208
MCanes	.023	-.139	-.350	-.089	-.146	-.109	-.008	.181	.234	1.000	-.140	-.004	-.117	.305
POL %	.085	.014	-.143	.024	-.066	-.175	.061	-.025	-.080	-.140	1.000	.899**	.994**	-.299
Purity	.025	-.056	-.171	-.054	-.156	-.198	.142	.044	-.056	-.004	.899	1.000	.928**	-.271
Recovery	.068	.006	-.126	.004	-.083	-.168	.084	.002	-.064	-.117	.994	.928	1.000	-.297
Yield	.165	.123	-.078	.245	.135	-.216	.167	.271	.208	.305	-.299	-.271	-.297	1.000

[a] *Germination Percentage* [b] *Crop Growth Rate* [c] *Millable Canes * Significant at P= .05 ** Significant at P= .01*

Table 2. Table of MSA's in succeeding analyses for the characters under study

		Germ1[a]	Germ2[a]	CGR1[b]	Tiller1	Tiller2	CGR2	Growth1	Growth2[c]	CGR3	MCanes[c]	POL Percent	Purity	Recovery	Yield	Overall MSA
1	MSA[a]	0.47	0.47	0.22	0.47	0.40	0.30	0.19	0.37	0.28	0.41	0.45	0.72	0.49	0.75[b]	0.409
2	MSA	0.57	0.63	--	0.69	0.61	--	--	0.50	--	0.301	0.50	0.61	0.5	--	0.57
3	MSA	0.55	0.63	--	0.69	0.60	--	--	--	--	0.21	0.53	0.63	0.52	--	0.58
4	MSA	0.58	0.669	--	0.679	0.590	--	--	--	--	--	0.540	0.610	0.520	--	0.595

[a] *Kaiser's Measure of Sampling Advocacy* [b] *Yield was also excluded as it had to go into regression analysis* [c] *Growth2 was also excluded because it loaded both on component1 and 2. While MCanes had a low MSA value.*

Table 3. Principal components (rotated) for the characters of 25 sugarcane genotypes

		PC1	PC2
Eigenvalues		3.278	2.892
% of Variance		46.830	41.316
Cumulative % of Variance		46.830	88.146
	Communalities	**Eigenvectors (Rotated)**	
Tiller1	.778	.955*	-.005
Germ2	.879	.938*	.003
Germ1	.913	.878*	.084
Tiller2	.717	.839*	-.114
Recovery	.969	.008	.993*
POL Percent	.927	.025	.984*
Purity	.987	-.060	.961*

*Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. a. Rotation converged in 3 iterations. * Value >0.40*

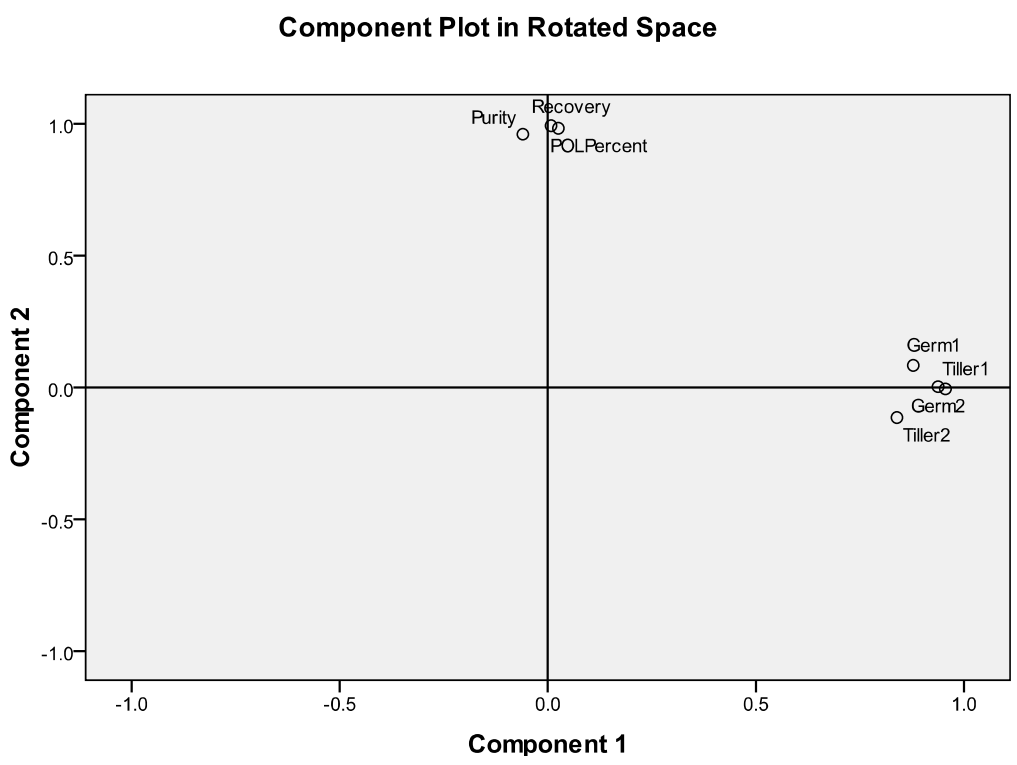

Fig. 1. The components extracted in rotated space

The newly created variables (using eigenvectors of the Principal Components and the original data) were regressed on yield using principal component regression (Table 4). It is obvious from the Table that the Vigor component caused 92.64% variation followed by Quality (7.35). It is inferred that Vigor is an important principal component in causing variation in yield.

Table 4. Principal components regression for the new variables created using principal components (percent variation accounted for by principal components)

Number of extracted factors	Model effects		Dependent variables	
	Current	Total	Current	Total
1	92.6495	92.6495	2.2482	2.2482
2	7.3505	100.0000	4.0639	6.3121

The cluster analysis procedure grouped the genotypes into 3 clusters at a linkage distance of 4.5 with cluster I having 11, cluster II having 3, and cluster III having 11 genotypes, respectively (Fig. 2). Cluster analysis showed that there was no correspondence between the clustering of the genotypes and their geographic origin. The germplasm contained N series, CP series and HO series genotypes and were grouped in separate clusters irrespective of their geographic location. This suggests that the genotypes of different locations have genetic similarity and could have been derived from the same breeding material. Similar results were obtained by Ram and Hemaprabha [7] wherein they found that the progenies of a cross clustered independently of their parents. The mean and standard deviation for the clusters are given in Table 5. Cluster I showed high mean values for Germ1, Germ2, Tiller1 and Tiller2 (i.e., Vigor Component), while Cluster II for POL%, Purity, and Recovery (i.e., Quality Component) and Cluster III showed genotypes with high mean yield.

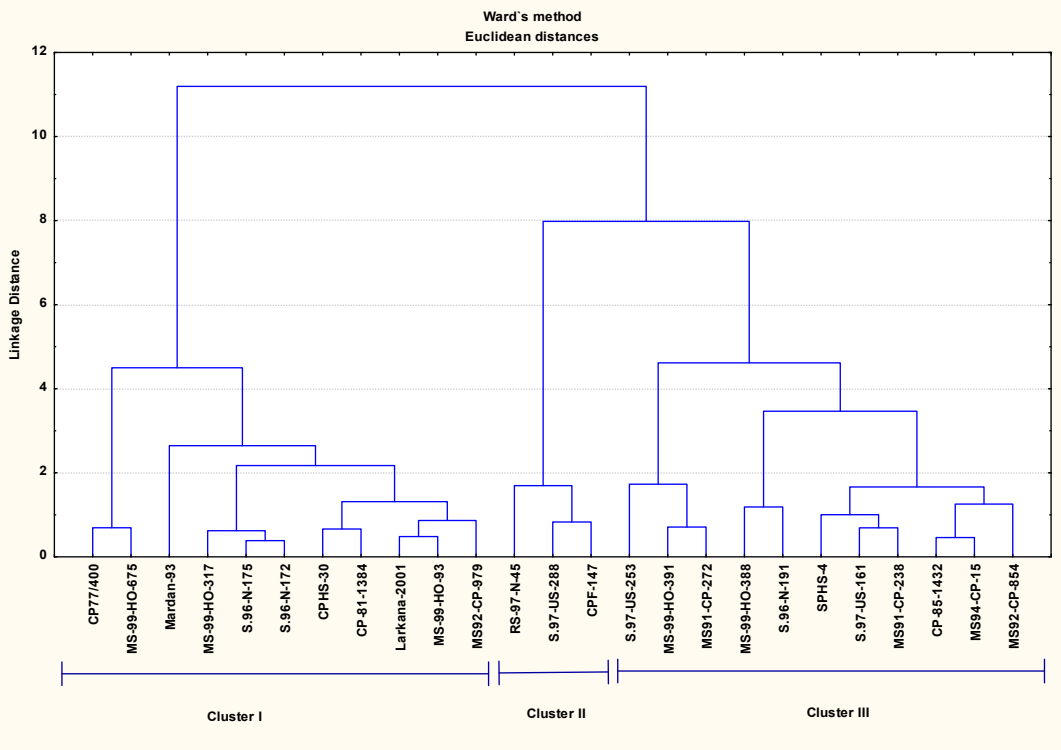

Fig. 2. Dendrogram showing 25 sugarcane genotypes using 3 variables estimated by 3 principal components and original data (the 3 newly created variables were standardized before clustering)

Table 5. Cluster means and standard deviations for the characters of 25 sugarcane genotypes

	Cluster I		Cluster II		Cluster III	
Germ1	44.6	± 3.8	30.5	± 9.0	36.1	± 6.7
Germ2	52.5	± 5.1	36.8	± 10.1	44.1	± 5.9
Tiller1	184.8	± 19.8	111.0	± 18.9	150.2	± 22.6
Tiller2	214.8	± 22.6	152.3	± 1.5	187.6	± 30.6
Pol%	17.9	± 0.5	18.1	± 0.4	17.3	± 0.4
Purity	84.1	± 1.6	85.1	± 1.1	82.7	± 1.2
Recovery	11.3	± 0.4	11.6	± 0.3	10.9	± 0.4
Yield t/ha	60.3	± 6.3	46.6	± 1.1	66.6	± 7.3

4. CONCLUSION

The findings of this study suggest that the important characters responsible for diversity in the sugarcane genotypes could be grouped in two Principal Components namely Vigor and Quality with Vigor traits being comparatively more important than quality. Similarly, the 14 genotypes clustered for high mean values of various traits could be exploited for improvement in Vigor and Quality characteristics either through selection or hybridization. The cluster having high mean values for yield could be selected for yield per se as well.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Mohammadi SA, Prasanna BM. Analysis of genetic diversity in crop plants-salient statistical tools and considerations. Crop Sci. 2003;43(4):1235-1248.
2. Malik SR, Bakhsh A, Asif MA, Iqbal U, Iqbal SM. Assessment of genetic variability and interrelationship among some agronomic traits in chickpea. Int. J. Agric. Biol. 2010;12:81–85.
3. Luo J, Hua Z, Liangnian X. Comparison and Cluster Analysis of Photosynthetic Characters of Different Sugarcane Varieties. Scientia Agricultura Sinica. 2005;38(8):1562–1569.
4. Ilyas MK. Genetic variability and contribution of some morphological traits in cane yield and sucrose recovery in (*Saccharum officinarum*). Accessed 20 August 2011. Available: http://agris.fao.org. AGRIS Record No. PK2004000303.
5. Chen Y, Haihua D, Jinan L. Differences of physiological and biochemical characters among the progenies of Erianthus arundinaceus. Journal of Huazhong Agriculture University. 2006;25(6):598-602.
6. Sudhir IK, Govindraj M, Kumar VK. Estimation of Genetic Diversity of new advanced breeding lines of Groundnut (*Arachis hypogaea L.)*. World J. Agric. Sci. 2010;6(5):547-557.
7. Ram B, and Hemaprabha G. Nature and pattern of genetic divergence of sugar yield and its components in the progenies of *Saccharum barberi*. Indian J. Gen. Plant Breed. 1998;58(2):193-199.

8. Muyco RR. Genetic diversity in sugarcane (*Saccharum* spp. L.) from the active germplasm collection of PHILSURIN [Philippine Sugar Research Inst.,] based on coefficient of parentage, agro-morphological traits and DNA microsatellite markers; 2000. Accessed 20 August 2011. Available: http://agris.fao.org. AGRIS Record No. PH2003001327.
9. Tai PYP, Miller JD. Germplasm diversity among four sugarcane species for sugar composition. Crop Sci. 2002;42:958–964.
10. Zhang G, Yang R, Liu H, Fang W. Principal component analysis for 7 quantitative traits and cluster analysis based on 7 quantitative traits of *Saccharum spontaneum* L. Southwest China J. Agric. Sci;2006. Accessed 20 August 2011. Available: http://en.cnki.com.cn/Article_en/CJFDTOTAL-NYX200606031.htm
11. Stevens J. Applied multivariate statistics for the social sciences. Hillsdale, NJ: Lawrence Erlbaum Associates;1986.
12. Karl LW. Multivariate analysis with SPSS. Department of Psychology, East Carolina University;2011. Accessed 20 August 2011. Available: http://core.ecu.edu/psyc/wuenschk/SPSS/SPSS-MV.htm.
13. Deepak S, Kishore R, Mohammad GHB, Maryke L. From sugar industry to cane industry: investigations on multivariate data analysis techniques in the identification of different high biomass sugarcane varieties. Euphytica. 2012;185:543–558.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=156&id=2&aid=807