academicJournals

Vol. 10(8), pp. 145-156, August 2016
DOI: 10.5897/AJPS2016.1419
Article Number: 66A6F5659493
ISSN 1996-0824

http://www.academicjournals.org/AJPS

African Journal of Plant Science

Full Length Research Paper

Multivariate analysis of sugar yield contributing traits in Sugarcane (*Saccharum officinarum* .L), in Ethiopia

Mebrahtom Ftwi[1]*, Firew Mekbib[2] and Eyasu Abraha[3]

[1]Ethiopian Sugar Corporation, Research and Training, Nazareth, Ethiopia.
[2]Haramaya University, P.O Box 138 Dire Dawa, Ethiopia.
[3]Tigray Agricultural Research Institute, P.O Box -492, Mekelle, Ethiopia.

Received 7 April, 2016; Accepted 17 June, 2016

Knowledge on performance of genotypes and interrelationships among traits is very important for sugarcane breeding programmes. The objectives of this study were to assess the phenotypic relationship among 49 sugarcane genotypes and the inter-relationships among traits considered. The cluster analysis demonstrated that the 49 sugarcane genotypes studied were clustered into nine groups and were highly different for Pol in juice, cane yield (tons $ha^{-1}m^{-1}$), number of tillers (ha^{-1}), purity% and milleable stalk population (ha^{-1}). The relationship among sugarcane genotypes was not dependent on geographic origin, suggesting that a high proportion of total genetic variation was retained within the groups of origin and active genetic ex-change was found between different origins. The principal component analysis indicated that cane yield, milleable stalk height and milleable stalk diameter were highly correlated with sugar yield while the correlation of quality traits with sugar yield was weak. In contrary, path and multiple regression analysis revealed that cane yield, recoverable sucrose percentage (%) and Pol contribute more to the variability of sugar yield; these are very important traits for high sugar yield that should be considered in sugarcane breeding programmes. Moreover, milleable stalk height and milleable stalk population via cane yield and Brix, Pol, purity and number of internodes via recoverable sucrose percentage had high indirect effects on sugar yield suggesting these traits should also be given consideration during selection for high sugar yield. Generally, similar and adequate information was generated following the use of cluster, principal component, linear discriminant, path coefficient and multiple regression analyses indicating the use of multivariate analyses was successful and results of the study were more substantial to give concrete recommendation.

Key words: Clusters, genotypes, multiple regression, path coefficient, phenotypic correlation.

INTRODUCTION

Sugarcane (*Saccharum officinarum* L.) is an important industrial crop and global major source of energy and is a major crop in most parts of tropical and subtropical regions (Khan et al., 2013). The increasing multiple use

*Corresponding author. E-mail: mebreftwi@gmail.com.

of sugarcane necessitates a strong breeding program that generates gene pools that enables identification of sugarcane genotypes with multiple uses. Hence, to have successful genetic improvement of sugarcane genotypes for multiple purposes, efficient and diversified selection procedures have to be followed. Selection will be efficient if the procedures consider many traits simultaneously during evaluation of sugarcane genotypes. As the appropriate methods that provide accurate evaluations and estimation of genetic diversity depend on genetic variation, sampling methods, the magnitude of data sets, and the statistical tools applied in the data analysis (Mohammadi and Prasanna, 2003), multivariate statistical analysis techniques like principal component analysis (PCA) and cluster analysis techniques are very important to study genetic diversity of sugarcane.

However, prior to starting selection, the genetic diversity among the genotypes need to be assessed using morphological and agronomic traits as the genetic diversity assessment is a basic tool to determine whether there could be enough genetic pool that enables generating desirable genes and genotypes. The information about the status of the genetic pool of a crop (introduction, existing commercial varieties and germplasm) using cluster and other analyses, determines the success of selection efficiency of breeding sugarcane (Malik et al., 2010). The quantified gene pool could be used for future breeding purposes such as combination and introgression of genes (Mohammadi and Prasama, 2003).

After the magnitude and pattern of existing genetic base of the crop determined for traits of interest, the most important thing most sugarcane breeding programs deserves to do is to follow efficient selection procedures that utilize both direct and indirect methods to improve quantitative traits. The method of path coefficients was first used for yield component analysis by Dewey and Lu (1959), and subsequently has become a common method to examine breeding strategy in the 'whole variety' context. As yield is affected by numerous components and is a complex resultant character, the internal adjustments between components causes' increment in one component and causes decrement in the other, causing no change in resultant yield (Wen and Zhu, 2005) and causal pathways existed when independent variables are co-related (Kozak et al., 2007).

Ong'ala et al. (2016) recommended PCA and linear discriminant analysis to identify representative traits for phenotypic characterization of sugarcane, and thereby to select superior clones in the breeding process. During phenotypic evaluation of sugarcane clone, many traits are simultaneously evaluated, which are often genetically linked. It is costly to evaluate all the traits which probably may be interrelated and does not ensure optimal selection gains. Path coefficient analysis is one of the most important tool that enable breeders to handle both selection methods simultaneously (Sidwell et al., 1976). Moreover, it enables to have an insight in to the correlation of these effects with the actions of additive and non-additive genes that govern the traits of interest.

High sugar yield are obtained from cane yield and sucrose content (Terzi et al., 2009) and therefore cane yield and sucrose content and their interaction are important parameters for developing superior genotypes (Zhu et al., 2000; Chohan et al., 2007). Several reports clarify about the relationship of cane yield components with cane yield and cane quality traits. For example, Ahmed et al. (2010) reported positive correlation between cane yield and its components (number of milleable stalks/m^2, milleable stalk height internodes/stalk and single weight) but negative association with milleable stalk diameter, Pol in juice and purity. Similarly, Tyagi and Lai (2007) reported that the weight of milleable stalks contributed high direct effect on cane yield followed by milleable stalk, height, number and thickness. Ei-Shafi and Ismail (2006) reported to use multiple regressions model and reported that the main contributors for sugar yield were cane yield, sugar recovery percentage and milleable stalk diameter. Generally, the results of different studies showed discrepancies to the level of sugar yield and cane yield. This phenomenon necessitates successive studies to be conducted to determine the relationship and association of the traits to increase the efficiency of selection.

One of the major sugarcane production constraints in Ethiopia is the lack of high yielding and stable sugarcane varieties across sugar estates. Based on these problems, the Ethiopian Sugarcane Research Sector is introducing, collecting and recycling sugarcane materials to increase the genetic base and efficient use of gene pool of the crop. However, under Ethiopian sugarcane research conditions, the existing diversity among most of the materials is not assessed and selection strategies that help to increase the selection efficiencies of traits have not been well developed. Moreover, multivariate analysis that helps to develop efficient selection strategies have not been efficiently exploited. Therefore, the objectives of this study were to assess the magnitude of genetic divergence among sugarcane genotypes and to study the interrelationships among traits using multivariate techniques.

MATERIALS AND METHODS

Description of experimental materials

Forty-three sugarcane genotypes along with six commercial varieties were grown across Ethiopian Sugar Estates (Wonji, Metahara and Finchaa) and Projects (Tendaho and Belles) over two successive plant cane and first ratoon crops in 2013 to 2016 production years (Table 1). Of which, 21, 3, 5, 7 and 7 genotypes were introduced from France, Philippines, Barbados, USA and Cuba, respectively. The rest 6 varieties were from commercial varieties which had been introduced into Ethiopia from India, South Africa and Barbados before 50 years and were included in this study for comparison purposes. Out of the introduced materials from France, those whose name starts with PG, are clones that

Table 1. Description of 49 sugar cane genotypes analyzed for sugar yield contributing traits.

Code	Genotypes	Origin	Code	Genotypes	Origin
1	PSR97 092	PHILSURIN (Philippines)	26	VMC95 212	USA
2	DB70047	WICSCBS (Barbados)	27	NCO-334	South Africa
3	DB66 113	WICSCBS (Barbados)	28	FG03 418	Cirad (France)
4	FG06 700	Cirad (France)	29	CO449	India
5	FG06 729	Cirad (France)	30	FG03 204	Cirad (France)
6	PSR97 087	Cuba	31	FG02 553	Cirad (France)
7	PSR97 051	PHILSURIN (Philippines)	32	FG03 103	Cirad (France)
8	HO95 988	USDA (Louisiana)	33	FG03 318	Cirad (France)
9	Cp99 1534	USDA (USA)	34	FG04 708	Cirad (France)
10	FG04 829	Cirad (France)	35	FG04 705	Cirad (France)
11	DB71 060	Cirad (France)	36	FG02 551	Cirad (France)
12	TCP93 4245	USDA (Texas/Canal Point)	37	FG03 173	Cirad (France)
13	CP001 252	USDA (USA)	38	FG04 187	Cirad (France)
14	VMC95 173	USA	39	FG03 372	Cirad (France)
15	FG03 447	Cirad (France)	40	FG03 214	Cirad (France)
16	CO 740	India	41	C86-56	Cuba
17	CP99 1894	USDA (USA)	42	SP70-1284	Cuba
18	FG03 425	Cirad (France)	43	C86-165	PHILSURIN
19	FG05 408	Cirad (France)	44	B78-505	Barbados
20	FG03 520	Cirad (France)	45	C132-81	Cuba
21	FGo4 754	Cirad (France)	46	C86-12	Cuba
22	FG04 466	WICSCBS (Barbados)	47	C90-501	Cuba
23	FG03 526	Cirad (France)	48	B52-298	Barbados
24	Mex54/245	Mexico	49	CO- 678	India
25	FG03 396	Cirad (France)	-	-	-

have been screened half way at the sugarcane breeding scheme in Cirad (France) and those varieties whose name starts with other than PG are advanced breeding clones at the final testing stage before a possible commercial release.

Experimental design and layout

The experiment was implemented with partially balanced square lattice design and was repeated (replicated) three times. Plot size for a genotype per replication was 8.7 × 6 m (52.5 m^2) with four test rows and two guard rows. Moreover, the design contains 7 blocks per replication and each block had an area of 8.7 m (width) × 48 m (length) = 417.6 m^2 and the experimental area per location was 0.78 hectares. Each replication was defined as replication nested in each location because the replications were unique for location and each block was nested within both replications and location. At planting, 18 two budded sets were laid on a furrow with 5 m length and cane was harvested at 17 and 13 months cane age for plant cane and ratoon crops, respectively. All recommended agronomic and cultural practices were uniform to raise the crop across all the sugar estates.

Morphological and agronomic characteristics

Cane and cane yield components

Sprout percentage: The percentage of setts which sprout 45 days after planting was calculated as the numbers of setts sprouted divided by the numbers of setts planted per plot and multiplied by 100, while the number of tillers the (ha^{-1}) per plot (from the central test rows) was counted 4 months after planting and was converted on hectare basis. For average numbers of internodes per stalk, milleable stalk diameter and stalk height (cm), 20 randomly selected milleable stalks per plot were considered and only the average values were reported. For estimation of cane yield (tons $ha^{-1}m^{-1}$), all milleable stalks from the central four rows per plot were hand trashed to remove the leaves and hand topped at the natural breakpoint of sugarcane stalk. The milleable stalks were then weighted using digital scale balance to the weights per plot and was extrapolated to tons ha^{-1} m^{-1}

Sugar yield and yield quality traits

Recoverable sucrose percentage refers to the total recoverable sugar percent in the cane and was calculated as recoverable sucrose percentage = [Pol% - (Brix - Pol%) 0.61] 0.75) as described by Berg (1972), where 0.61 = non-sugar factor, representing the amount of sucrose lost in final process and 0.75 = cane factor, representing the correlation factor between theoretical yields of molasses mixed juice and primary juice for the same genotype and the same cut of cane determined by milling test. Pol and Brix in cane refers to Pol and Brix percentage in cane and were determined as Pol in juice × (100-(fiber%+5))/100 and Brix in juice × (100-(fiber%+3))/100, respectively. Moreover, sugar yield (ton/ha) was estimated as the product of cane yield per hectare and average estimated recoverable sucrose percentage, and was computed as sugar yield = [Cane Yield (t/ha) × Recoverable

Table 2. Combined analysis of variances for 49 sugar cane genotypes (G) evaluated across 12 test environments (Location × Crop Years).

Parameter	Sources of variation						CV	Mean
	Environment	Genotype	GxE	Rep (Env)	Block (Rep*Env)	IBE		
DF	**11**	**48**	**528**	**24**	**144**	**432**		
Traits								
Sprout%	51676.93**	1117.7**	799.67**	1090.17**	108.1ns	97.25	15.13	65.17
TN	1695948.57**	1.04E+^{12}*	7954.96**	5159.09**	1.9x10^{7}*	8.9x10^{7}	14.37	207900
DM	8.18**	0.59**	0.17**	0.62**	0.05**	0.14	7.18	2.67
MSH	94.11**	1.29**	0.39**	0.61*	0.09ns	0.08	11.63	2.39
MSP	76279.85**	2985.5**	1206.57**	605.99ns	442.3**	219.4	14.79	100150
NIPS	2167**	71.45**	22.44 ns	59.46*	17.08ns	20.41	17.03	26
CYLD	1646.65**	40.95**	16.36**	13.58**	4.27ns	3.82	19.83	9.85
Pol%	141.65**	12.74**	9.98**	7.79ns	3.34**	2.38	8.59	18.16
Brix%	242.36**	9.54**	5.08*	2.81ns	1.44**	3.45	5.67	20
Purity%	1068.96**	22.64**	14.43**	10.72ns	4.99**	3.8	2.21	90.02
RS%	70.74*	6.33**	3.56**	2.64ns	1.22**	0.68	6.62	12.47
SYLD	22.06**	0.60*	0.28**	0.18*	0.08ns	0.065	20.87	1.22

**Highly significant at p<0.01; *Significant at P<0.05; ^{ns}Non-significant; rep: Replications; IBE: intra block error; DF: degree of freedom; DM: Diameter; NT: number of tillers (ha^{-1}); Pol: Pol in juice (%); RS%: recoverable sucrose percentage (%); Brix: Brix in juice (%); MSH: milleable stalk height (m); DM: milleable stalk diameter (mm); MSP: milleable stalk population (ha^{-1}); NIPS: numbers of internodes per stalk; SYLD: sugar yield (t ha^{-1}m^{-1}); CYLD: cane yield (t ha^{-1}m^{-1}).

sucrose percentage] / 100. As the plant cane crops and ratoon crop were harvested at 17 and 14 months age of cane, respectively, data for cane and sugar yield were converted to t ha^{-1}m^{-1} (tons per hectares per month) to bring the crops types to the same productivity unit.

Statistical analysis

The data collected for each trait were subjected to combined analysis of variances (ANOVA) and using SAS program and data quality was checked to meet the assumptions of ANOVA and block effects were using SAS software package, 2009. Genotypic means were adjusted for the lack of orthogonality (intra or inter block) in the data depending on the relative magnitude of the block variance relative to the residual error as suggested by Federer et al. (2001) and the adjusted means were used for multivariate analyses. For cluster analysis, average linkage was obtained by specifying METHOD=AVERAG as adopted by Sokal and Michener (1958), while Euclidean distance and linear discrimination analysis were computed using Minitab v. 17. Moreover, multiple regressions were analyzed using GENSTAT (Edition 13th), while the path coefficient analysis was done using the SAS software package (SAS, 2009) and SAS program of PROC MATRIX and PROCIML as suggested by Kang (1994).

RESULTS AND DISCUSSION

The pooled analysis of variance (Table 2) showed that the variances for genotypes were highly significant (P<0.01) for all traits, suggesting there was an ample genetic variability among the genotypes. As the variability among genotypes was highly significant for all traits studied, conducting of multivariate analyses using these traits will be relevant to generate further analysis. Hence, means for sprout percentage (%), number of tillers (ha^{-1}), milleable stalk height (m), milleable stalk diameter (cm), milleable stalk population (ha^{-1}), numbers of internodes per stalk, cane yield (tons ha^{-1} m^{-1}), Brix in juice, Pol in juice, purity, recoverable sucrose percentage and sugar yield (t ha^{-1} m^{-1}) were adjusted. The adjusted means of the traits studied were subjected to cluster, linear discriminant, principal component, path coefficient and multiple regression analyses (Table 3).

Cluster analysis

Based on the adjusted means of 12 sugarcane traits presented in Table 3, results of the cluster analysis indicated that the 49 sugarcane genotypes formed 10 distinct groups or clusters where 3 of the groups contained a single genotype (Table 4). Starting from left to right of the Dendrogram (Figure 1), clusters number II comprised much of the genotypes studied (17 genotypes) followed by cluster V (8) and I (7). Cluster 1 consisted of 7 genotypes that have different origins but introduced from France. Cluster II consists of 17 genotypes (9, 36, 14, 29, 40, 13, 37, 23, 31, 42, 41, 44, 32, 13, 30, 43, 18, 45 and 48) which were a mixture of commercial varieties and introduced genotypes. It was also observed that most of the commercial varieties were grouped in one cluster except genotypes 24 and 49 form another separate group. Genotypes 7, 19 and 34 were ungrouped, suggesting these genotypes were outliers for lower or higher mean values of the traits studied.

Based on the grouping, the relationships observed

Table 3. Adjusted means for 12 traits in 49 sugarcane genotypes used for multivariate analyses.

Genotypes		Traits											
Cd	**Name**	**Sprout%**	**NT**	**Pol %**	**Purity**	**RS%**	**Brix%**	**Cyld**	**HT**	**DM**	**MSTP**	**NIPS**	**SYLD**
1	PSR97 092	61.6	184034.4	17.69	89.46	12.08	19.63	12.7	2.643	3.03	87813.33	14.9	1.521
2	DB70047	64.5	159785.84	18.92	90.52	13.26	21.06	9.38	2.639	2.79	86312.57	16	1.226
3	DB66 113	80.5	242950.29	17.25	89.66	11.91	19.13	11.9	2.436	2.66	118483.1	12.89	1.409
4	FG06 700	58.6	241346.91	18.5	90.08	12.52	20.1	8.35	2.483	2.22	105329.5	17.52	1.023
5	FG06 729	62.7	183176.18	19.67	91.44	13.46	21.05	9.4	2.703	2.67	86440.39	17.65	1.262
6	PSR97 087	71.6	172225.55	18.75	89.45	12.65	20.48	9.81	2.305	2.97	73518.3	17.26	1.229
7	PSR97 051	86.6	251177.41	19.02	89.65	13.05	21.04	10.6	2.38	2.78	91003.73	15.19	1.37
8	HO95 988	70.2	216704	18.65	91.56	12.98	20.29	8.94	2.242	2.26	137152.6	16.21	1.154
9	Cp99 1534	55.3	185746.53	17.71	90.12	12.36	19.87	8.75	2.107	2.63	95732.89	14.42	1.078
10	FG04 829	58.1	196837.72	19.3	90.11	13.26	21.27	9.31	2.203	2.49	116522.9	16.91	1.232
11	DB71 060	56.5	181090.09	18.63	90.18	12.8	20.52	10.8	2.235	3.12	75663.11	15.03	1.377
12	TCP93 4245	71.7	202129.37	19.82	91.84	13.65	21.19	9.42	2.625	2.47	104872.6	18.98	1.291
13	CP001 252	61.3	200516.75	17.93	91.2	12.54	19.67	9.37	2.451	2.59	100506.4	13.87	1.171
14	VMC95 173	63.7	207430.05	17.41	90.42	12.03	19.12	9.27	2.546	2.62	87714.31	14.36	1.128
15	FG03 447	49.7	233412.46	18.47	91.2	12.94	20.46	9.79	2.614	2.54	105984.5	15.4	1.245
16	CO 740	55.5	200983.62	17.47	89.28	11.78	19.07	11.2	2.257	2.8	107522.9	12.25	1.31
17	CP99 1894	75.3	245477.14	18.56	88.87	12.69	20.78	7.17	2.131	2.39	108191.2	16.71	0.91
18	FG03 425	59.1	205614.96	17.41	87.82	11.73	19.68	9.13	2.441	2.71	94010.43	15.28	1.064
19	FG05 408	55	236234.46	18.72	93.04	13.23	20.03	8.33	2.01	2.66	97390.73	15.37	1.121
20	FG03 520	78.5	229228.81	18.04	88.84	12.41	20.15	12	2.954	2.77	96965.76	15.14	1.49
21	FGo4 754	67.2	214073.9	17.78	88.75	12.28	20.16	11.5	2.322	2.94	95710.55	13.18	1.399
22	FG04 466	53.3	172227.5	18.38	90.15	12.66	20.25	11.8	2.284	2.84	101854.3	12.68	1.475
23	FG03 526	65.4	203918.56	18.55	90.8	12.82	20.33	9.47	2.434	2.7	100986.7	13.77	1.21
24	Mex54/245	54.4	172808.67	17.07	88.48	11.63	19.16	10	2.832	2.58	94102	13.36	1.145
25	FG03 396	74.7	194664.81	19.19	90.93	13.31	21	9.82	2.448	2.63	99619.4	15.21	1.309
26	VMC95 212	77.3	257855.02	17.85	90.03	12.21	19.57	10.9	2.354	2.58	110613.5	14.24	1.32
27	NCO-334	65.6	237816.99	16.96	88.34	11.51	19	11.1	2.548	2.44	136396.9	13.76	1.271
28	FG03 418	69.8	184731.11	19.06	91.23	13.24	20.79	11.7	2.527	2.64	108611.4	16.52	1.54
29	CO449	63.4	181656.15	17.83	90.31	12.27	19.59	8.88	2.477	2.62	100092.2	13.29	1.08
30	FG03 204	54.9	191504.07	18.21	91.06	12.62	19.9	10.2	2.63	2.69	98034.82	15.83	1.274
31	FG02 553	67	202959.95	17.96	90.29	12.34	19.77	9.67	2.124	2.65	102663.2	14.43	1.17
32	FG03 103	68.1	225035.79	17.71	90.35	12.31	19.57	10.7	2.643	2.65	90884.98	14.41	1.314
33	FG03 318	56.2	163633.75	17.92	89.55	12.29	19.9	10.2	2.725	2.78	82001.93	14.27	1.268
34	FG04 708	77.8	301267.25	18.61	89.41	12.73	20.69	6.45	1.676	2.53	113309.4	14.48	0.826
35	FG04 705	75.7	263476.38	19.18	89.91	13.23	21.23	8.85	2.199	2.49	103371.2	17.48	1.156

Table 3. Contd.

36	FG02 551	55.3	160240.66	17.46	90.2	12.21	19.55	9.42	2.281	2.71	97298.79	14.72	1.147
37	FG03 173	62.3	217494.25	18.21	90.65	12.59	20.01	9.25	2.277	2.56	99489.64	13.13	1.164
38	FG04 187	84.1	232984.5	17.14	88.46	11.62	19.2	13.1	2.669	2.58	113411.6	13.04	1.506
39	FG03 372	65.4	169686.89	18.53	90.05	12.75	20.45	11.3	2.844	2.66	97016.87	14.9	1.436
40	FG03 214	64.2	189855.88	17.64	89.49	12.12	19.56	9.03	2.445	2.52	99992.32	13.44	1.091
41	C86-56	71.5	241200.44	17.33	90.41	12.13	19.46	9.92	2.227	2.72	103595.6	12.77	1.199
42	SP70-1284	67.1	240728.83	17.91	90.17	12.3	19.74	9.32	2.143	2.75	109318.3	14.9	1.135
43	C86-165	50.7	221037.66	17.49	91.33	12.17	19.26	10.3	2.278	2.6	102467.5	14.98	1.25
44	B78-505	60.2	224108.15	17.93	90.23	12.39	19.74	9.82	2.376	2.9	93891.2	14.47	1.195
45	C132-81	71.4	198182.45	17.45	88.71	11.94	19.49	9.76	2.132	2.94	86807.38	13.65	1.152
46	C86-12	65	164182.53	18.76	90.34	12.75	20.36	9.74	2.23	2.77	97970.94	18.4	1.244
47	C90-501	68.2	167499.59	20.55	90.36	12.74	20.32	9.81	2.135	2.76	94214.57	16.49	1.256
48	B52-298	73	252441.62	17.18	88.46	11.67	19.28	9.04	2.148	2.62	100737	13.22	1.044
49	CO- 678	48.2	168277.54	16.36	87.9	11	18.46	10.3	2.688	2.69	95908.11	12.62	1.123

*NT: Number of tillers (ha^{-1}); Pol: Pol in juice (%); RS: recoverable sucrose percentage (%); Brix: Brix in juice (%); MSH: milleable stalk height (m); DM: Milleable stalk diameter (mm); MSTP: milleable stalk population (ha^{-1}); NIPS: numbers of internodes per stalk; SYLD: sugar yield (t $ha^{-1}m^{-1}$); CYLD: cane yield (to $ha^{-1}m^{-1}$); CD: code.

Table 4. Clusters of 49 sugar cane genotypes based on traits contributing to sugar yield.

Clusters	Genotypes	Freq.	%
I	1, 21, 22, 16, 33, 39, 20	7	14.28
II	9, 36,14, 29, 40, 13, 37, 23, 31, 42, 41, 44, 32, 13, 30, 43, 18, 45, 48	17	34.69
III	24, 49	2	4.08
IV	3, 26 38, 27	4	8.16
V	2, 5, 25, 28, 12, 10, 46, 47	8	16.32
VI	7	1	2.04
VII	6, 11	2	4.08
VIII	19	1	2.04
IX	4, 17, 35, 8	4	8.16
X	34	1	2.04

*Freq: Frequency; 1-49=codes of the genotypes and are given in Tables 1 and 3.

among these genotypes had no any correspondence with the geographic origin (from where they were introduced). This suggested that the genotypes of different geographic origin had genetic similarity and genotypes of the same geographic origin had different genetic background,

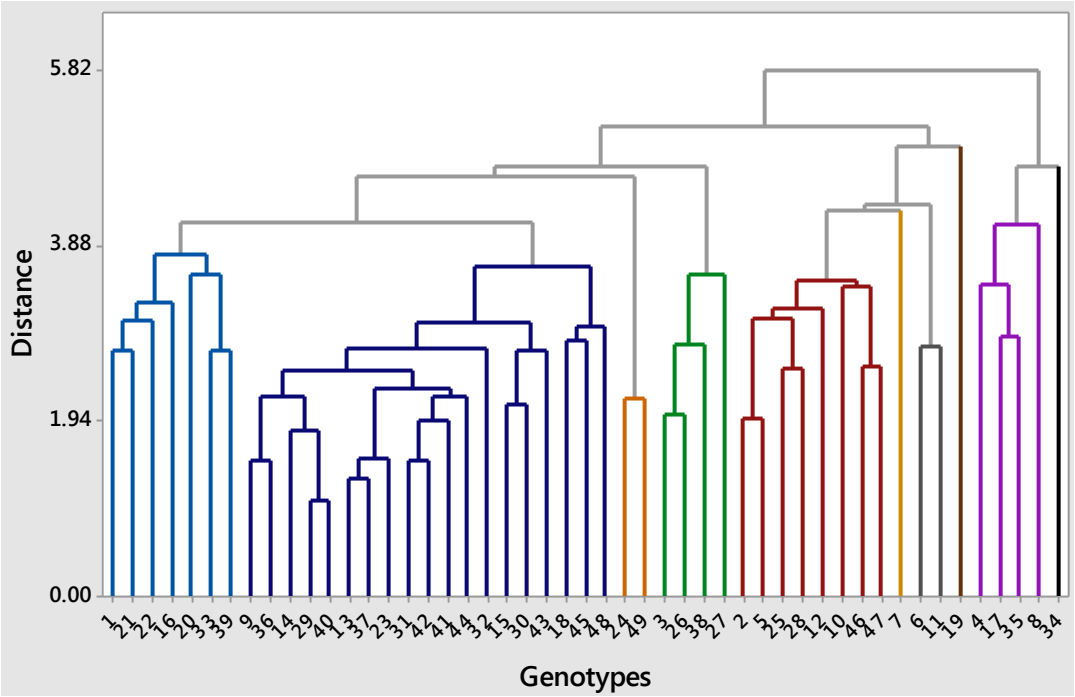

Figure 1. Dendrogram of 49 sugar cane genotypes based on Euclidean distance.

Table 5. Euclidean distances between cluster groups.

CL	I	II	III	IV	V	VI	VII	VIII	IX	X
I	0	4.1265	3.7805	5.7723	3.4236	4.3103	2.8287	5.8858	4.285	8.158
II	4.12653	0	6.0381	3.7624	3.6033	3.7174	3.8664	4.21976	7.0447	7.0159
III	3.78046	6.0381	0	5.5644	6.4863	5.193	3.8829	6.99782	4.9947	7.4587
IV	5.77226	3.7624	5.5644	0	6.0394	4.7534	3.909	4.33017	7.0749	4.3024
V	3.42356	3.6033	6.4863	6.0394	0	4.1943	3.964	5.32772	6.3177	7.8924
VI	4.31028	3.7174	5.193	4.7534	4.1943	0	4.5438	5.78581	7.6557	7.6557
VII	2.82874	3.8664	3.8829	3.909	3.964	4.5438	0	4.04076	4.0382	5.989
VIII	5.8858	4.2198	6.9978	4.3302	5.3277	5.7858	4.0408	0	7.685	5.8337
IX	4.28496	7.0447	4.9947	7.0749	6.3177	7.6557	4.0382	7.68504	0	8.9809
X	8.15798	7.0159	7.4587	4.3024	7.8924	7.6557	5.989	5.83368	8.9809	0

*Cl: Clusters.

suggesting that a high proportion of total genetic variation was retained within the groups of origin and active genetic ex-change was found between different origins. This relationship suggests the introduction strategy was successful in terms of improving the base of the crop in Ethiopia and increases the chances of selection efficiency during parental selection in the future using the traits that contributed more to the existed phenotypic diversity. Similar results were also observed by Ram and Hemaprabha (1998) and Tahir et al. (2013) in which they found the progenies of a cross clustered independently of their parents. Hence, our introduction strategy was appropriate in terms of broadening the narrow genetic pool of sugarcane in Ethiopia.

Euclidean distances between clusters groups and contributions of variable (Traits) to diversity

Distances between clusters groups based on the Euclidean Distances statistic (Table 5) revealed that groups I, II, III, V, VI and IX had the highest distances to group X with a single genotype (34) suggesting the genotype was an outlier. Moreover, cluster group IX was

Table 6. Step wise order inclusion of variables in the discriminant analysis.

Step number	Multivariate statistics					Statistic	
	Trait	Trait entered	R-Square	Partial R-Square	Trait removed	Wilks' Lambda	Pillai's Trace
1	Pol	Pol	0.7790**	-	No	0.0022**	0.7798**
2	Pol	Pol	0.7790**	0.7676**	No	0.0615**	1.4974**
	Cyld	Cyld	0.7352**	0.7215**			
3	TN	TN	0.6629**	0.6438**	No	0.0219**	2.0788**
	Pol	Pol	0.7790**	0.7610**			
	Cyld	Cyld	0.6426**	0.7083**			
4	TN	TN	0.6529**	0.6066**	No	0.0107**	2.5482**
	Pol	Pol	0.7790**	0.7619**			
	Cyld	Cyld	0.7352**	0.6602**			
	MSP	MSP	0.6424**	0.5086**			
5	TN	TN	0.6629**	0.6190**	No	0.0055**	2.9653**
	Pol	Pol	0.7790**	0.7545**			
	Purity	Purity	0.4923**	0.4830*			
	Cyld	Cyld	0.7352**	0.6862**			
	MSP	MSP	0.6424**	0.5218*			
6	Sprout	No		0.2130ns	No further steps possible		
	RS			0.1711ns			
	Brix			0.1918ns			
	Ht			0.1623ns			
	DM			0.1926ns			
	NIPS			0.2843ns			
	SYLD			0.2530ns			

**Highly significant at P<0.01; *Significant at P<0.05; ^{ns}Non-significant; DM: Diameter; NT: number of tillers (ha^{-1}); Pol: Pol in juice (%); RS%: recoverable sucrose percentage (%); Brix: Brix in juice (%); MSH: milleable stalk height (m); DM: milleable stalk diameter (mm); MSP: milleable stalk population (ha^{-1}); NIPS: numbers of internodes per stalk; SYLD: sugar yield (t ha^{-1}m^{-1}); CYLD: cane yield (t ha^{-1}m^{-1}).

more distanced from cluster groups II (7.016), III (7.461), IV (7.070), VI (7.651) and VIII (7.680); genotypes of each group had ample diversity and can be crossed with genotypes in groups IX. On the contrary, the smallest distance was observed between cluster groups I and VII, the diversity between the groups was narrow. Generally, the smallest and larger distances among cluster groups suggest high probability of getting divergent genotypes that are useful for crossing purposes.

A step wise discriminant analysis by minimizing the Wilk's criteria (Table 6) resulted in significant F-values for Pol%, cane yield, number of tillers; Purity% and milleable stalk population, suggesting that these traits contributed more to the discrimination among the groups.

Results of different studies demonstrate that the linear discrimination function is a usefully tool for screening and evaluating variation among sugarcane genotypes studied. Moreover, the step wise discrimination procedure provided in Table 6 indicated that Pol in juice, cane yield (t ha^{-1}m^{-1}), number of tillers (ha^{-1}), purity% and milleable stalk population (ha^{-1}) significantly explained total variability (R^2) of 77.90, 73.52, 66.29, 64.24 and 49.23%, respectively; revealing these traits contribute more to the diversity which existed among the 49 sugarcane genotypes. This result was inconsistent with findings reported by Kang et al. (2013) in which Brix and juice contents contributed more to divergence among genotypes. It can be concluded that the 49 sugarcane genotypes were diversified for Pol in juice; cane yield (t ha^{-1}m^{-1}), number of tillers (ha^{-1}), purity% and milleable stalk population (ha^{-1}).

Principal component analysis

Biplot of principal component analysis based on

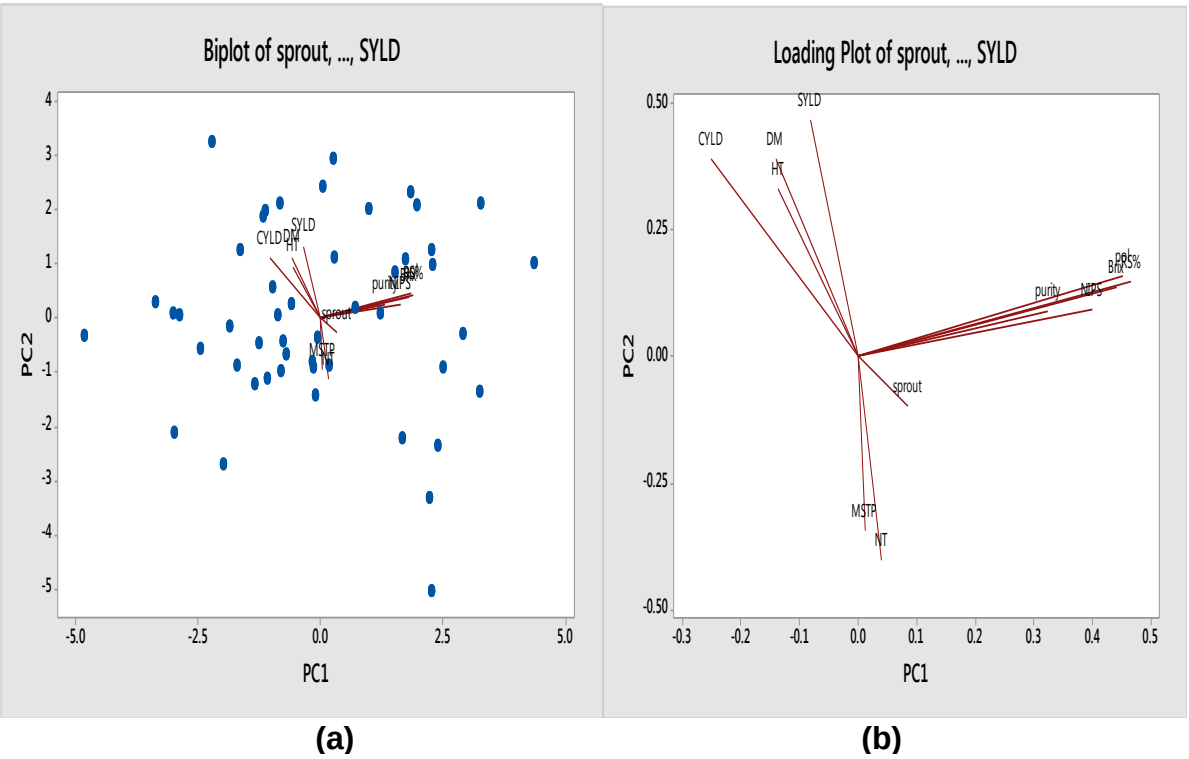

Figure 2. (a) Biplot distribution of 49 sugar cane genotypes displaying 12 traits and (b) loading plot distribution of 12 traits. DM: Diameter; Sprout NT: number of tillers (ha^{-1}); Pol: Pol in juice (%); RS: recoverable sucrose%; Brix: Brix in juice (%); HT: milleable stalk height (m); DM: milleable stalk diameter (mm); MSTP: milleable stalk population (ha^{-1}); NIPS: numbers of internodes per stalk; SYLD: sugar yield (t $ha^{-1}m^{-1}$); CYLD: cane yield (t $ha^{-1}m^{-1}$).

correlation matrix is depicted in Figure 2. It was sufficient enough to show correlation among variables considered. 58, 96 and 100% of the variation existed among the genotypes was explained by the first two, eight and ten principal components, respectively (Figure 2a). The Biplot and loading plot (Figure 2b) were able to separate cane yield and its components (large PC2 score) from the quality traits (large PC1) highlighting characterization of genotypes in terms of traits would be possible using principal component analysis.

As far as the relationships (correlation) existed among the traits is concerned, cane yield, milleable stalk height, diameter and sugar yield were characterized with large PCA2 score and were more correlated (angles among the specified traits were acute) which agreed with reports of Rewati and Joshi (2005) in which milleable stalk height and diameter were positively correlated with cane yield. Selection of sugarcane genotypes based on cane yield, milleable stalk height and diameter increased sugar yield which is consistent with the results reported by Khan et al. (2013) and Masri et al. (2015). On the contrary, quality parameters such as purity%, pol in juice, Brix in juice and recoverable sucrose percentage were characterized with large PCA 1 and had small angles among themselves; indicating strong and positive correlation. Moreover, milleable stalk population and tiller number were characterized with small PCA score and were negatively correlated with cane yield and sugar yield which disagreed with the results of Punia et al. (1983). Sprout percentage was positively correlated with milleable stalk population which was consistent with reports of Sahu et al. (2008) in which germination% showed a positive and significant correlation with number of millable canes.

Weak and positively correlation was observed between sugar yield Pol%, Purity%, Brix and numbers of internodes per single stalk. Furthermore, milleable stalk diameter made obtuse angle with numbers of internodes per stalk implying negative correlation disagreed with report of Kumar and Kumar (2014). The inconsistencies of the results might be attributed to the nature of quantitative traits which are more affected by environment and sampling error. Generally, weak correlation existed between sugar yields and recoverable sucrose percentage which is not expected as recoverable sucrose percentage is the main component of sugar yield. Thus, additional analyses such as path coefficient analysis, which enable to compute indirect effects of secondary traits on dependent trait, should be used.

Results obtained from PCA analysis were supported by linear discriminant analysis in that Pol% and cane yield which contributed more to the total variability in the linear discriminant analysis had long vectors in the Biplot

Table 7. Phenotypic path coefficients showing direct (diagonal) and indirect effects (off diagonal) of 11 sugarcane traits on sugar yield.

Traits	Sprout%	Tiller	Pol	Purity	RS%	Brix	CYLD	Height	MSP	DM	NIPS
sprout	**0.01358**	-0.00722	0.003191	-0.0001	0.058078	-0.01348	0.08917	-0.000157	0.000323	-0.00373	-0.00048
Tiller	0.006932	**-0.01414**	-0.00165	-0.00016	-0.01341	0.000183	-0.23835	-0.000519	0.001292	-0.00822	0.000398
Pol	0.002567	0.001379	**0.016882**	2.71E-05	0.378162	-0.04761	-0.26173	-0.000217	0.000231	0.001669	-0.00373
Purity	-0.00247	0.001003	0.00939	0**.000933**	0.300446	-0.02188	-0.2226 -	-0.000156	0.000549	-0.0005	-0.002
RS%	0.001891	0.000455	0.015309	0.000375	**0.41702**	-0.04938	-0.29027	-0.000172	0.000372	0.001119	-0.0035
Brix	0.003425	4.85E-05	0.015042	5.85E-05	0.385327	**-0.05344**	-0.2809	-0.000174	0.000223	0.00166	-0.00355
CYLD	0.001178	0.003278	-0.0043	-0.0002	-0.11771	0.014596	**1.028372**	0.0007413	-0.00146	0.000587	0.001556
Height	-0.00152	0.005261	-0.00262	-0.0001	-0.05149	0.006641	0.545963	**0.0013964**	-3.3E-05	0.003156	3.85E-05
MSP	-0.00134	0.005588	-0.00119	-0.00016	-0.04739	0.003649	0.45961	0.0000142	**-0.00327**	0.011761	0.001077
DM	0.002948	-0.00677	-0.00164	2.71E-05	-0.02719	0.005165	-0.03515	-0.000257	0.002239	**-0.01717**	0.000358
NIPS	0.001305	0.001134	0.012663	0.000375	0.293461	-0.03824	-0.32222	-0.000011	0.000709	0.001238	**-0.00497**

*Coefficient of determination=0.99 and residual=0.06; NIPS: number of internodes per stalk; CYLD: cane yield; DM: milleable stalk diameter; TN: numbers of tillers; SYLD: sugar yield; MSP: milleable stalk population; HT: milleable stalk height.

(Figure 2a) and loading plot (Figure 2b), indicating these traits contributed more to the total variation explained by the first two dimensions. This suggests that the PCA and linear discriminate analysis were similar in identifying the traits which were dominant in explaining the existing variability among the genotypes. Moreover, these traits can be further used to discriminate cluster groups and are helpful for parent selection in sugarcane breeding programs as the variation existed among the genotypes was highly significant for these traits (Table 2).

Path coefficient analysis

The gap in principal component or correlation analysis with respect to causal relationships among traits necessitates path coefficient analysis to be used in selection to utilize both direct and indirect relationships among traits (Kang, 2015). Thus, the use of path coefficient analysis will be mandatory to increase the efficiency of our selection. The path coefficient analysis presented (Table 7) indicated that the highest positive direct effect on sugar yield was exerted by cane yield (1.028) followed by recoverable sucrose percentage (0.417) and Pol (0.016) which is consistent with report of Khan et al. (2013). In the contrary, small negative direct effects on sugar yield were exerted by number of tillers, Brix, stalk diameter, milleable stalk population and numbers of internodes per stalk. Furthermore, milleable stalk height and milleable stalk population exerted indirect effect of 0.546 and 0.459, respectively via cane yield, while Brix% in juice, Pol% in juice, Purity% and number of internodes had indirect effects of 0.385, 0.378, 0.300 and 0.293, respectively via recoverable sucrose percentage. Hence, these traits should be given due consideration during selection for high sugar yield. This result was similar with reports of Khan et al. (2012) in which higher number of tillers, good weight, endowed with better available sugar in the cane (Pol%), commercial cane sugar (CCS)% and purity% were the important characters which should be considered in selection of higher sugar yield in sugarcane genotypes.

Selection of sugarcane genotypes on the basis of cane yield and recoverable sucrose percentage (%) would be beneficial for increasing sugar yield in sugarcane. Our result was in agreement with report of Hussein et al. (2012) except for the effect of number of milleable stalks to sugar yield in our study was negative and negligible. Moreover, the coefficient of determination and the residual effect in this study were 0.990 and 0.061, respectively suggesting that most of the variability in sugar yield was best explained by the traits studied (causal factors) and the error was negligible and thus, no additional traits is necessary to be included in selection. Generally, the path coefficient analysis, in the present study was sufficient enough to increase the efficiency of selection. For example, the weak correlation between sugar yield and recoverable sucrose%,

Table 8. Multiple linear regression model to explain sugar yield variation using its related characters.

Variable name	Estimate	Standard error
Sprout (%)	0.000244	0.000255
Numbers of tillers (ha^{-1})	-6.96E-08	7.26E-08
Pol%	0.00271	0.00540
Purity%	0.00092	0.00957
Recoverable sucrose%%	0.1080*	0.0421
Brix%	-0.0099	0.0258
Cane yield	0.12199**	0.00256
Milleable stalk height (m)	0.0008	0.0118
Milleable stalk populations(ha^{-1})	-0.000000212	0.000000262
Milleable stalk diameter (cm)	-0.0026	0.0218
Number of internodes per stalk (ha^{-1})	-0.00039	0.00157
Intercept	-1.232	0.883
Model significance	<001	-
R^2	0.995	-
R^2 of eliminated traits	0.00413	-

**Highly significant at p<0.01;*Significant at p<0.05.

in the principal component analysis, was ruled out by path coefficient analysis in such a way that it was able to compute the direct effect of recoverable sucrose percentage (0.417) and its highest indirect effects via Brix% in juice (0.385), Pol% in juice (0.378) and Purity% (0.300).

Multiple linear regressions analysis

Although path coefficient analysis provides a picture of the pattern of association, it cannot construct a prediction equation for dependent variable using its components (El-Shafi and Ismail, 2006). For this reason, multiple regressions were used to develop the regression model. SY (Sugar yield) =-1.232+0.000244 (sprout %) +-6.96E-08 (Tiller Numbers) +0.00271 (Pol %) +0.00092 (Purity%) +0.1080 (Yield %) +-0.0099 (Brix %) +0.12199 (cane yield) + 0.0008 (stalk height) + -0.000000212 (stalk population) + -0.0026 (stalk diameter)-0.00039 (number of internodes). The result presented in Table 8 demonstrated that 99.5% of the variability is explained as R^2 and the rest 0.41% is attributed to unknown variation. Furthermore, the multiple linear regressions indicated that recoverable sucrose percentage and cane yield significantly contributed to sugar yield which is similar to the results reported by Hussein et al. (2012) in which recoverable sucrose percentage and stalk weight contributed more to sugar yield.

Conclusion

The multivariate analysis generates relevant information about the performance of the genotypes, relationship among genotypes and interrelationships among traits which is very important for sugarcane breeding programmes. The cluster analysis demonstrated that the 49 sugarcane genotypes studied were clustered into ten groups and were highly significantly different for Pol in juice, cane yield (t $ha^{-1}m^{-1}$), number of tillers (ha^{-1}), purity% and milleable stalk population (ha^{-1}). The relationship existed among sugarcane genotypes studied was not related to their geographic origin, suggesting that a high proportion of total genetic variation was retained within the groups of origin and active genetic ex-change was found between different origins indicating that the introduction strategy was successful.

Generally, regarding with the interrelationships among the traits, more information was generated following the use of principal component, linear discriminant, path coefficient and multiple regression analyses indicating the use of multivariate analyses was successful. Path coefficient was unique in generating information about the indirect effects of traits on sugar yield which was very important to provide substantial information about indirect effects of traits that are very relevant to increase selection efficiency in sugarcane plant breeding programs.

Conflict of interests

The authors have not declared any conflict of interests.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the Sugar Corporation

Research and Training for providing necessary inputs on time. Due appreciation goes to Sugar cane Research Stations Team leaders, for the follow up of the experiments.

REFERENCES

Ahmed AO, Obeid A, Dafallah B (2010).The influence of characters association on behavior of sugarcane genotypes (Saccharum Spp.) for cane yield and juice quality. World J. Agric. Sci. 6(2):207-211.

Chohan M, Panhwar RN, Memon MA, Unar GS, Mari AH (2007). Performance of new sugarcane varieties for cane yield and quality under different agro-climatic conditions of Sindh. Pak. J. Sci. Res. 59:28-33.

Dewey DR, Lu KH (1959). A correlation and path coefficient analysis of components of crested wheat grass seed production. Agron. J. 51:515-518.

El-Shafai AM, Ismail AMA (2006). Effect of row spacing on yield and quality of some promising sugarcane genotypes. Egypt. J. Appl. Sci. 21(11):32-46.

Hussein M, Al-Sayed H, Fateh S, Fares WM, Attaya S (2012). Crops Multivariate Analysis of Sugar Yield Factors in Sugar Cane. Sugar Res. Inst., Agricultural Researches Center, Giza, Egypt. Am. Eurasian J. Sustain. Agric. 6(1):44-50.

Kang MS (1994).Applied quantitative genetics. Baton Rouge. Agron. J. 84:917-918.

Kang MS (2015). Efficient SAS Programs for Computing Path Coefficients and Index Weights for Selection Indices. Department of Plant Pathology, Kansas State University, Mahattan, Kansas, USA. J. Crop Improve. 29:6-22.

Kang MS, Sosa O, Miller JD (1989). Path analysis for percent fiber andcane and sugar yield in sugarcane. Crop Sci. 29:1481-1483.

Khan IA, Bibi S, Yasmin S, Khatri A, Seema N, Abro AS (2012). Correlation studies of the agronomic traits for higher sugar yield in sugarcane. Pak. J. Bot. 44(3):969-971.

Khan IA, Nighat, S, Raza S, Yasmine S, Sajida B (2013). Environmental Interaction of Sugarcane Genotypes and Yield Stability Analysis of Sugarcane. Nuclear Institute of Agriculture, Tando Jam, Pakistan, Pak. J. Bot. 45(5):1617-1622

Kozak M, Kang MS, Stepien M (2007). Causal pathways when independent variables are co-related: New interpretational possibilities. Plant Soil Environ. 53(6):267-275

Kumar S, Kumar D (2014). Correlation and path coefficient analysis in sugarcane germplasm under subtropics. Department of Genetics and Plant Breeding. CCR (P. G.) College, Muzaffar Nagar (U.P), India. Afr. J. Agric. Res. 9(1):148-153.

Malik SR, Bakhsh A, Asif MA, Iqbal U, Iqbal SM (2010). Assessment of genetic variability and interrelationship among some agronomic traits in chickpea. Int. J. Agric. Biol. 12:81-85.

Masri MI, Shaban SHA, El-Hennaway, Abu El-lail (2015). Phenotypic and Genotypic Correlations in Sugarcane at First Clonal Selection Stage. Agronomy Depart. Fac. of Agric. and Natural Resources. Aswan Univ. Aswan, Egypt. Sugarcane Research institute. Egypt. J. Plant Breed. 19(2):297-321.

Mohammadi SA, Prasanna BM (2003). Analysis of genetic diversity in crop plants-salient statistical tools and considerations. J. Crop Sci. 43(4):1235-1248.

Ong'ala J, Mwanga D, Nuani F (2016). Principal Component Analysis in Sugarcane Clone Selection. Kenya Agricultural and Livestock Research Organization, Sugar Research Institute, Nairobi, Kenya. J. Indian Soc. Agric. Stat. 70(1):33-39.

Punia MS, Paroda R, Hooda RS (1983). Correlation and path analysis of cane yield in sugarcane. Indian J. Genet. Plant Breed. 43:109-12.

Ram B, Hemaprabha G (1998). Nature and pattern of genetic divergence of sugar yield and its components in the progenies of *Saccharum barberi*. Indian J. Gen. Plant Breed. 58(2):193-199.

Rewati RC, Joshi DK (2005).Correlation and Path Coefficient Analyses in Sugarcane. Sugarcane Research Programme-NARC, Jitpur-Bara-Nepal. Nepal. J. Agric. Res. 6:25-26.

Sahu RS, Pandey SS, Kamat DN (2008). Correlation studies between seedlings and their settings in inter varietal crosses of sugarcane for yield and quality traits. Indian Sugar 50(2):27-35.

Sidwell RJ, Smith EL, McNew RW (1976). Inheritance and interrelationships of grain yield and selected yield-related traits in a hard red winter wheat cross. Crop Sci. 16:650-654.

Sokal RR, Michener CD (1958). A Statistical Method for Evaluating Systematic Relationships. Univ. Kans. Sci. Bull. 38:1409-1438.

Tahir MH, Rahman RG, Ali A, Khalid M (2013).Genetic Divergence in Sugarcane Genotypes. Sugar Crops Research Institute, Mardan, KPK, Pakistan. Am. J. Exp. Agric. 3(1):102-109.

Terzi FS, Rocha FR, Vencio RZ, Felix JM, Bran DS, Waclawovsky AJ (2009). Sugarcane gene associated with sucrose content. Retrieved from BMC Genomics 10(1):1.

Tyagi AP, Lai P (2007). Correlation and path coefficient analysis in sugarcane. South Pac. J. Nat. Sci. 1:1-10.

Wen YX, Zhu J (2005). Multivariable conditional analysis for complex trait and its components. Acta Genet. Sin. 32(3):289-296.

Zhu YJ, Albert HH, Moore PH (2000). Differential expression of soluble acid invertase genes in the shoots of high-sucrose and low-sucrose species of Saccharum and their hybrids. Austr. J. Plant Physiol. 27:193-199.