International Journal of Biosciences | IJB |
ISSN: 2220-6655 (Print) 2222-5234 (Online)
http://www.innspub.net
Vol. 13, No. 1, p. 412-419, 2018

INNSPUB

RESEARCH PAPER OPEN ACCESS

Sugarcane clone selection by FUZZ (True seed) at SUCAF/Ferké in Côte d'Ivoire

Oura Otchou Jean-Didier Thierry[1], Kouamé Konan Didier[*2], Péné Bi Crépin[1]

[1]Research and Development Department at SUCAF CI/SOMDIAA, Ferkessédougou, Côte d'Ivoire

[2]Félix HOUPHOUËT BOIGNY University of Côte d'Ivoire, Abidjan, Côte d'Ivoire

Key words: Multivariate analysis, Cluster analysis, Sugarcane, Breeding process

http://dx.doi.org/10.12692/ijb/13.1.412-419 Article published on July 30, 2018

Abstract

Knowledge on performance of genotypes and interrelationships among traits is very important for sugarcane breeding program. Therefore, many characters are evaluated simultaneously in the sugarcane phenotypic evaluation process. In this study, we used the Principal Component Analysis (PCA) to identify representative traits for phenotypic characterization of sugarcane, and thereby to select superior clones in the breeding process. Five quantitative and two qualitative traits of sugarcane were studied from the PCA. These major components represented for 58.13% of the variance. Cluster analysis has allowed us to divide the 148 clones of sugarcane into 5 groups. The high genotypes diversity of selected sugarcane is reflected by the genetic diversity revealed within the population.

* **Corresponding Author:** Kouamé Konan Didier ✉ didykonan@yahoo.fr

Introduction

Sugarcane (*Saccharum* spp.) is the source of about three quarters of the world's sugar, and is grown widely in the tropic and sub-tropics. It contributes approximately 80% of sugar to the world, greatly exceeding sugar beet as a sugar (Dahlquist, 2013). It is also used as a bioenergy crop due to its phenomenal dry matter production (Chohan *et al.*, 2014).

New sugarcane genotypes are produced by sexual means and vegetative reproduction. Each year, a new population of original seedlings consisting of thousands of new varieties is produced through fuzz (true seed). These are screened clonally through several selection stages, their numbers being reduced at each stage and the selected ones tested in larger plots in which their performance can be evaluated more reliably.

Major criteria for advancement from one stage to the next are high yields, quality index, disease resistance or tolerance and agronomic traits (Glaz *et al.*, 2002). A key goal of sugarcane breeding programs is to increase sugar yield by increasing sugarcane production per area, which is closely associated which is closely associated with height, diameter and number of the stalk, along with sugar accumulation in the stalk. Glaz and Miller (1982) reported that each stage results reasonably well predicted the commercial yields of released genotypes. The increasing multiple use of sugarcane requires a powerful breeding program allowing the identification of promising genotypes. Therefore, to achieve successful genetic improvement of sugarcane genotypes, effective and diversified breeding procedures must be followed. Selection will be effective if procedures take into account many traits simultaneously when evaluating sugarcane genotypes.

The Principal Component Analysis or canonical root analysis is a multivariate statistical technique attempt to simplify and analyze the inter relationship among a large set of variables in terms of a relatively a small set of variables or components without losing any essential information of original data set.

The PCA reduces relatively a large series of data into smaller number of components by looking for groups that have very strong inter-correlation in a set of variables and each component explained per cent (%) variation to the total variability.

The appropriate methods that provide accurate evaluations and estimation of genetic diversity depend on genetic diversity, 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.

Ong'ala *et al.* (2016) recommended PCA 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.

The PCA is one powerful statistical method widely applied to classify phenotypic traits in crop germplasm into groups based on similarities (Rukundo *et al.*, 2015). The purpose of principal component analysis is to find the best low-dimensional representation of the variation in a multivariate data set. We can carry out a principal component analysis to investigate whether we can capture most of the variation between samples using a smaller number of new variables (principal components), where each of these new variables is a linear combination of all or some of the traits.

PCA reduces the original variables into a new set of uncorrelated variables known as principal components (PCs). These PCs clarify the connections between traits and divide the total variance of the original traits into a small number of uncorrelated new variables (Wiley and Lieberman, 2011).

Since PCA extract all the important components and highlight their contribution towards the total variability, it can be the choice as an important tool to speed up the breeding program. The multivariate analysis that allows the development of efficient selection strategies allowed to group the different clones according to their characteristics. The objective of this study was therefore to use PCA to identify the principal traits of an effective phenotypic characterization of sugarcane in the identification of higher clones for the next stages of selection.

Materials and methods

Experimental site

The study was conducted at the Experimental Station of the Sugar Complex of Ferké 2 in northern Côte d'Ivoire between 9°20' and 9°60' north latitude on the one hand, and, 5°22' and 5°40' west longitude, with an average altitude 325 m above sea level. The climate prevailing in the study area is of dry tropical type with two seasons; one dry season, from November to April and the other wet, from May to October. The rainfall pattern is unimodal and centered on the months of August-September which accumulate nearly half of the average annual rainfall height of about 1200 mm. To compensate for the water deficit of sugarcane, the water supply through irrigation approaches on average 700 mm (Péné *et al.*, 2012). The vegetation of Ferké 2 is a Guinean savanna (or sub-Sudanese) of wooded type, with variable levels containing small fragments of detached forests. The soils are predominantly ferrallitic, with shallow topsoil (40 to 60 cm) limited by indurations

Plant material

New genotypes have been developed from bi-parental cross breeding of fuzz (true seed) of Reunion origin at Reunion Island sugarcane breeding center (eRcane). The large number of seedling clones was developed and year wise tested/screened under several selection stages in single plant, 1st stage, 2nd stage and 3rd stage. Their numbers were reduced at each stage and only promising clones were promoted to the next selection stage on the basis of better stalk diameter (mm), number of Stalks on 3 meters,% smut, Stalk height (m), % scald leaf, flowering rate% and% Brix. The one hundred and forty-eight (148) clones used in this study come from a population of 984 of 3rd stage of the selection scheme (Table 1). All genotypes were planted following families and compared between control commercial varieties SP70-1006.

Table 1. sugarcane genotypes tested in 3rd selection stage.

N°	Genotypes	N°	Genotypes	N°	Genotypes	N°	Genotypes	N°	Genotypes
1	RCI14/11	31	RCI14/131	61	RCI14/161	91	RCI12/191	121	RCI13/1121
2	RCI13/12	32	RCI14/132	62	RCI11/162	92	RCI12/192	122	RCI13/1122
3	RCI13/13	33	RCI10/133	63	RCI11/163	93	RCI13/193	123	RCI13/1123
4	RCI14/14	34	RCI11/134	64	RCI10/164	94	RCI13/194	124	RCI13/1124
5	RCI12/15	35	RCI11/135	65	RCI11/165	95	RCI13/195	125	RCI14/1125
6	RCI13/16	36	RCI13/136	66	RCI11/166	96	RCI13/196	126	RCI14/1126
7	RCI13/17	37	RCI13/137	67	RCI11/167	97	RCI13/197	127	RCI14/1127
8	RCI14/18	38	RCI13/138	68	RCI11/168	98	RCI13/198	128	RCI11/1128
9	RCI12/19	39	RCI13/139	69	RCI11/169	99	RCI14/199	129	RCI11/1129
10	RCI13/110	40	RCI13/140	70	RCI11/170	100	RCI14/1100	130	RCI12/1130
11	RCI14/111	41	RCI13/141	71	RCI14/171	101	RCI14/1101	131	RCI13/1131
12	RCI11/112	42	RCI13/142	72	RCI13/172	102	RCI14/1102	132	RCI13/1132
13	RCI11/113	43	RCI13/143	73	RCI13/173	103	RCI14/1103	133	RCI13/1133
14	RCI11/114	44	RCI13/144	74	RCI13/174	104	RCI14/1104	134	RCI13/1134
15	RCI11/115	45	RCI13/145	75	RCI13/175	105	RCI14/1105	135	RCI13/1135
16	RCI13/116	46	RCI14/146	76	RCI13/176	106	RCI14/1106	136	RCI13/1136
17	RCI13/117	47	RCI14/147	77	RCI13/177	107	RCI14/1107	137	RCI14/1137
18	RCI13/118	48	RCI14/148	78	RCI13/178	108	RCI14/1108	138	RCI14/1138
19	RCI13/119	49	RCI12/149	79	RCI13/179	109	RCI14/1109	139	RCI14/1139
20	RCI13/120	50	RCI13/150	80	RCI13/180	110	RCI11/1110	140	RCI14/1140
21	RCI13/121	51	RCI13/151	81	RCI13/181	111	RCI11/1111	141	RCI14/1141
22	RCI13/122	52	RCI13/152	82	RCI13/182	112	RCI11/1112	142	RCI14/1142
23	RCI13/123	53	RCI13/153	83	RCI13/183	113	RCI11/1113	143	RCI14/1143
24	RCI13/124	54	RCI14/154	84	RCI13/184	114	RCI11/1114	144	RCI14/1144

N°	Genotypes	N°	Genotypes	N°	Genotypes	N°	Genotypes	N°	Genotypes
25	RCI13/125	55	RCI14/155	85	RCI13/185	115	RCI12/1115	145	RCI14/1145
26	RCI13/126	56	RCI14/156	86	RCI13/186	116	RCI12/1116	146	RCI11/1146
27	RCI14/127	57	RCI14/157	87	RCI13/187	117	RCI13/1117	147	RCI14/1147
28	RCI14/128	58	RCI14/158	88	RCI14/188	118	RCI13/1118	148	RCI13/1148
29	RCI14/129	59	RCI14/159	89	RCI14/189	119	RCI13/1119		
30	RCI14/130	60	RCI14/160	90	RCI11/190	120	RCI13/1120		

Methodology and experimental Design

The experimental design used at one-row screening stage was a randomized block comprising 148 sugarcane clones, each being planted in single rows of 3 m long with 1.5 m of inter-row spacing. Clones were not replicated apart from the control variety which was replicated many times every 5 rows of clones subjected to visual screening. The clones were divided into 60 families (or crosses) and the control varieties were planted in November 2015 in 17 matched sub-blocks of 3 m over a total area of 5,000 m². To facilitate comparison of the clones with the control, this was repeated every 7.5 m (after 5 individuals). To avoid edge effects, the field trial was surrounded by a buffer zone 3 m wide and 30 m long planted with a commercial variety SP70-1006.

Statistical analysis

The data collected were analyzed using the JASP 0.8.6.0 version 2018 software. These data were the subject of descriptive analysis, principal components analysis PCA, which is an analysis that consists of transforming the variables related to each other into new variables uncorrelated from each other. Cluster analysis based on the Ward method using Euclidean distance (Kumar *et al.*, 2009) was performed using the XLSTAT version 2017 statistical software. The Euclidean distance has been determined and makes it possible to identify the parameters close to each other (Elmore and Richman, 2001).

Results and discussion

Variability of all traits studied

Statistical parameters such as mean, standard deviation, minimum, maximum, coefficient of variation (CV%), Kurtosis and trust level (95%) for different traits in this study are shown in Table 2. Among morphological traits such as Nb_stalk/3m with CV% = 22.6, had a higher variation while Stalk diameter (mm) and Stalk height (m) with CV% = 10.9, 10.8 respectively, had a minimum variation. Among a phonological trait the% Flowering with CV%=86.0, had a higher variation while the technological quality Brix% had CV%=8.7. As the variables were normally distributed, the shape of the distribution was characterized by the kurtosis. The kurtosis values were negative for all morphological traits and varied from −0.1, 0.38 to −0.3 for Nb_stalk/3m, stalk Diameter and stalk height, respectively and negative value for the phonological trait the% Flowering -1.2. The Kurtosis for the Brix% had a positive value of 1.2. Therefore, the shape of the distributions was classified as platykurtic given that it was relatively less peaky than the Gaussian distribution. In order to avoid possible bias due to the shape of the curves, the classification proposed by Costa *et al.*, (2002) was also performed. However, the two methods showed similar classification of the CV values for each response variable (data not shown). This result is in accordance with previous reports by Costa *et al.* (2002), Carvalho *et al.* (2003) and Oliveira *et al.* (2009), who concluded that, when the variable is normally distributed, both methods are equivalent.

Table 2. Statistical parameters of 148 genotypes.

Traits	N	Means	SD	CV%	Kurtosis	Mini	Maxi	Trust level (95%)
Nb_stalk/3m	148	58.5	13.2	22.6	-0.1	32	100	2.1
Stalk_Diam(mm)	148	24.9	2.7	10.9	-0.4	18.6	31.6	0.4
Stalk_Height(m)	148	2.7	0.3	10.8	-0.3	2.1	3.6	0.0
Brix%	148	20.5	1.8	8.7	1.2	14.2	24.4	29
% Flowering	148	34.5	29.7	86.0	-1.2	0	94.4	4.8

Principal Component Analysis

Principal component analysis was performed to assess the variability among 148 sugarcane genotypes, using quantitative traits to reveal the outlier genotypes. The primary purpose of PCA was to define the underlying structure in a data. As a data reduction or exploratory methods, these procedures were used to reduce the number of variables and to detect structural relationship between these variables. PCA is a technique for finding putative variables which gives interpretation for as much of the variables in a multivariate data as possible. PCA is a unique mathematical solution; it performs simple reduction of the data set to a few components, for plotting and clustering purposes, and can be used to assume that the most essential components have association with some other underlying variables (Acquaah, 2012). The remaining variance of other principal components did not have significant eigenvalues. First 2 principal components (PCs) have significant eigenvalues for all 5 quantitative traits compared; hence they were all included in the model. PC1 contributed maximum variance (31.9%) and eigenvalue was 1.597 in the data set followed by the PC2 (26.2%) with eigenvalue of 1.309 (Table 3).

Table 4. Component loadings for all quantitative traits.

Traits	PC1	PC2	Uniqueness
% Flowering	0.5	0.6	0.3
Brix%	-0.5	-0.4	0.6
Nb_stalk/3m	-0.6	0.5	0.3
stalk_Diam. (mm)	0.6	-0.5	0.2
stalk_height. (m)	0.4	0.5	0.5

Component loadings of PCs for quantitative traits

Loadings of PC1

Yamamoto *et al.*, (2014) in his work had defined the component loading as the correlation coefficients between the PC scores and the variables. Loading of first principal component (PC1) presented in the Table 4, which describes that Stalk diameter (mm) showed maximum positive loadings (0.6) followed by the Stalk height (m) (0.4). The number of stalk/3m showed minimum loadings (-0.6). From the results it can be inferred that Stalk diameter (mm) and Stalk height (m) have positive correlation among themselves while these parameters have negative correlation with stalk/3m. While in the Component loading of PC2% flowering has maximum loading (0.6) in the PC. The Brix% showed minimum loading (-0.4). The Brix% have a negative correlation with % flowering.

PC1 versus PC2 biplot for 5 quantitative traits of 148 sugarcane genotypes

The first 2 PCs (PC1 and PC2) generated 58.2 percent of the total variance (Table 3) among the 148 genotypes for 5 quantitative traits under study and is represented in the Fig. 1 and 2. Brix% falls on opposite axis with respect to%Flowering and stalk height (m) in the correlation diagram which means that these parameters have negative correlation. Nb_stalk/3m has negative correlation stalk_diam (mm).

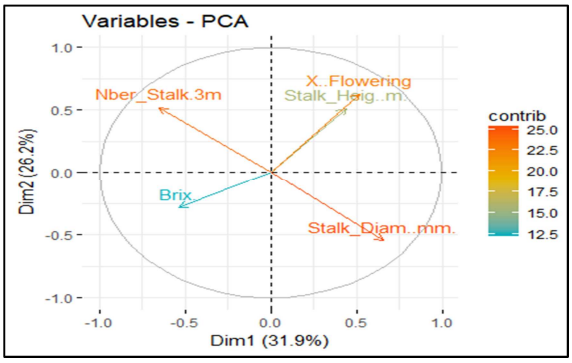

Fig. 1. Correlation circle resulting from projection of 148 sugarcane genotypes and quantitative traits.

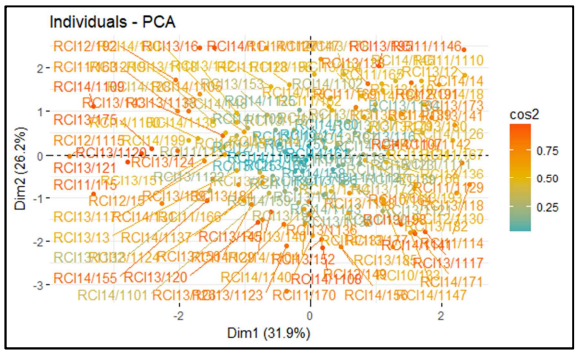

Fig. 2. Individuals factor map (PCA) from projection of 148 sugarcane genotypes.

The Scree plot

Scree plot helps us in deciding the number of components to retain (Lukibisi and Lanyasunya, 2010). The eigenvalues, associated with each component, are plotted and then searched for a breakdown between the components with large eigenvalues and those with smaller eigenvalues.

The components that come before the break are considered to be important and are retained, while the components that come after the break are taken to be unimportant and are therefore not retained. In Fig. 3, it can be seen easily the break after two components and it is approved that these two components are meaningful and important.

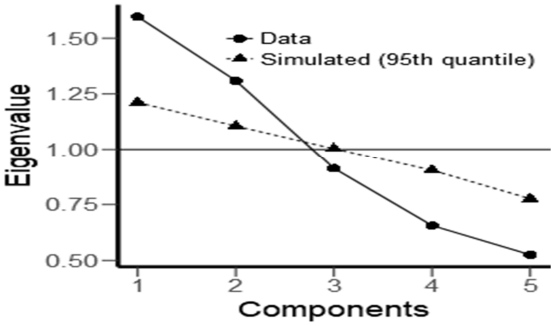

Fig. 3. Scree plot of components versus Eigenvalue.

Cluster analysis

For cluster analysis, hierarchal clustering with Ward's method was used. Based on the adjusted means of 5 sugarcanes traits results of the cluster analysis indicated that the 148 sugarcanes genotypes formed 5 distinct groups or clusters. Cluster III was the largest, containing 44 genotypes, followed by cluster I with 32 genotypes. Cluster IV contained 25 genotypes followed of cluster II with 23 genotypes and cluster V contained 20 genotypes (Fig. 4). No association between geographic origin of genotypes and cluster group was observed. Maximum genetic variability was present in all clusters. The means and CV% of different characteristics associated with each cluster are shown in Table 5. The genotypes appearing in cluster I is characterized with high Nb_stalk/3m, with high Brix %. The cluster II is characterized with high stalk number/3m, stalk height and early genotype.

Table 5. Means and CV% of 148 sugarcane genotypes for all quantitative traits.

	Nb_stalk/3m		stalk_Diam (mm)		stalk height (m)		Brix %		% Flowering	
	Mean	CV%	Mean	CV%	Mean	CV%	Mean	CV%	Mean	CV%
Cluster I	71.4	15.4	22.1	8.8	2.6	9.7	21.8	5.1	27.2	98.2
Cluster II	64.5	18.3	24.4	7.5	3.0	9.9	18.4	11.1	63.6	37.9
Cluster III	52.3	19.9	25.7	9.5	2.7	9.1	19.6	5.0	14.4	93.7
Cluster IV	49.5	17.0	26.4	5.3	2.8	8.8	22.4	3.3	26.4	101.1
Cluster V	50.0	11.4	25.9	9.7	2.7	7.6	20.1	4.4	63.1	24.9

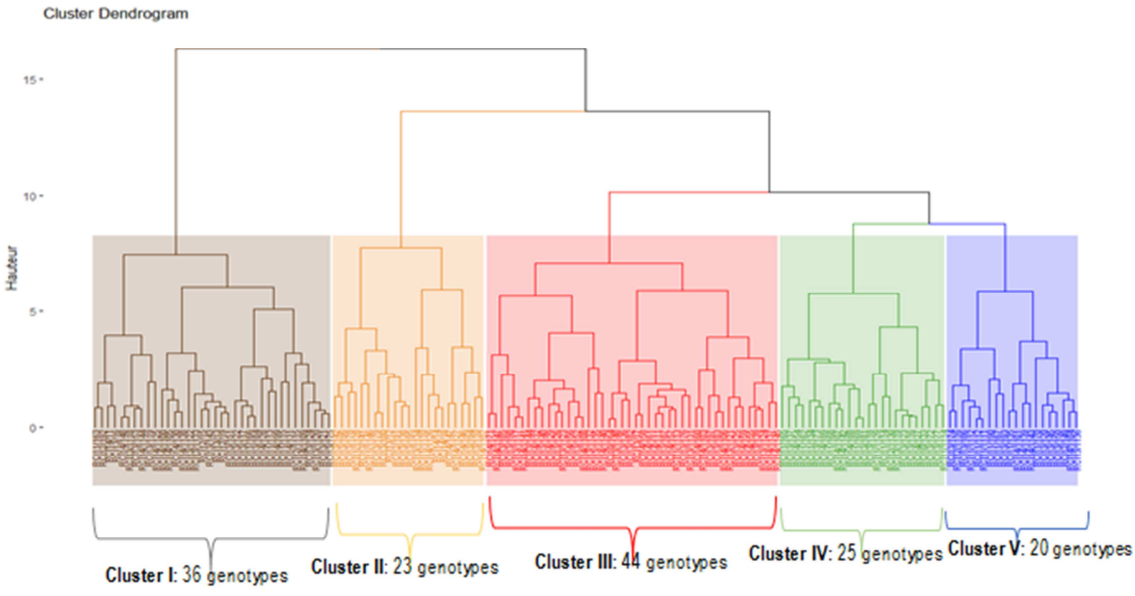

Fig. 4. Cluster diagram of 148 sugarcane genotypes on the based on Euclidean distance and 5 traits.

Genetic distance

The distances between the clusters based on the Euclidean distance statistics revealed that the cluster II had a large distance to the clusters III and V (respectively 60.4 and 64.5) whereas the group 3 had a large distance to the Cluster IV (76.0). In addition, Cluster IV was distant from Cluster V (72.2) (Table 6). The clones in each cluster represented a great diversity. In contrast, the smallest Euclidean distances were observed between cluster I and II, between cluster IV and II, and between cluster group IV and III, which shows a close diversity between clusters. In general, the smallest and largest distances among cluster groups suggest a high probability of getting to select promising clones.

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 148 sugarcane genotypes studied were clustered into 5 groups and were highly significantly different for Brix, Stalk height, Stalk diameter, nomber of stalk and % flowering.

Morphological diversity among the sugarcane clone germplasm was well defined by both principal component and clustering analyses. Considering the different morpho-bio-agronomic descriptors, it has been possible to observe a noteworthy inter and intra-group diversity. This recommends the likelihood of attaining, through selection, suitable genotypes combining the high yield with desirable traits. The selected genotypes were thought to have great promise to be good commercial varieties in future and were advanced for further testing and progression.

References

Acquaah G. 2012. Principles of plant genetics and breeding. John Wiley and Sons, Ltd, Publication **2,** 718-719.

Carvalho CGP, Arias CAA, Toledo JFF, Almeida L, Kiihl RA, Oliveira MF, Hiromoto DM, Takeda C. 2003. Proposal to categorize coefficients of variation for yield and plant height in soybean. Pesquisa Agropecuária Brasileira **38,** 187-193.

Chohan MUA, Talpur S, Junejo GS, Unar RNBPA. 2014. Selection and evaluation of the diverse sugarcane genotypes in 4th stage. The J. Anim. Plant Scis **24(1),** 197-203.

Costa NHAD, Seraphin JC, Zimmermann FJP. 2002. A new method of variation coefficient classification for upland rice crop. Pesquisa Agropecuária Brasileira **37,** 243-249.

Dahlquist E. 2013. Biomass as Energy Source: Resources, SyStalks and Application, CRC Press, Taylors & Francis, London.

Elmore KL, Richman MB. 2001. Euclidean distance as a similarity metric for principal component analysis. Mon. Wea. Rev **129 (3),** 540-549.

Glaz B, Miller DJD, Tai PYP, Deren CW, Kang MS, Lyrene PM, Gill BS. 2002. Sugarcane Genotype Repeatability in Replicated Selection Stages and Commercial Adoption. J. Am. Soc. Sugarcane. Tech **22,** 73-88.

Glaz B, Miller JD. 1982. Comparison of commercial and experimental yields in sugarcane. Proceed. III Inter-American Sugar Cane Seminar: Varieties and Breeding 139- 143p.

Kumar B, Gaibriyal ML, Ruchi Upadhyay A. 2009. Genetic variability, diversity and association of quantitative traits with grain yield in bread wheat (*Triticum aestivum* L.). Asian J. Agric. Sci **1(1),** 4-6.

Lukibisi FB, Lanyasunya T. 2010. Using principal component analysis to analyze mineral composition data. Biennial Kenya Agricultural Research Institute, Scientific Conference on Socio Economics and Biometrics.

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

Péné CB, Ouattara HM, Koulibaly GS. 2012. Late sugarcane performance as affected by soil water regime at the yield formation stage on commercial farms in northern Ivory Coast. J. of life Sci. Vol **6 (6),** 644-651.

Rukundo P, Hussein S, Mark L, Daphrose G. 2015. Application of principal component analysis to yield and yield related traits to identify sweet potato breeding parents. Tropical Agric. (Trinidad) **92 (1).**

Welde MG, Alamerew S, Kufa T, Benti T. 2013. Genetic Diversity Analysis of Some Ethopian Specialty Coffee (*Coffea Arabica* L.) Germplasm Accessions Based on Morphological Traits. Time Journals of Agriculture and Veterinary Sciences **1, 4,** pp. 47-54.

Wiley E, Lieberman B. 2011. Phylogenetics: Theory and Practice of Phylogenetic SyStalkatics. John Wiley and Sons, New York.

Yamamoto H, Fujimori T, Sato H, Ishikawa G, Kami K, Ohashi Y. 2014. Statistical hypothesis testing of factor loading in principal component analysis and its application to metabolite set enrichment analysis. BMC Bioinformatics **15 (1),** 1-9.