Principal component analysis for seed yield traits in roselle (*Hibiscus sabdariffa* L.)

N. Hari Satyanarayana*, S. Mukherjee, V. Visalakshmi[1], Bhanu priya and K.K. Sarkar
Dept. of Genetics and Plant Breeding, BCKV, Mohanpur, Nadia, West Bengal – 741 252
[1]Agricultural Research Station (ANGRAU), Ragolu, Srikakulam District, A. P. – 532 484
E-mail: namburihari@ymail.com

Abstract
Sixty genotypes of roselle (*Hibiscus sabdariffa* L.) were evaluated at three different agro-climatic zones *viz.*, North Coastal Zone of Andhra Pradesh, Indo-Gangetic Zone and Terai Zone of West Bengal for seed and its contributing characters during *kharif*, 2013 and 2014. Principal Component Analysis (PCA) indicated that three components with eigen values more than one accounted for about 79 % of the total variation among ten quantitative characters responsible for seed yield in roselle genotypes. The principal components PC_1, PC_2 and PC_3 contributed about 48.35 %, 21.22 % and 9.57 %, respectively to the total variation. The first principal component had high positive loading for test weight, plant height @130 DAS, seeds per pod, plant height @75 DAS, pods per plant and days to 50% flowering which contributed more to the diversity. The second principal component had high loading for pods per plant, seed yield/plant, test weight, plant height @130 DAS and @75 DAS. The genotypes *viz.*, AR-71, AR-55, REX-34, AS-80-26, AS-81-1, AS-80-19, ER-60, AS-81-22, AS-81-2 and REX-38 were identified as the most diverse genotypes, based on component scores and these genotypes could be utilized as parents in hybridization programme.

Introduction
Roselle (H*ibiscus sabdariffa* L.) is an annual or perennial plant belonging to the large family Malvaceae and is cultivated in tropical and sub-tropical regions for bast fibre, paper pulp or edible calyces, leaves and seeds (Osman *et al*., 2011). In India, roselle is one of the most important bast fibre crop which occupies second place in area and production after jute (Hariram and Appalaswamy, 2014). The major growing states of this crop are Andhra Pradesh, Bihar, Orissa, West Bengal and Maharastra with an area of 84 thousand ha (during 2012-13) that produced 6.16 lakh bales (1bale = 180 kg) with an average productivity of about 13.20 q/ha. (Sen and Karmakar, 2014). Till now ten varieties for bast fibre have been released in India; six from Agricultural Research Station, Amadalavalasa (ANGRAU) and four from CRIJAF (ICAR), Barrackpore, Kolkata (Kar *et al*., 2013). Seed yield is quantitatively inherited and influenced by genetic factors as well as environments.

Presence of sufficient genetic variability is a pre-requisite to formulate breeding programme aimed at improvement in yield and other characters, since the crosses made between the parents with maximum genetic divergence would more likely to yield desirable recombinants in the progeny. However, it is desirable to select suitable genetically divergent parents, based on information about the genetic variability and genetic diversity present in the available germplasm. Further, selection of diverse parents for hybridization programme will be effective by the identification of characters responsible for the total genetic diversity among the populations (Murthy and Arunachalam, 1966). Adequate analysis of genetic diversity is essential for understanding and utilization of genetic variability among accessions and their characters.

Principal component analysis helps researchers to distinguish significant relationship between traits. This is a multivariate analysis method that aims to explain the correlation between a large set of variables in terms of a small number of underlying independent factors (Badenes *et al.,* 2000). In determining the potentiality of genetically different lines and cultivars, breeders have to observe many different characters that influence yield (Pradhan *et al.,* 2015). Using genetic diversity vast improvement is possible in roselle crop, however, due to narrow genetic base in roselle, lot of potential yet to be explore. Hence, the present investigation was undertaken with the main objective to assess the potential genetic diversity among sixty genotypes of roselle through principal component analysis based on ten seed yield attributing traits for selection of parents in hybridization programme to obtain desirable segregants in advanced generation.

Materials and methods
In the present investigation, sixty roselle (*Hibiscus sabdariffa* L.) genotypes (Table 1) consisting of eleven exotic lines, four released varieties and 45 indigenous accessions (Table 2) were evaluated at three different agro-climatic environments *viz.*, North Coastal Zone, Andhra Pradesh; Indo-

Gangetic and Terai Zones of West Bengal. During the first year, in *kharif* 2013, trails were laid at Agricultural Research Station, Ragolu (Latitude 18^0 24$'$ N; Longitude 83. 84^0 E at an altitude of 27m above mean sea level); Instructional Farm, Bidhan Chandra Krishi Vishwavidyalaya (BCKV), Jaguli (Latitude 22^0 93$'$ N; Longitude 88. 59^0 E at an altitude of 9.75m above mean sea level) and University Farm, Uttar Banga Krishi Vishwavidyalaya (UBKV), Pundibari (Latitude 26^0 19$'$ N; Longitude 89. 23^0 E at an altitude of 43m above mean sea level) in the three agro-climatic zones, respectively; whereas, in the second year, in *kharif* 2014, trails were laid out at Agricultural Research Station, Ragolu;   Teaching farm, Mondauri, BCKV (Latitude 22^0 87$'$ N; Longitude 88. 59^0 E at an altitude of 9.75m above mean sea level) and University Farm, UBKV, Pundibari in the three agro-climatic zones, respectively.

The experimental trial was laid out in randomized block design with a plot size of four rows of 2m length in two replications and adapted a spacing of 30 x 10cm under rainfed conditions and followed recommended package of practices to raise a good crop. Data on the basis of five randomly selected competitive plants were recorded on plant height @75 DAS (cm), base diameter @75 DAS (mm), plant height @130 DAS (cm), base diameter @130 DAS (mm), days to 50% flowering, pods per plant, seeds per pod, test weight (g) and seed yield/plant (g). The principal component analysis using INDOSTAT software was followed in the extraction of the components.

Results and discussion
The analysis of variance revealed significant differences among sixty genotypes for all the ten characters. This indicated the existence of significant amount of variability among the sixty genotypes for the characters studied. Three principal components were identified through Principal Component Analysis (PCA) with eigen values more than one having contributed about 79 per cent of total variance (Table3).

The first principal component (PC_1) contributed maximum towards variability (48.35%) with significant loading of test weight (-0.620) which was negatively correlated and plant height @130 DAS (0.473), seeds per pod (0.347), plant height @75 DAS (0.342), pods per plant (0.272) and days to 50 per cent flowering (0.227) which were positively correlated. The second principal component (PC_2) accounted for 21.22 percent of total variance and it reflected significant loading of pods per plant (0.539), seed yield/plant (0.493), test weight (0.336), plant height @130 DAS (0.293) and plant height @75 DAS (0.273) which were positively correlated and seeds per pod (-0.400) and days to 50 per cent flowering (-0.164) which were negatively correlated. The third

principal component (PC_3) accounted for 9.57 percent of cumulative variance and it was characterized by significant loading of seeds per pod (-0.684), seed yield/plant (-0.369), test weight (-0.330), days to 50% flowering (-0.326) and pods per plant (-0.261) which were negatively correlated and plant height @75 DAS (0.239) and plant height @130 DAS (0.172) which were positively correlated.

The characters plant height @130 DAS (0.473), seeds per pod (0.347) and plant height @75 DAS (0.342) contributed maximum to the diversity in PC_1. To the diversity in the PC_2, pods per plant (0.539), seed yield/plant (0.493), test weight (0.336), plant height @130 DAS (0.293) and plant height @75 DAS (0.273) contributed maximum to the diversity. In PC_3, plant height @75 DAS (0.239) and plant height @130 DAS (0.172) contributed maximum to the diversity. Sharma and Prasad (2010) evaluated twenty genotypes of *okra* and observed three principal components with eigen value more than one, contributed for a cumulative variation of 62.83 per cent. Kandil *et al.* (2012) studied twenty one genotypes of *flax* and twenty F_1 hybrids and reported that first five principal components accounted for 78.2 per cent of total variation. Similarly, Ayo-Vaughan *et al.* (2013) reported three principal components contributing to 62.70 per cent of total variation in *kenaf* by studying twenty genotypes. Likewise, Ghosh *et al.* (2013) reported five principal components contributing to 75.93 percent of total variation in 63 accessions of *jute* including 39 white jute and 24 tossa jute accessions.

Cluster analysis: Five group patterns were observed in three dimensional (3D) PCA score plot by the first two principal components which accounted for 69.58 per cent of total variation (Fig 1). Hence, based on PCA analysis, sixty genotypes of roselle were grouped into five clusters based on ten quantitative characters. Out of five clusters, three clusters were having two genotypes each, one cluster with four genotypes and the last cluster with fifty genotypes which were closely related with each other within the cluster and distinctly related with genotypes of other clusters. High degree of variability was existing within the accessions and hence these characters will help for further improvement of the roselle crop by evolving high seed yielding cultivars. Hence, ten genotypes belonging to four diverse clusters *viz.*, AR-71, AR-55, REX-34, AS-80-26, AS-81-1, AS-80-19, ER-60, AS-81-22, AS-81-2 and REX-38 may be selected as parents for future breeding programmes. Crosses between the genotypes of the diverse clusters will help to produce high fibre yielding segregants for the crop improvement.

The Principal Component Analysis ultimately showed the amount of variability present in the

genotypes which could be used for improvement of roselle crop. Principal component analysis in germplasm collection facilitates reliable classification of genotypes, identification of diverse parents for specific breeding purpose. Under the present investigation, the genotypes viz., AR-71, AR-55, REX-34, AS-80-26, AS-81-1, AS-80-19, ER-60, AS-81-22, AS-81-2 and REX-38 were identified as the most diverse genotypes, based on component scores and they could be utilized as parents in hybridization programme for roselle crop improvement.

References
Ayo-Vaughan, M.A., Alake, C.O., Ariya, O.J. and Adeniyan, O.N. 2013. Classification of variation patterns in kenaf. *African Crop Sci. Conf. Proc.*, **11**: 559-562.

Badenes, M.L., Martinez-Calvo, J. and Llacer, G. 2000. Analysis of a germplasm collection of loquat (*Eriobotrya japonica* Lindl.). *Euphytica*, **114**: 187-194.

Ghosh, R.K., Sreewongchai, T., Nakasathien, S. and Phumichai, C. 2013. Phenotypic variation and the relationship among jute (*Corchorus* sps) genotypes using morpho-agronomic traits and multivariate analysis. *Australian J. Crop Sci.*, **7**(6): 830-842.

Hariram Kumar, B. and Appalaswamy, A. 2014. Principal Component and Cluster Analysis of fibre yield in roselle (*Hibiscus sabdariffa* L.). *Int. J. Agric, Environ. Biotech.*, **7**(2): 255-260.

Kandil, A.A., Shareif, A.E., Abo-Zaied, T.A. and Moussa, A.G.T. 2012. Multivariate analysis of some economic characters in flax. *Pakistan J. Bio. Sci.*, **15**(2): 85-91.

Kar, C.S., Mitra, J., Satya, P., Pandey, S.K., Sharma, H.K., Chaudhury, S.B. and Anil Kumar. 2013. Status and Prospectus of Improvement of Jute and Allied Fibres. Training Manual on Advances in Improved Production Technology and fibre quality assessment of Jute and Allied Fibre crops held at CRIJAF, Barrackpore, Kolkata from 24th to 29th June, 2013, pp. 11-22.

Murthy, B. and Arunachalam, V. 1966. The nature of genetic divergence in relation to breeding systems in crop plants. *Indian J. Genet. Pl. Breed.*, **26**: 188-189.

Osman, M., Faruq, G., Saberi, S., Majid, N.A., Nagoor, N.H. and Zulqarnain, M. 2011. Morpho-agronomic analysis of three roselle (*Hibiscus sabdariffa* L.) mutants in tropical Malaysia. *Australian J. Crop Sci.*, **5**(10): 1150-1156.

Pradhan, A.M., Sarkar, K.K. and Konar, A. 2015. Analysis of genetic diversity in Indian potato genotypes by using Principal Component Analysis. *Elect. J. Plant Breed.*, **6**(3): 782-786.

Sen, H.S. and Karmakar, P.G. 2014. Public-Private Partnership from the perspectives of CRIJAF for development of Jut and Allied Fibres. Today's natural fibres for tomorrow. *In Proc.* International Conference on Natural Fibres organized by The Indian Natural Fibre Society, Kolkata pp. 40-44.

Sharma, R.K. and Prasad, K. 2010. Classification of promising okra (*Abelmoschus esculentus*) genotypes based on Principal Component Analysis. *J. Tropical Agric. Food Sci.*, **38**(2): 161-169.

Table 1. List of sixty roselle (*Hibiscus sabdariffa* L.) genotypes

S. No.	Name of the genotype	S. No.	Name of the genotype
1	AR - 14	31	R - 318
2	AR - 19	32	R - 322
3	AR - 42	33	ER - 56
4	AR - 45	34	ER - 57
5	AR - 48	35	ER - 60
6	AR - 50	36	ER - 68
7	AR - 55	37	AS - 80 - 6
8	AR - 66	38	AS - 80 - 7
9	AR - 67	39	AS - 80 - 19
10	AR - 71	40	AS - 80 - 26
11	AR - 79	41	AS - 80 - 29
12	AR - 80	42	AS - 81 - 1
13	AR - 81	43	AS - 81 - 2
14	AR - 85	44	AS - 81 - 3
15	AR - 88	45	AS - 81 - 5
16	R - 16	46	AS - 81 - 9
17	R - 29	47	AS - 81 - 14
18	R - 30	48	AS - 81 - 17
19	R - 37	49	AS - 81 - 22
20	R - 48	50	REX - 6
21	R - 67	51	REX - 14
22	R - 68	52	REX - 34
23	R - 77	53	REX - 38
24	R - 86	54	REX - 45
25	R - 134	55	REX - 52
26	R - 180	56	REX - 63
27	R - 225	57	HS – 4288
28	R - 243	58	AMV - 4
29	R - 271	59	AMV - 5
30	R - 284	60	AMV – 7

Table 2. Grouping of sixty roselle (*Hibiscus sabdariffa* L.) genotypes

Group	Name of the genotypes
Indigenous (45)	AR-14, AR-19, AR-42, AR-45, AR-48, AR-50, AR-55, AR-66, AR-67, AR-71, AR-79, AR-80, AR-81, AR-85, AR-88, R-16, R-29, R-30, R-37, R-48, R-67, R-68, R-77, R-86, R-134, R-180, R-225, R-243, R-271, R-284, R-318, R-322, AS-80-6, AS-80-7, AS-80-A9, AS-80-26, AS-80-29, AS-81-1, AS-81-2, AS-81-3, AS-81-5, AS-81-9, AS-81-14, AS-81-17 and AS-81-22
Exotic (11)	ER-56, ER-57, ER-60, ER-68, REX-6, REX-14, REX-34, REX-38, REX-45, REX-52 and REX-63
Varieties (4)	HS-4288, AMV-4, AMV-5 and AMV-7

Table 3. Character loading of three principal components for sixty genotypes of roselle

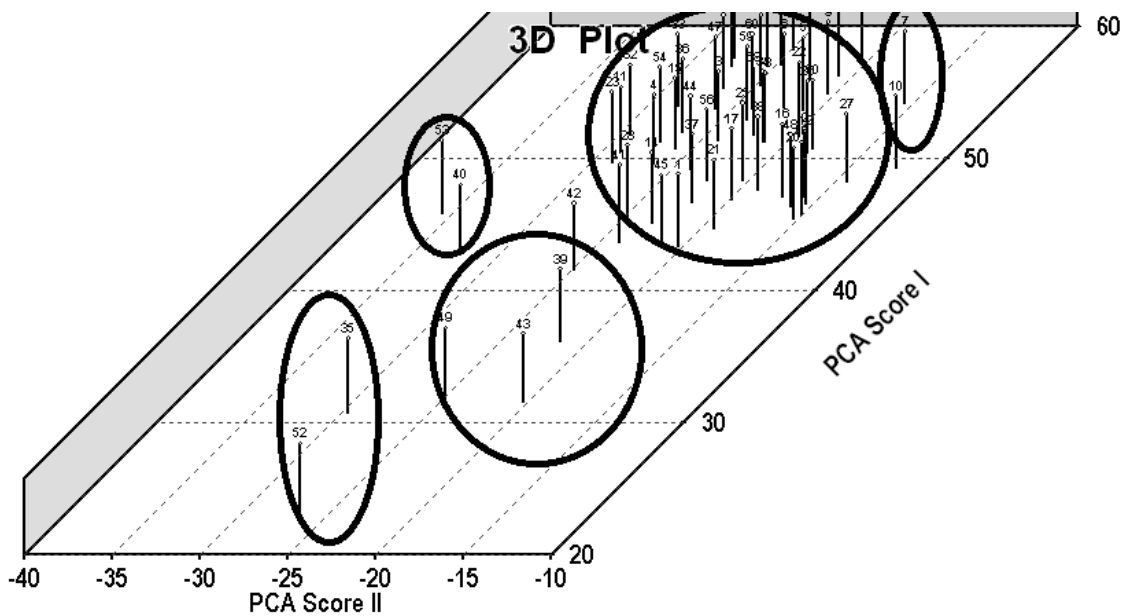| Characters	PC $_1$	PC $_2$	PC $_3$
Plant height @ 75 DAS (cm)	0.342	0.273	0.239
Base diameter @ 75 DAS (mm)	-0.083	0.038	-0.034
Plant height @ 130 DAS (cm)	0.473	0.293	0.172
Base diameter @ 130 DAS (mm)	-0.103	0.070	-0.138
No. of Nodes / plant	-0.104	-0.028	-0.078
Days to 50% flowering	0.227	-0.164	-0.326
No. of pods / plant	0.272	0.539	-0.261
No. of seeds / pod	0.347	-0.400	-0.684
Test weight (g)	-0.620	0.336	-0.330
Seed yield / plant (g)	0.028	0.493	-0.369
Eigen value (Root)	54.31	35.98	24.16
% total variance	48.35	21.22	9.57
Cumulative variance	48.35	69.58	79.15

Fig. 1. Three (3D) dimensional PCA score plot for sixty genotypes of roselle for ten seed yielding traits