International Journal of Chemical Studies

P-ISSN: 2349–8528
E-ISSN: 2321–4902
IJCS 2017; 5(4): 989-993

Received: 25-05-2017
Accepted: 26-06-2017

Vivek Kumar
Department of Vegetable Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar (Uttarakhand), India

Dhirendra Kumar Singh
Department of Vegetable Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar (Uttarakhand), India

Ankit Panchbhaiya
Department of Vegetable Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar (Uttarakhand), India

Neeraj Singh
Department of Vegetable Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar (Uttarakhand), India

Correspondence
Vivek Kumar
Department of Vegetable Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar (Uttarakhand), India

Genetic divergence studies in mid-season cauliflower (*Brassica oleracea* var. *botrytis* L.) through Principal Component Analysis (PCA) and D^2 analysis

Vivek Kumar, Dhirendra Kumar Singh, Ankit Panchbhaiya and Neeraj Singh

Abstract
The present study was undertaken at Vegetable Research Centre, G.B. Pant University of Agriculture and Technology, Pantnagar, Udham Singh Nagar (Uttarakhand) during September- December, 2015. Fifty seven genotypes of cauliflower were grown in Augmented Block Design II including five checks which were arranged in four blocks with eighteen genotypes in each block inclusive of five checks. The study was primarily focused on assessing Principal Component Analysis (PCA) and D^2 analysis. Genetic diversity plays an important role in plant breeding because hybrids between lines of diverse origin generally display a great heterosis than those between closely related strains. Geographical diversity may not be the only factor responsible for causing genetic diversification, thus focus should be laid on selecting the parents based on genetic diversity rather on geographical diversity. The Eigen root of first principal component was accounted approximately 45.696% of total variation. The first five PC axes explained 81.72% of the variation, suggesting considerable diversity among the genotypes for all the characters. Based on cluster analysis, the results revealed maximum inter-cluster distance between cluster IV and VII followed by cluster IV and VIII, thereby paving the chance for them to be used in hybridization breeding programmes.

Keywords: Principal component analysis, D^2 analysis, Tocher method, Diversity, Cauliflower

Introduction
Cauliflower (*Brassica oleracea* var. *botrytis* L., 2n = 2x = 18) is an important vegetable crop cultivated throughout India. Cauliflower is the most popular vegetable among the Cole crops. It is believed that it has been originated in the island of Cyprus from where it moved to other areas like Syria, Turkey, Egypt, Italy, Spain and North Western Europe. It is originated from wild cabbage known as 'Cole warts', through mutation, human selection and adoption. Dr. Jemson at Saharanpur introduced it to India in 1822 during the period of East India Company (Swarup and Chatterjee, 1972) [11]. The present tropical Indian cauliflower developed as a result of inter crossing between European and Cornish type, perhaps the first to introduced in India, has itself gone out of cultivation after contributing many genes to Indian varieties like resistance to black rot, self-incompatibility, curd flavour, open plant habit, exposed yellow loose curds etc (Swarup and Chatterjee, 1972) [11]. The highly suppressed pre-floral apical meristem commonly known as 'curd' is the edible part **(**Sidki, 1962**)** [10]. For a good cauliflower crop, high yield, compact, white colour and medium-sized curds, free from any disease or disorder, are desired (Varalaxmi, 2009) [12]. Yield in cauliflower is a complex character influenced by various component characters namely, marketable curd weight, plant height, number of leaves per plant, curd diameter and curd depth which inherit polygenically and highly subject to environmental variations (Devaraju *et al*., 2010) [1].
Principal component analysis being a data reduction technique for investigating the interdependence attempts to simplify complex and diverse relationships existing among a set of observed variables, by revealing common dimensions or components that link seemingly unrelated variables. Genetic diversity plays an important role in plant breeding because hybrids between lines of diverse origin generally display a great heterosis than those between closely related strains. The primary objective behind a plant breeding programme is to create variability and select best recombinants with desirable characters.

But selection is effective only when a tremendous diversity exists in a population. Among different ways of creating variability, hybridization is one in which genetically diverse parents with desirable characters have been utilized to develop better genotypes. Therefore an attempt was made to study the genetic diversity in mid-season cauliflower.

Materials and methods

The experiment was conducted at Vegetable Research Centre (V.R.C.) of G.B. Pant University of Agriculture & Technology, Pantnagar during the year 2015-2016. The experimental material comprises of fifty seven genotypes of mid season cauliflower with five checks. The experiment was laid out in Augmented Block Design with four blocks, each block comprises of eighteen genotypes, also inclusive of five check genotypes planted at a row to row and plant to plant spacing of 60 cm by 50 cm. Observations were recorded for sixteen quantitative characters *viz*., leaf length (cm), leaf width (cm), Petiole length (cm), plant height (cm), plant spread (cm), number of leaves per plant, stalk length (cm), days to curd maturity, gross plant weight (g), marketable curd weight (g), net curd weight (g), curd length (cm), curd breadth (cm), harvest index (%), curd sixe index (cm²) and curd yield per hectare (q/ha).

The concept of principal component analysis, which is a multivariate technique, was developed by Hotelling (1933) [4] after its original concept was given by Pearson (1901) [7]. The D^2 analysis is a measure of the distance between a point P and a distribution D, introduced by P.C. Mahalanobis (1936) [6]. This distance is zero if P is at the mean of D and grows as P moves away from the mean.

Results and discussion

The percent variations explained by each Eigen roots are presented in Table 1 and figure 1. The Eigen root of first principal component was accounted approximately 45.696 % of total variation followed by second to fifth components which accounted 15.592, 8.477, 6.291 and 5.667 % of total variation presented among the genotypes, respectively. The first five PC axes explained 81.72% of the variation, suggesting considerable diversity among the genotypes for all the characters, the rest of the components not considered. These were interpreted as relative weight of the variables in each component. The important variables are those which have high positive or negative relative weight values.

Table 1: Eigen vector, Eigen root and associated variation for principal component in cauliflower based on economic traits

S.N.	Characters	PC 1	PC 2	PC 3	PC 4	PC 5
1	Leaf length (cm)	0.308	0.074	0.213	0.300	0.140
2	Leaf width (cm)	0.260	0.113	0.232	0.301	0.032
3	Petiole length (cm)	0.221	-0.004	0.188	0.115	0.557
4	Plant height (cm)	0.296	-0.089	0.238	-0.076	-0.038
5	Plant spread (cm)	0.283	0.176	0.067	0.064	-0.212
6	No. of leaves per plant	0.197	-0.267	0.111	-0.090	-0.489
7	Stalk length (cm)	0.048	0.367	-0.033	0.537	0.078
8	Days to maturity	0.117	-0.262	0.024	-0.404	0.585
9	Gross plant weight (g)	0.351	-0.049	-0.035	-0.102	-0.0384
10	Marketable curd wt. (g)	0.343	-0.121	-0.136	-0.040	-0.069
11	Net curd weight (g)	0.293	-0.121	-0.435	0.130	0.026
12	Curd length (cm)	0.011	0.567	-0.041	-0.317	0.040
13	Curd breadth (cm)	0.277	0.072	-0.371	-0.119	-0.024
14	Harvest index	-0.159	-0.181	-0.597	0.324	0.147
15	Curd size index (cm²)	0.142	0.512	-0.256	-0.295	0.014
16	Curd yield (q/ha.)	0.343	-0.125	-0.142	-0.032	-0.064
	Eigen Root	7.311	2.495	1.356	1.007	0.907
	Percent variation	45.696	15.592	8.477	6.291	5.667

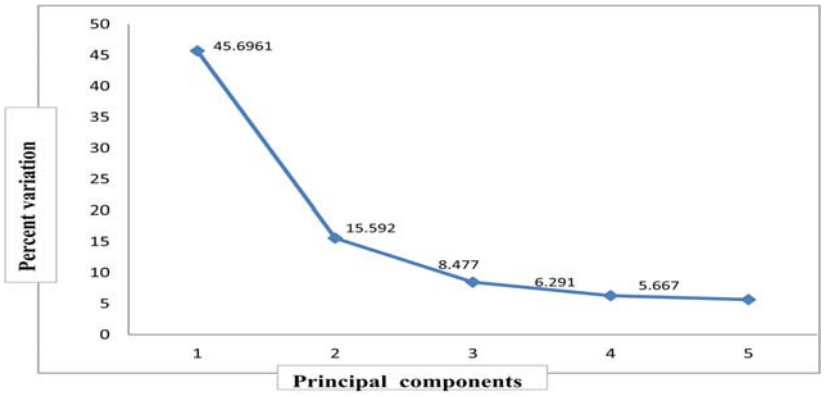

Fig 1: Cattel scree graph for variation explained by variation principal components based on yield attributing traits in Cauliflower genotypes

The principal component score based on the correlation of 16 quantitative characters of 57 genotypes of cauliflower along with the percent variation explained by each eigen roots are also presented in Table 1. The first principal component had high positive gross plant weight (0.351) followed by marketable curd weight (0.343) and curd yield per hectare (0.343), while high negative harvest index (-0.159). The second principal component had high positive curd length (0.567) followed by curd size index (0.511) and stalk length (0.367), while high negative number of leaves per plant (-

0.267) followed by days to maturity (-0.262) and harvest index (-0.181). The third principal component exhibited high positive plant height (0.238) followed by leaf width (0.232) and leaf length (0.213), while high negative harvest index (-0.597) followed by net curd weight (-0.435) and curd breadth (-0.371). The fourth principal component exhibited high positive stalk length (0.537) followed by harvest index (0.324) and leaf width (0.301), while high negative days to maturity (-0.404) followed by curd length (-0.317) and curd size index (-0.295). The fifth principal component exhibited high positive days to maturity (0.585) followed by petiole length (0.557) and harvest index (0.147), while high negative number of leaves per plant (-0.489) followed by plant spread (-0.212) and marketable curd weight (-0.069).

Principal component analysis being a data reduction technique for investigating the interdependence attempts to simplify complex and diverse relationships existing among a set of observed variables, by revealing common dimensions or components that link seemingly unrelated variables. The results so obtained revealed that the important characters responsible for genetic divergence in vector 1 were curd weight, curd length, curd yield and days to curd initiation in conformation with the present results as reported by Quamruzzaman *et al*. (2007) [8]. Similarly identical to the results Verma *et al.* (2016) [13] suggested leaf length, leaf width and days to maturity as higher loading displaying variables.

Cluster analysis by Tocher method based on principal component analysis was found to be more useful measure for estimating divergence. Genetic divergence was studied in fifty seven genotypes in respect of various economic traits using cluster analysis by Tocher method. Distribution of genotypes in each cluster is presented in Table 2. Cluster number I had highest number of genotypes (17) followed by cluster VI (14), cluster V (10), cluster III (9), cluster VIII (4), cluster II and cluster VII (1 each).

Table 2: Distributing pattern of fifty seven genotypes of cauliflower into eight clusters

Cluster number	Number of genotypes	Genotype included
I	17	PCF108, PCF 244, PCF 251, DC 541-5, PC 77, INB-9-2, PCF 70, PCF 250, PCF-93, PCF 246, PCF 252, CAUVAR-5, PCF 248, INBPCF 120, Pusa Early Synthetic, Pusa Sharad, 2013/CAUMVAR-6
II	1	PCF 249
III	9	DC-98-4-2, Pant Gobi-4, INB-20, INBPCF-3, INB 16-2, PCF-87, C1, PCF 245, PCF 256
IV	1	Pusa Pausjha
V	10	PG-3-1-1, PCF 4-2, PG-6, PCPGR-2004, INB 10-2, Comp-3, PCF-65, PG-5, PCF 253, PCF 243
VI	14	INBPCF 117, INB 20-3, PCF- 117, PC 100, PC 20-8, PCF 14-2, PCF 1-1, Pant Gobi-3, DC-94, PC 98-4-3, PCF-86, PCF 255, PCF 247, PCF 254
VII	1	PCF-7
VIII	4	CAUVAR-2, INB 79, PC-98, PCF 29

The genotypes were grouped into 8 clusters and the averages inter and intra-cluster distances have been presented in Table 3 and figure 2. Maximum intra-cluster distance was noted in cluster VIII (22.290) followed by cluster V (21.302), cluster I (15.324), cluster VI (10.087), cluster III (9.185), while minimum intra-cluster distance was recorded in cluster II, cluster IV and cluster VII (0.000).

Table 3: Inter and intra-cluster distances

Cluster	I	II	III	IV	V	VI	VII	VIII
I	15.324	37.346	20.04	73.295	31.188	38.440	104.504	48.449
II		0.000	46.964	107.522	51.483	61.818	132.242	65.578
III			9.185	102.243	26.493	17.670	55.921	26.218
IV				0.000	125.519	147.527	240.078	168.073
V					21.302	32.348	88.273	33.519
VI						10.087	33.953	22.234
VII							0.000	47.217
VIII								22.290

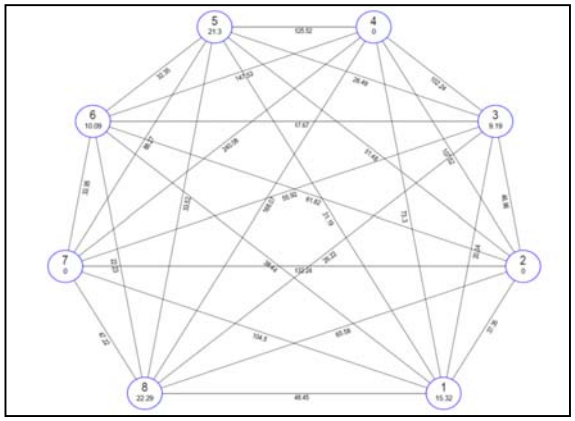

Fig 2: Average distance of intra and inter-cluster centroids based on various traits in cauliflower genotypes

Maximum inter-cluster distance was noted between cluster IV and VII (240.078) followed by cluster IV and VIII (168.073), cluster IV and VI (147.527), cluster VII and II (132.242), cluster IV and V (125.519), cluster IV and II (107.522), cluster VII and I (104.504), cluster IV and III (102.243), cluster VII and V (88.273), cluster IV and I (73.295), cluster VIII and II (65.578), cluster VI and II (61.818), cluster VII and III (55.921), cluster V and II (51.483), cluster VIII and I (48.449), cluster VIII and VII (47.217), cluster III and II (46.964), cluster VI and I (38.440), cluster II and I (37.346), cluster VII and VI (33.953), cluster VIII and V (33.519), cluster VI and V (32.348), cluster V and I (31.188), cluster V and III (26.493), cluster VIII and III (26.218), cluster VIII and VI (22.234), cluster III and I (20.040), cluster VI and III (17.670).

The cluster means for all the sixteen characters are given in Table 4. Leaf length had shown highest cluster mean value in cluster II (64.940) and lowest in cluster VII (27.540). Highest cluster mean value for leaf width was observed for cluster II (28.500) and minimum for cluster VII (13.00). Cluster IV (13.900) showed maximum cluster mean value for petiole length while cluster VII (6.670) the least. Plant height showed highest cluster mean value in cluster IV (60.850) and minimum for cluster VII (33.715). The character plant spread had the highest cluster mean value in cluster IV (65.175) and the lowest cluster mean value in cluster VII (41.370). For number of leaves per plant cluster IV (21.000) had the highest cluster mean value and cluster VIII (14.540) the least. Stalk length was noted to have maximum cluster mean value in cluster VIII (6.940) while minimum cluster mean value in cluster VII (3.580). Days to maturity had shown highest cluster mean value in cluster IV (123.750) while lowest cluster mean value in cluster VIII (89.100). Maximum cluster mean value for gross plant weight was observed for cluster IV (2101.750) and minimum in cluster VII (271.950). Cluster IV (905.500) showed highest cluster mean value for marketable curd weight while cluster VII (213.900) possessed the least cluster mean value. For net curd weight cluster IV (514.500) acquired maximum cluster mean value and cluster VII (144.900) the minimum. Cluster mean value for curd length was observed to be highest for cluster V (9.692) and lowest for cluster VII (5.560). The character curd breadth had the highest cluster mean value in cluster IV (14.175) and minimum for cluster VII (9.600). Cluster VII (56.780) showed the highest cluster mean value for harvest index whereas cluster IV (24.500) possessed the minimum cluster mean value. Cluster mean value for curd size index was found to be highest for cluster V (113.951) and minimum cluster mean value for cluster VII (54.445). Similarly highest cluster mean value for curd yield per hectare was observed in cluster IV (271.500) while minimum cluster mean value in cluster VII (63.950).

Table 4: Cluster mean for different economic traits in Cauliflower

S.No.	Characters	Cluster							
		I	II	III	IV	V	VI	VII	VIII
1	Leaf length (cm)	46.278	64.940	41.398	49.050	42.456	36.675	27.540	37.490
2	Leaf width (cm)	20.247	28.500	17.671	21.275	18.988	16.922	13.000	16.600
3	Petiole length (cm)	12.939	13.330	12.077	13.900	11.384	9.568	6.670	12.175
4	Plant height (cm)	53.970	46.715	47.093	60.850	45.365	43.212	33.715	41.915
5	Plant spread (cm)	62.982	55.970	56.166	65.175	60.838	52.506	41.370	50.220
6	Number of leaves per plant	17.374	18.160	16.015	21.000	16.108	16.550	16.920	14.540
7	Stalk length (cm)	4.613	6.800	3.989	5.550	5.730	4.006	3.580	6.940
8	Days to maturity	92.588	92.700	94.150	123.750	89.800	92.557	96.500	89.100
9	Gross plant weight (g)	1160.409	878.750	848.794	2101.750	898.710	657.407	271.950	650.350
10	Marketable curd weight (g)	564.691	465.900	485.256	905.500	419.860	347.639	213.900	319.400
11	Net curd weight (g)	336.853	275.700	314.022	514.500	241.960	202.921	144.900	232.300
12	Curd length (cm)	7.017	6.640	6.919	6.600	9.692	6.680	5.560	7.500
13	Curd breadth (cm)	12.769	11.560	11.781	14.175	11.886	10.393	9.600	10.630
14	Harvest index	28.633	31.080	36.881	24.500	26.238	31.360	56.780	36.180
15	Curd size index (cm^2)	88.967	78.505	81.781	94.525	113.951	69.669	54.445	82.025
16	Curd yield (q/ha.)	169.162	140.750	145.756	271.500	125.230	104.093	63.950	96.350

In the present investigation it was found that maximum inter-cluster distance exists between cluster IV and VII followed by cluster IV and VIII. Considering all the group distances and agronomic performance, the inter-genotypic crosses between the member of cluster IV and VII and cluster IV and VIII would exhibit high heterosis and helps in the development of good recombinants. Cluster IV and VII consists of one genotype while four genotypes are included in cluster VIII respectively. PCF-7 is included in cluster VII and Pusa Pausjha in cluster IV. Whereas CAUVAR-2, INB 79, PC-98, PCF 29 are included in cluster VIII. Therefore it becomes necessary to emphasize more on cluster IV and cluster VII for selecting inbreds in cauliflower hybridization programme.

Geographical diversity may not be the only factor responsible for causing genetic diversification, thus focus should be laid on selecting the parents based on genetic diversity rather on geographical diversity. The results so obtained were found to be in conformation with the findings of Quamruzzaman *et al*. (2007) [8], Dey *et al.* (2011) [3], Santhosha *et al.* (2011) [9], Kumar *et al*. (2013) [5] and Dey *et al*. (2015) [2].

Conclusion

From the above results it can be concluded that the Eigen root of first principal component was accounted approximately 45.696% of total variation. The first five PC axes explained 81.72% of the variation, thereby suggesting considerable diversity among the genotypes for all the characters. The results revealed maximum inter-cluster distance between cluster IV and VII followed by cluster IV and VIII, thereby paving the chance for them to be used in hybridization breeding programmes. Cluster IV and VII consists of one genotype while four genotypes are included in Cluster VIII respectively. PCF-7 is included in cluster VII and Pusa Pausjha in cluster IV. Whereas CAUVAR-2, INB 79, PC-98, PCF 29 are included in cluster VIII.

References

1. Devaraju Varalakshmi B, Savithramma DL. Genetics of yield and its component traits in Early Cauliflower. Indian J. Hort. 2010; 67(3):339-342.
2. Dey SS, Bhatia R, Sharma S, Sharma K, Kumar Raj, Sharma VK *et al*. Genetic divergence and interaction among CUPRAC, FRAP and Total phenolics content in cauliflower (*Brassica oleracea* var. *botrytis* L.). Indian J. Genet. Plant Breed. 2015; 75(1):79-85.
3. Dey SS, Sharma SR, Parkash C, Barwal RN, Bhatia R. Genetic Divergence in Snowball Cauliflower (*Brassica oleracea* var. *botrytis* L.). Indian J. Plant Genet. Resour. 2011; 24(1):48-51.
4. Hotelling H. Analysis of complex of statistical variables into principal components. J. Educ. Psychol. 1933; 24:417-441.

5. Kumar V, Chand P, Bahuguna DK, Tyagi N, Kumar S. A study of character association and genetic divergence in Indian mustard (*Brassica juncea* L.). Prog. Agri. 2013; 13(1):187-192.
6. Mahalanobis PC. On the generalized distance in statistics. Proc. Natl. Acad. Sci. 1936; 2:49-55.
7. Pearson K. Onlines and planes of closest fit to systems of point in space. Philos. Mag. 1901; 2:559-572.
8. Quaimruzzaman AKM, Rahman MM, Uddin MN, Dalower M. Genetic diversity in cauliflower (*Brassica oleracea* L. var. *botrytis*). Indian J. Hort. 2007; 64(1):50-52.
9. Santhosha HM, Varalakshmi B, Narase Gowda NC. Genetic diversity in early cauliflower (*Brassica oleracea* var. *botrytis* L.) germplasm. J. Hortl. Sci. 2011; 6(1):21-24.
10. Sidki S. Morphology of the curd of cauliflower. Amer. Jour. Bot. 1962; 49:290-297.
11. Swarup V, Chatterjee SS. Origin and genetic improvement of Indian cauliflower. Econ. Bot. 1972; 26(4):381-393.
12. Varalaksmi B. Heterosis and combining ability for yields and its components in cauliflower. Indian J. Hortic. 2009; 66:198-203.
13. Verma U, Thakral NK, Neeru. Genetic diversity analysis in Indian mustard (*Brassica juncea* L.). Internat. J. Appl. Mathematics and Statistical Sciences 2016; 5(1):25-34.