

PRINCIPAL COMPONENT ANALYSIS IN OAT (*AVENA SATIVA L.*) GENOTYPES FOR GREEN FODDER YIELD AND ITS ATTRIBUTING TRAITS

ATMAN POONIA¹, D. S. PHOGAT^{1,*}, VERSHA¹ AND SHIVAM KUMAR²

¹Department of Genetics and Plant Breeding

²College of Agriculture

CCS Haryana Agricultural University, Hisar (125004), India

*(e-mail : dsphogatccshau@gmail.com)

(Received : 12 March 2021; Accepted : 30 March 2021)

SUMMARY

The present study on principal component analysis was conducted with 92 genotypes of oat grown in a randomized block design with three replications at the Research Farm of Forage Section, Department of Genetics and Plant Breeding, CCS HAU, Hisar during *rabi* 2015-16. The analyses of variance revealed highly significant differences among the genotypes for all the characters at genotypic level. Positive and significant correlations of dry matter yield were recorded with green fodder yield, plant height and seed yield. Path coefficient analysis revealed that characters green fodder yield, plant height and number of tillers/plant had positive and direct effects on dry matter yield. Genetic divergence classified genotypes into nine major clusters. The eight principal components had eigen values more than one and explained 70.01% of the total accumulated variability. The first principal component explained 12.78 % of the total variation followed by the second, third, fourth, fifth, sixth, seventh and eighth with 10.86, 9.81, 8.34, 8.08, 7.20, 6.89 and 6.04%, respectively.

Key words : Principal components, multivariate, genetic divergence, fodder yield, oat

Fodder oat is used as a multipurpose crop globally and nowadays attention is increasing towards higher grain yield components. They are usually winter sown, grazed prior to stem elongation and left to mature for its use as feed or milling grains. Multiple cuts are usually taken, after which part or all of the crop may be saved for seed (Kumar *et al.*, 2010). Oat has adequate soluble carbohydrates and fibres (Peterson *et al.*, 2005; Ratan *et al.*, 2016). Genetic diversity is a pre-requisite for any crop improvement programme as the genetic variability is fast depleting due to anthropological factors and continuous use of available variability in the on-going crop improvement programmes.

Principal component analysis is a multivariate statistical technique used to reduce the data with large number of correlated variables into a substantially smaller set of new variables, through linear combination of the variables that account for most of the variation present in the original variables (Chaudhary *et al.*, 2015). This is estimated either from correlation matrix or variance-covariance matrix but when the variables are measured with different units; scale effects can influence the composition of derived components. In such situations, it becomes desirable

to standardize the variables and the correlation matrix comes to the rescue. Principal factor analysis was carried out as it has many added advantages over principal component analysis. It is closely related to principal component analysis, but differs in that it assumes a definite model where each observed variable is expressed linearly in terms of common factor and unique factor. Moreover, in principal component analysis total variation contained in a set of variables is considered whereas, in factor analysis interest centers only on that part of variance, which is shared by the common factors.

MATERIALS AND METHODS

The experimental material comprising of ninety two (92) oat (*Avena sativa L.*) genotypes including checks, HJ-8 and OS-403 was conducted in randomized block design at the forage research farm, Department of genetics and plant breeding, CCS Haryana Agricultural University, Hisar during *Rabi* 2015-2016. The experimental material from ten different species of oat available in the gene pool at CCS HAU, Hisar collected from seven oat breeding institutes of India and three other countries (Algeria,

TABLE 1
List of 92 oat genotypes with respective species, genomic constitution and origin

S. No.	Centre of Origin	Ploidy (No. of genotypes)	Name of genotype, species and genomic constitution
1.	Algeria	Hexaploid (1)	Algerian (A. sativa L.-AACCCDD)
2.	Australia	Hexaploid (1)	Kent (A. sativa L.-AACCCDD)
3.	Bulgaria	Hexaploid (3)	Dunav, Kalojan, Dulo (A. sativa L.-AACCCDD)
4.	CSK HPKV, Palampur	Hexaploid (1)	PLP-1 (A. sativa L.-AACCCDD)
5.	GBPUA&T, Pantnagar	Hexaploid (2)	UPO-94, UPO-212 (A. sativa L.-AACCCDD)
6.	IGFRI, Jhansi	Hexaploid (5)	JHO-2006-4, JHO-822, JHO-851, JHO-99-1, JHO-2006-2 (A. sativa L.-AACCCDD)
7.	JNKVV, Jabalpur	Hexaploid (1)	JO-1 (A. sativa L.-AACCCDD)
8.	PAU, Ludhiana	Hexaploid (2)	OL-125, OL-10 (A. sativa L.-AACCCDD)
9.	SKAUST, Shri Nagar	Hexaploid (2)	Sabzar, SKO-90 (A. sativa L.-AACCCDD)
10.	CCS HAU, Hisar	Diploid (2) Tetraploid (4) Hexaploid (68)	HFO-305 (<i>Avena nuda</i> -AA), HFO 864 (<i>Avena brevis</i> -AA) HFO-58 (<i>A. barbata</i> -AABB), HFO-865 (<i>A. insularis</i> -AACC), HFO-867 (<i>A. maroccana</i> -AACC), HFO-870 (<i>A. vaviloviana</i> -AABB) HFO-60 (<i>A. byzantina</i>), HFO-504 (<i>A. fatua</i>), HFO-872 (<i>A. sterilis</i>) - AACCCDD HFO-267, FOS-1/29, OS-6, HFO-505, HFO-975, HFO-69, HFO-876, HFO-877, HFO-879, HFO-502, HFO-78, HFO-603, HFO-414, HFO-409, HFO-508, HFO-498, HFO-875, HFO-878, HFO-874, HFO-523, HFO-433, HFO-885, HFO-896, HFO-839, HFO-836, HFO-863, HFO-884, HFO-841, HFO-851, HFO-880, HFO-893, HFO-852, HFO-862, HFO-845, HFO-883, HFO-831, HFO-114, HFO-603, HFO-832, HFO-605, HFO-611, HFO-704, HFO-715, HFO-610, HFO-706, HFO-703, HFO-575, HFO-614, HFO-707, HFO-905, HFO-908, HFO-914, HFO-909, HFO-904, HFO-910, HFO-913, HFO-921, HFO-924, HFO-919, HFO-906, HFO-912, HFO-902, HFO-920, HJ-8 & OS-403 (<i>A. sativa</i> L.-AACCCDD)

Australia and Bulgaria). Each genotype was sown in three rows of 3 m length with a row-to-row distance of 25 cm and plant to plant spacing of 10 cm with recommended agronomic practices.

Climate and Weather conditions

The experimental site is situated at 29°10' N latitude and 75°44' E longitude with an altitude of 228 meters above the mean sea level and is a semi-arid and subtropical region with prevailing hot and dry winds during summer months. Both, winter and summer are usually harsh to bear upon. The mean minimum and maximum temperature exhibit wide variations. A maximum temperature zooming 44 to 48°C during summer and temperature dipping as low as to freezing point accompanied with chill frost in winter is of common occurrence (Fig. 1).

Statistical analysis : Principal component analysis, basically a data reduction technique, initially floated by Pearson (1901) and later on developed by Hotelling (1936) offers solution to the complex problem of large and unmanageable data by transforming the original set of variables into a smaller set of linear combinations that account for most of the variability of the original set. The INDOSTAT software was used to analyze the data.

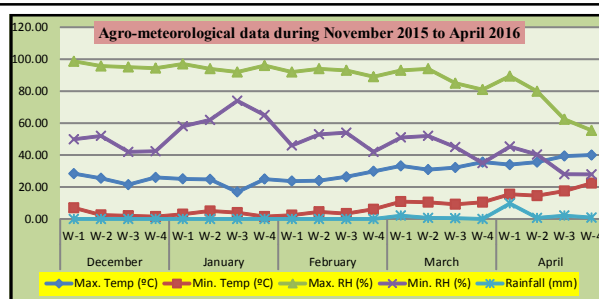


Fig. 1. Agro-meteorological data during the period of experimentation from November 2015 to April 2016.

RESULTS AND DISCUSSION

Analysis of variance and mean values of green fodder yield attributing traits

Analysis of variance revealed the presence of significant amount of variation for all the traits studied, which indicated the suitability of the data for further analysis (Poonia *et al.*, 2017b). In the present investigation, mean values of green fodder yield and its component characters of 92 oat genotypes were also considered for selecting the superior genotypes for different yield attributing traits. The number of days to 50% flowering was maximum in HFO 611 (100 days) and the minimum HFO 502 and HFO 874 (86

days) followed by HFO-267, OL 125, OL 10, JHO 851, DULO, UPO 212, OS 6, HFO 505, HFO-975, HFO-876, HFO-877, HFO-and HFO- 904 while maximum (133.7 days) and the minimum (110.3 days) number of days to maturity were observed in genotypes, HFO-877 and HFO-893, respectively and genotypes attained maturity early are HFO-975, HFO-504, HFO-884, HFO-502, HFO-885, HFO-880, HFO-852, HFO-862, HFO-883, HFO-603, HFO-614, HFO-924 and HFO-906.

The maximum plant height was recorded in the genotype, HFO-864 (150 cm) followed by JO-1, OL-125, JHO-822, FOS-1/29, DULO, UPO-212, Kent, HFO-603, HFO-409, HFO-605, HFO-611, HFO-614 and HFO-118 and minimum in HFO-267 (70 cm) while HFO-832 (11.5) and HFO-909 (7.2 cm) showed highest and lowest number of tillers/plant. The highest seed yield/plant was recorded by HFO-912 (174.4 g) followed by JO-1, Algerian, Kent, HFO-504, HFO-78, HFO-867 and HFO-704 whereas lowest yield was recorded in Sabzar (41.3 g). The highest 100 seed weight was observed HFO-878 (4.56 g) and the lowest in Sabzar (2.01 g).

The highest mean value of green fodder yield was observed in genotype HFO-58 (1.847 Kg) and the lowest in HFO-852 (0.468 Kg). The genotypes JO-1, JHO-99-1, DULO, UPO-212, Kent, HFO-864, HFO-879, HFO-58, HFO-878 and HFO-409 found superior than checks for green fodder yield while highest mean value of dry matter yield was found in JO-1 (0.357 Kg) followed by Algerian, Sabzar, JHO-851, JHO-99-1, FOS-1/29, DULO, HFO-876, Kent, HFO-504, HFO-877, HFO-879, HFO-508, HFO-845, HFO-883, HFO-611, HFO-715, HFO-908, HFO-914, HFO-910, HFO-913 and HFO-912 and the lowest in HFO-69 (0.107 Kg).

Correlation Analysis

The correlation analysis reported that the plant height, dry matter yield, number of tillers per plant and seed yield exhibited the highly significant and positive correlation with the green fodder yield (Poonia *et al.*, 2017a). Thus, the improvements in characters found significantly and positively correlated will improve green fodder yield directly and indirectly. While negative direct effect was observed by traits like peduncle length, number of spikelets/panicle, number of leaves per plant and number of days to 50% flowering which will dilute the positive and direct effect of earlier traits on green fodder yield. So the negatively correlated traits are required to be kept

within a threshold limit. Path coefficient analyses revealed that characters green fodder yield, plant height and number of tillers/plant had positive and direct effects on dry matter yield.

Divergence Analysis

The D² analysis of 92 oat genotypes clustered them into nine clusters and cluster IX exhibited maximum (18) while cluster I had the minimum (5) number of genotypes. Highest inter-cluster distance was observed between cluster IV and IX whereas lowest was found between cluster I and VIII. The distribution of mean values revealed that the clusters I, II and IV had better cluster means for most of the characters (Poonia and Phogat, 2017). So, the genotypes JO 1, Algerian, HFO 912, DULO, UPO 212, OS 6, HFO 305, HJ 8, OS 403, HFO 879, HFO 878, JHO-2006-4, JHO 851, HFO 832, HFO 409, HFO 831 and FOS-1/29 were grouped into these clusters could be used in crossing programme to obtain high heterotic response and thus better segregants in subsequent generations for dry matter yield in forage oat.

Principal component analysis

In the present investigation, principal components were used to identify the most relevant characters and present them in more visualized dimensions through linear combinations of variables that account for most of the variation present in original set of variables. The eight principal components had eigen values more than one and altogether explained 70.01% of the total accumulated variability (Table 2). The first principal component explained 12.78 % of the total variation and followed by the second, third, fourth, fifth, sixth, seventh and eighth principal components explained 10.86, 9.81, 8.34, 8.08, 7.20, 6.89 and 6.04% of the total variance, respectively.

Further, principal factor analysis was carried out, because the principal component analysis does not assume a definite model. In principal factor analysis each observed variable is expressed linearly in terms of a common factor and a unique factor. The common factors account for the correlation among the variables. While each unique factor accounts for the remaining variance (including error) of that variable. Moreover, in principal component analysis, the total variation contained in a set of variables is considered, whereas, in factor analysis interest centres on that part of variance which is shared by the common factors. Also in contrast to principal component analysis, here the

TABLE 2
Total variance explained by different principal components in 92 oat genotypes

Component	Eigen values	Variation Explained %	Cumulative variation Explained %
1	4.18	12.78	12.78
2	2.73	10.86	23.64
3	2.35	9.81	33.45
4	1.88	8.34	41.80
5	1.77	8.08	49.88
6	1.47	7.20	57.08
7	1.37	6.89	63.97
8	1.05	6.04	70.01

variables axes are allowed to interact resulting in distortion of mutual orthogonality.

Principal factor analysis

The principal factor analysis carried out without any rotation did not derive the clear picture of interaction among the characters. This did not provide a clear picture regarding the idea of character

association with respective principal factor as some factors had very high loading of variable and some have none. This indicated the next alternative of factor analysis i.e. rotation method of Kaiser (1958). To select the relevant characters in various principal factors, the correlation values ($> (\pm 0.5)$) were considered as relevant for that principal factor. Factor loading of different characters with varimax rotation have been presented in Table 3.

Data presented in Table 3 clearly indicated that PF-1 was loaded on germination %, seedling length and seed vigour index-I, PF-2 showed a strong and positive factor with the traits like plant height, green fodder yield and dry matter yield while the PF-3 was loaded with leaf width and PF-4 showed the diversity among accessions based on number of leaves per plant, numbers of tillers per plant and days to maturity. PF-5 confirms the variation observed by number of spikelets per panicle and seed index whereas PF-6, PF-7 and PF-8 were loaded with days to 50% flowering, seed yield and leaf length, respectively.

TABLE 3
Factor loading of different characters with respect to different principal factor (Varimax rotation) in 92 oat genotypes

Character	Components							
	PF-1	PF-2	PF-3	PF-4	PF-5	PF-6	PF-7	PF-8
PH	0.22	0.618	-0.037	-0.227	0.291	-0.064	-0.096	0.038
DF	-0.163	0.08	0.095	-0.577	-0.04	0.545	0.126	0.226
DM	-0.482	0.031	0.173	0.278	0.216	0.359	0.04	-0.159
FLL	0.342	0.435	0.384	0.084	-0.134	-0.123	-0.062	-0.223
IL	0.392	0.197	0.257	0.408	-0.251	-0.161	0.377	0.016
TPL	-0.357	0.164	-0.155	0.549	0.022	0.492	0.078	-0.042
PL	0.371	-0.152	0.187	0.407	-0.366	-0.245	-0.277	0.096
AL	0.496	0.239	0.404	0.152	-0.204	-0.047	0.237	0.281
NOS	0.044	0.119	-0.303	0.28	0.650	-0.054	-0.002	0.105
SY	0.293	0.377	-0.058	0.135	0.007	0.047	0.605	0.335
SI	0.386	-0.032	-0.233	0.252	0.460	-0.24	0.053	0.378
LL	0.309	0.35	0.317	0.226	-0.217	0.244	-0.008	0.513
LW	0.069	0.021	0.506	0.042	-0.063	0.254	0.379	0.033
GFY	0.221	0.773	0.011	-0.177	0.24	0.093	0.045	-0.099
DMY	0.273	0.720	0.01	-0.142	0.138	0.048	0.032	-0.184
NOLS	-0.409	0.2	-0.238	0.628	0.023	0.323	-0.06	0.098
CPg	0.316	-0.258	-0.252	0.172	0.048	-0.025	0.391	-0.305
CPf	0.467	-0.114	-0.098	-0.031	-0.305	0.298	-0.502	-0.026
GP	0.547	0.126	-0.494	-0.13	-0.066	-0.022	0.319	0.027
SL	0.697	-0.246	-0.363	-0.035	-0.142	0.386	-0.013	0.023
SDW	0.39	-0.432	0.598	-0.027	0.452	0.186	-0.115	-0.083
SVI	0.752	-0.141	-0.483	-0.077	-0.139	0.292	0.113	0.024
SVII	0.491	-0.409	0.521	-0.046	0.448	0.182	-0.063	-0.088
EC	-0.631	0.195	0.066	-0.205	-0.278	0.019	0.042	0.164

PH- Plant height (cm), DF - No. of days to 50% flowering, DM- Number of days to maturity, FLL- Flag leaf length (cm), IL- Internode length (cm), TPL-Number of tillers per plant, PL- Peduncle length (cm), AL- Axis length (cm), NOS- Number of spikelets per panicle, SY- Seed yield(g), SI-100 seed weight(g), LL- Leaf length(cm), LW- Leaf width(cm), GFY- Green fodder yield(kg), DMY- Dry matter yield(kg), NOLS- Number of leaves per plant, CPg- Crude protein in grain (%), CPf- Crude protein in forage (%) GP- Germination %, SL- Seedling length(cm), SDW- Seedling dry weight(mg), SVI- Seed vigour index-1, SVII- Seed vigour index-2, EC- Electrical conductivity(mS/cm/seed).

Among these principal factors, PF-2 and PF-7 regarded as fodder yield factors and seed yield factors. In PF-2 the PH (0.618), GFY (0.773) and DMY (0.720) values were highest suggesting that PF-2 can be regarded as fodder yield factor as all the green yield contributing factors are present in it while the PF-7 showed highest value for SY (0.605) and called as seed yield factor.

Vaisi *et al.* (2013) and Krishna *et al.* (2014) also conducted principal component analysis in oat and suggested to transfer many correlated variables into a few independent principal components explaining much of the variability of the original set. Hemavathy (2020) conducted principal component analysis study in sweet corn and reported the similar results.

The principal factor analysis was carried out using principal component method, which does not require assumption of multivariate normal distribution of population in contrast to the other methods like maximum likelihood method (Jaiswal, 2000). Initially the data were analyzed without any rotation but it failed to load all the variables meaning thereby that it could not provide much information regarding the idea of correlation between the variables and the principal factors.

The failure of principal factor analysis without rotation to draw sensible conclusions prompted to go for analysis with rotation. Varimax method of orthogonal rotation (Kaiser, 1958) was utilized to rotate the factor axes. This is the most commonly used method and can be placed in a meaningful biological context.

The present investigation showed significantly high amount of variability for all traits that could be exploited for crop improvement. The green fodder yield and seed yield associated traits were identified which directly or indirectly responsible for their improvement. The clustering pattern could be utilized in finding the best cross combinations for generating variability with respect to various characters under study. Clustering pattern of genotypes indicated that there was no association between geographical distribution of accessions and genetic divergence (Murthy and Arunachalam, 1966). Cluster analysis groupings are useful to breeders in identifying desirable genotypes that may be used as parents in breeding for improvement of desirable traits under study. Therefore, all the information generated through principal component and factor analysis will reduce the overall time required by plant breeders to screen large populations for potential breeding stock.

REFERENCES

- Chaudhary, S., P. Sagar, B. K. Hooda, and R. K. Arya, 2015 : Multivariate analysis of pearl millet data to delineate genetic variation. *Forage Res.*, 40, 201-208.
- Hotelling, H, 1936 : Relations between two sets of varieties. *Biometrika*, 28(3-4):321-377.
- Jaiswal, U. C., 2000 : Principal components analysis and its interpretation. II. Factor analysis and its application in social sciences. In course compendium of "Use of statistical methods in agricultural and allied fields" Jan. 19- Feb. 15, 2000, *Academy of Agricultural Research and Management*, CCS Haryana Agricultural University, Hisar. 166-181.
- Hemavathy, A. T, 2020 : Principal component analysis in sweet corn (*Zea mays* L. *saccharata*.). *Forage Res.*, 45(4): 264-268.
- Kaiser, H.F. 1958 : The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23 : 187.
- Krishna, A., S. Ahmed, H. C. Pandey, and V. Kumar, 2014 : Correlation, path and diversity analysis of oat (*Avena sativa* L.) genotypes for grain and fodder yield. *Journal of Plant Science & Research*, 1: 1-9.
- Kumar, A., R. K. Arya and S. K. Pahuja 2010 : Effect of cutting and fertilizer management on seed quality parameters of oat (*Avena sativa* L.). *Forage Res.*, 35 : 198-200.
- Murthy, B. R. and V. Arunachalam, 1966 : The nature of genetic divergence in relation to breeding system in crop plants. *Ind. J. Genet. and Pl. Breed.*, 26 : 188-198.
- Pearson, K. 1901 : On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2 : 559.
- Peterson, D. M., D. M. Wesenberg, D. E. Burrup, and C. A. Erickson, 2005 : Relationships among agronomic traits and grain composition in oat genotypes grown in different environments. *Crop Science*, 45(4) : 1249-1255.
- Poonia, A., D. S. Phogat, and P. Kumari, 2017b : Genetic variability parameters for morphological and quality attributes in fodder oat (*Avena sativa* L.). *Trends in Biosciences*, 10(37) : 7687-7691.
- Poonia, A., and D. S. Phogat, 2017 : Genetic divergence in fodder oat (*Avena sativa* L.) for yield and quality traits. *Forage Research*, 43 : 101-105.
- Poonia, A., D. S. Phogat, S. K. Pahuja, A. Bhuker, and R. S. Khatri, 2017a : Variability, Character Association and Path Coefficient Analysis in fodder Oat for yield and quality traits. *Forage Res.*, 43(3) : 239-243.
- Ratan, N., U. N. Singh, and H. C. Pandey, 2016 : Yield and quality of oat (*Avena sativa* L.) as influenced by nitrogen and varieties in Bundelkhand region (UP) India. *Agricultural Science Research Journal*, 6(1): 27-30.
- Vaisi, H., A. R. Golparvar, A. Resaie, and S. Bahraminejad, 2013 : Factor analysis of some quantitative attributes in oat (*Avena sativa* L.) genotypes. *Scientia Agriculturae*, 3(3) : 62-65.