Character association and principal component analysis for yield and its attributing characters in greengram

*G. NAYAK, D. LENKA, M. DASH AND ¹S.K. TRIPATHY

Department of Plant Breeding and Genetics, ¹Department of Agricultural Biotechnology, Odisha University of Agriculture and Technology, Bhubaneswar-751003, Odisha

ABSTRACT

In the present investigation, 66 greengram genotypes were subjected to study the character association and principal component analysis for yield and its attributing characters. Seed yield plant⁻¹ had highly significant positive correlation with hundred seed weight, number of seeds plant⁻¹, pod length, number of pods plant⁻¹, number of clusters plant⁻¹ and plant height, whereas characters like plant height, number of pods plant⁻¹, pod length, number of seeds pod⁻¹ and hundred seed weight had direct positive effect on seed yield. Principal component analysis revealed that the first five principal components accounted for 80.54 % of the total variance for all the traits, with values 24.66%, 20.83%, 13.47%, 12.56%, and 9.02% for PC-1, PC-2, PC-3, PC-4 and PC-5 respectively. In the scatter plot, the genotypes HUM-10, PDM-139 and ML-613 were found to be far off from other genotypes, indicating that they may be distinct from other genotypes and hence a good candidate for hybridization.

*Keywords***:** *Character association, genotype, greengram, path coefficient and principal component*

Greengram (*Vigna radiata* (L.) Wilczek) is an important and well-known leguminous crop grown in India. It belongs to the family *Fabaceae*, subfamily *Faboideae* and tribe *phaseoleae*. It is a protein-rich food containing around 24% protein. The protein is rich in lysine, which is mostly deficit in cereal grains (Baskaran *et al*., 2009). It is popular due to its high nutritional value and adaptability to multiple cropping systems. Thus it can be grown as a sole crop and as an intercrop or mixed crop with cereals forming an essential constituent of crop rotation. They are regarded as ideal crops for crop intensification and diversity due to their short duration and photo-thermo insensitivity (Ali and Kumar, 2006). Major greengram producing states in India are Andhra Pradesh, Karnataka, Maharashtra, Odisha, Rajasthan and Uttar Pradesh (Jakhar and Kumar, 2018). In India, area under greengram cultivation is 29.03 million hectares with a total production of 23.40 million tonnes and average productivity of 8.06 quintals ha⁻¹. In Odisha, its productivity is 559 kg ha⁻¹ (Directorate of Economics and Statistics, 2019).

Though greengram is considered as an important pulse crop, the productivity is low and the areas for greengram cultivation in Odisha are limited due to lack of genetic expertise, low genetic diversity, poor yield, and vulnerability to a wide range of insect pests and diseases in monoculture (Panigrahi and Baisakh, 2014). As there was not much systematized breeding work undertaken for breeding of superior high yielding genotypes in greengram, an attempt was made for yield and genetic improvement of greengram through character association and principal component analysis (PCA), which might lead to a better understanding of the diversity in different landraces and the creation of diverse lines for any further breeding work.

MATERIALS AND METHODS

During *rabi* 2019, the experiment was conducted in a randomized block design with three replications at EB-II, Dept. of Plant Breeding and Genetics, Odisha University of Agriculture and Technology, Bhubaneswar. The 66 greengram germplasm lines used in this study included widely adopted local landraces, important pre-released varieties and standard ruling varieties. The crop was raised using recommended agronomic packages and practices, and it was harvested in December 2019. Five random plants were selected to take the observations on days to 50% flowering (DF), days to maturity (DM), plant height (PH) in cm, number of branches plant⁻¹ (B/P), number of clusters plant⁻¹ (C/P), number of pods plant⁻¹ (P/P), pod length (PL) in cm, number of seeds pod⁻¹ (S/P), 100 seed weight (HSW)in gram and seed yield plant⁻¹ (SY/P) in greengram.

The genotypic and phenotypic correlation coefficients between each pair of characters were estimated using the formula proposed by Al-Jibouri *et al.* (1958), and the path coefficient analysis was accomplished using seed yield plant⁻¹ as the dependent variable against other independent variables, as suggested by Dewey and Lu (1959). PCA was done using XLSTAT software to estimate the contribution of characters to the total variability for all the characters under study and to measure divergence among

Email: gyanishanayak@gmail.com

genotypes in terms of spatial distance in a two-way pictorial graph fixing the relative position of each genotype.

RESULTS AND DISCUSSION

Character association

For majority of the characters in the current study, genotypic correlation coefficients (r_g) were larger in magnitude than phenotypic values (r_p), showing that the quantitative characters were influenced by the environment (Singh and Narayanan, 2017) (Table1).Thus, larger genotypic correlations in the case of positive association and lower genotypic correlations in the case of negative association would emphasize phenotypic selections. Seed yield plant[-1] had highly significant positive correlation with hundred seed weight (0.377), number of seeds pod[-1] (0.253), pod length (0.196), number of pods plant[-1] (0.527), number of clusters plant[-1] (0.380) and plant height (0.494) at the phenotypic level. Days to maturity had positive correlation with seed yield plant[-1] (0.073). Similarly, seed yield plant[-1] showed highly significant positive correlation with hundred seed weight (0.401), number of seeds pod[-1](0.281), pod length (0.231), number of pods plant[-1] (0.575), number of clusters plant[-1] (0.413) and plant height (0.570) at genotypic level also. Further, it was positively correlated with number of branches plant[-1] (0.154) at 5% level of significance.Marappa *et al.* (2010), Tabasum *et al.* (2010) and Patel *et al.* (2014) also reported similar results. Hence, improving these characters could improve yield. Further, yield was negatively correlated with days to 50% flowering and days to maturity at either level (Mishra *et al.,* 2014 and Raturi *et al.,*2015).

Path coefficient analysis was used to split the phenotypic and genotypic correlation of seed yield with the 9 component characters into direct and indirect effects of component characters on yield (Table 2 and 3). At both the phenotypic and genotypic levels, the number of pods per plant had the greatest direct positive effect on seed yield (0.450, 0.651). This was comparable with the reports of Mohon *et al.* (2019) for days to 50% flowering. Further, plant height (0.293, 0.413), pod length (0.110, 0.184), number of seeds pod[-1] (0.137, 0.115) and hundred seed weight (0.272, 0.246) showed direct positive effect on seed yield at both phenotypic and genotypic levels, which have been earlier observed by several workers(Tabasum *et al.,* 2010, Patel *et al.,* 2014 and Mohon *et al.,* 2019). Characters like days to 50% flowering (-0.116,-0.071), days to maturity (-0.080,-0.199) number of branches plant[-1](-0.043, -0.018) and number of clusters plant[-1](-0.028, -0.283) exhibited negative direct effects on seed yield at phenotype and genotype level. This finding was similar to the finding

Table 1: Phenotypic (above diagonal) and Genotypic (below diagonal) correlation coefficients among 10 characters of 66 greengram genotypes

Characters	DF	DM	PH	B/P	C/P	P/P	PL	S/P	HSW	SY/P
DF		0.917**	0.101	0.180*	0.053	-0.016	0.126	0.096	-0.156*	-0.212**
DM	0.953**		0.119	0.063	-0.009	-0.102	0.131	0.134	-0.098	-0.193**
PH	0.109	0.137		0.052	0.449**	0.430**	0.058	0.214**	0.030	0.494**
B/P	0.208**	0.057	0.025		0.246**	0.203**	0.134	0.049	0.077	0.073
C/P	0.049	-0.013	0.577**	0.301**		0.704**	-0.111	-0.015	-0.094	0.380**
P/P	-0.029	-0.086	0.493**	0.295**	0.829**		-0.118	-0.039	-0.006	0.527**
PL	0.196**	0.242**	0.031	0.259**	-0.192**	-0.194**		0.427**	0.334**	0.196**
S/P	0.122	0.176*	0.229**	0.093	-0.018	-0.054	0.536**		0.172*	0.253**
HSW	-0.223**	-0.151*	0.010	0.091	-0.155*	-0.066	0.452**	0.198**		0.401**
SY/P	-0.256**	-0.235**	0.570**	0.154*	0.413**	0.575**	0.231**	0.281**	0.401**	

Note: ** and * indicate significance at 1% and 5% levels probability respectively
[DF: Days to 50% flowering; DM:Days to maturity;PH: Plant height; B/P: Number of branches plant[-1]; C/P : Number of clusters plant[-1];P/P:Number of pods plant[-1];PL.: Pod length;S/P:Number of seeds pod[-1]; HSW: Hundred **seed weight; SY/P:Seed yield plant[-1]**]

Table 2: Path analysis of 9 component characters on yield in greengram at phenotypic level

Characters	DF	DM	PH	B/P	C/P	P/P	PL	S/P	HSW	r$_p$ withSY/P
DF	**-0.116**	-0.073	0.029	-0.008	-0.002	-0.027	0.014	0.013	-0.042	-0.212**
DM	-0.106	**-0.080**	0.035	-0.003	0.001	-0.046	0.015	0.018	-0.027	-0.193**
PH	-0.012	-0.010	**0.293**	-0.002	-0.014	0.194	0.006	0.029	0.008	0.494**
B/P	-0.021	-0.005	0.015	**-0.043**	-0.007	0.091	0.015	0.007	0.021	0.073
C/P	-0.006	0.001	0.146	-0.010	**-0.028**	0.317	-0.012	-0.002	-0.026	0.380**
P/P	0.007	0.008	0.125	-0.009	-0.020	**0.450**	-0.013	-0.005	-0.018	0.527**
PL	-0.015	-0.010	0.017	-0.005	0.003	-0.053	**0.110**	0.059	0.090	0.196**
S/P	-0.011	-0.010	0.063	-0.002	0.000	-0.017	0.047	**0.137**	0.046	0.253**
HSW	0.018	0.008	0.009	-0.003	0.003	-0.030	0.037	0.023	**0.272**	0.377**

Residual effect P(R) = 0.443

Note: Diagonal bold figure shows direct effects

[DF: Days to 50% flowering; DM:Days to maturity;PH: Plant height; B/P: Number of branches plant[-1]; C/P: Number of clusters plant[-1]; P/P:Number of pods plant[-1];PL: Pod length; S/P:Number of seeds pod[-1]; HSW: Hundred seed weight; SY/P:Seed yield plant[-1]]

Table 3: Path analysis of 9 component characters on yield in greengram at genotypic level

Characters	DF	DM	PH	B/P	C/P	P/P	PL	S/P	HSW	r$_g$ with SY/P
DF	**-0.071**	-0.189	0.044	-0.003	-0.014	-0.018	0.036	0.014	-0.055	-0.256**
DM	-0.067	**-0.199**	0.057	-0.001	0.004	-0.056	0.044	0.020	-0.037	-0.235**
PH	-0.008	-0.027	**0.413**	-0.000	-0.163	0.320	0.005	0.026	0.002	0.570**
B/P	-0.015	-0.011	0.010	**-0.018**	-0.085	0.192	0.048	0.011	0.022	0.154*
C/P	-0.003	0.003	0.238	-0.005	**-0.283**	0.539	-0.035	-0.002	-0.038	0.413**
P/P	0.002	0.017	0.204	-0.005	-0.234	**0.651**	-0.036	-0.006	-0.016	0.575**
PL	-0.014	-0.048	0.013	-0.004	0.054	-0.127	**0.184**	0.062	0.111	0.231**
S/P	-0.009	-0.035	0.095	-0.002	0.005	-0.035	0.098	**0.115**	0.049	0.281**
HSW	0.016	0.030	0.004	-0.002	0.044	-0.043	0.083	0.023	**0.246**	0.401**

Residual effect P(R) = 0.272

Note: Diagonal bold figure shows direct effects

[DF: Days to 50% flowering; DM:Days to maturity;PH: Plant height; B/P: Number of branches plant[-1]; C/P: Number of clusters plant[-1]; P/P:Number of pods plant[-1];PL: Pod length; S/P:Number of seeds pod[-1]; HSW: Hundred seed weight; SY/P:Seed yield plant[-1]]

Table 4: Eigen value, proportion of total variance, cumulative percent variance and component loadings of 66 greengram genotypes

PCA	PC-1	PC-2	PC-3	PC-4	PC-5
Eigen Value	2.466	2.083	1.347	1.256	0.902
Variance (%)	24.66	20.83	13.47	12.56	9.02
Cumulative Variance (%)	24.66	45.49	58.96	71.52	80.54
DF	0.14	0.56	0.08	0.27	0.09
DM	0.03	0.57	0.26	-0.08	0.23
PH	0.41	-0.03	0.45	-0.28	0.05
B/P	-0.08	-0.21	0.02	0.76	-0.23
C/P	0.54	-0.16	0.09	0.19	0.07
P/P	0.49	-0.12	0.19	0.31	0.20
PL	-0.37	0.22	0.23	0.32	0.33
S/P	-0.09	0.19	0.49	0.04	-0.73
HSW	-0.30	-0.27	0.33	0.08	0.45
SY/P	-0.19	-0.35	0.54	-0.16	0.03

of Tabasum *et al.* (2010) for number of primary branches, plant height, number of pods cluster^{-1} and number of clusters plant^{-1}. When the correlation coefficient is positive but the direct effect is negative or small, the indirect effect appears to be the reason of the association, necessitating the consideration of indirect causative factors. The number of branches plant^{-1} and the number of clusters plant^{-1} in this study both exhibited these types of relationships at either level. Plant height through number of pods plant^{-1} had the highest indirect effect on seed yield, followed by number of clusters plant^{-1} at any level, as previously observed by Marappa *et al.* (2010). The residual effect in genotypic path coefficient analysis was low (0.272), but high in phenotypic path coefficient (0.443), which might require consideration of some additional characters for more accurate result (Roy *et al.,* 2020).

Principal component analysis

In the present study, the first five principal components (PC-1 to PC-5) with eigen values of 2.466, 2.083, 1.347, 1.256 and 0.902, respectively accounted for 80.54% of the total variance for all the traits (Table 4). Higher the absolute value in the PC, higher would be the contribution of characters towards the divergence (Singh *et al.*, 2017). The first principal component (PC-1) provided 24.66% of the variability, with the most important contributing variables being the number of clusters plant^{-1} (0.54), number of pods plant^{-1} (0.49) and plant height (0.41). Similarly, substantial loadings of days to maturity (0.57) and days to 50% flowering (0.56) were found in the second principal component (PC-2), which explained 20.83% of total variance. Seed yield plant^{-1} (0.54), number of seeds pod^{-1} (0.49) and plant height (0.45) collectively contributed 13.47% of total

variance to the third principal component (PC-3). With 12.56% of total variance, the fourth principal component (PC-4) was characterized by high loadings for number branches plant^{-1} (0.76), pod length (0.32) and number of pods plant^{-1} (0.31), indicating that these traits could be effectually used for selection among germplasms for yield enhancement in greengram. Principal components with eigen values larger than 1 should be kept as potential diversity contributors (Kaiser, 1960). With an eigen value of less than 1.0 and a variability of 9.02%, the fifth principal component (PC-5) contributed less to divergence. Hence, traits like number of clusters plant^{-1}, number of pods plant^{-1}, plant height, days to maturity, days to 50% flowering and seed yield plant^{-1} contribute the most to divergence. This finding was similar with the results of Jakhar and Kumar (2018) for plant height and Mahalingam *et al.* (2019) for seed yield plant^{-1}. With an eigen value of less than 1.0 and a variability of 9.02%, the fifth PC contributed less to divergence. The findings were similar to those of Jakhar and Kumar (2018) and Mahalingam *et al.* (2019).

The first 3 principal components contributed 58.96% to the total variance with eigen value more than one. As a result, scores of PCA for 66 genotypes in the first three principal components were calculated and displayed onto a graph to produce the 2D scatter diagram taking PCA I in X-axis and PCA II in Y-axis (Table 5, Fig. 1). The 2D structure showed that the genotypes numbered 18 (HUM-10), 61 (PDM-139) and 30 (ML-613) were scattered relatively away from other genotypes. The combinations between these genotypes would give better recombinants (Thippani *et al.,* 2017).

The concurrent studies on character association and their direct and indirect effects on seed yield in present research led to the conclusions that traits like number of pods plant^{-1}, plant height, hundred seed weight,

Table 5: PCA scores of 66 greengram genotypes

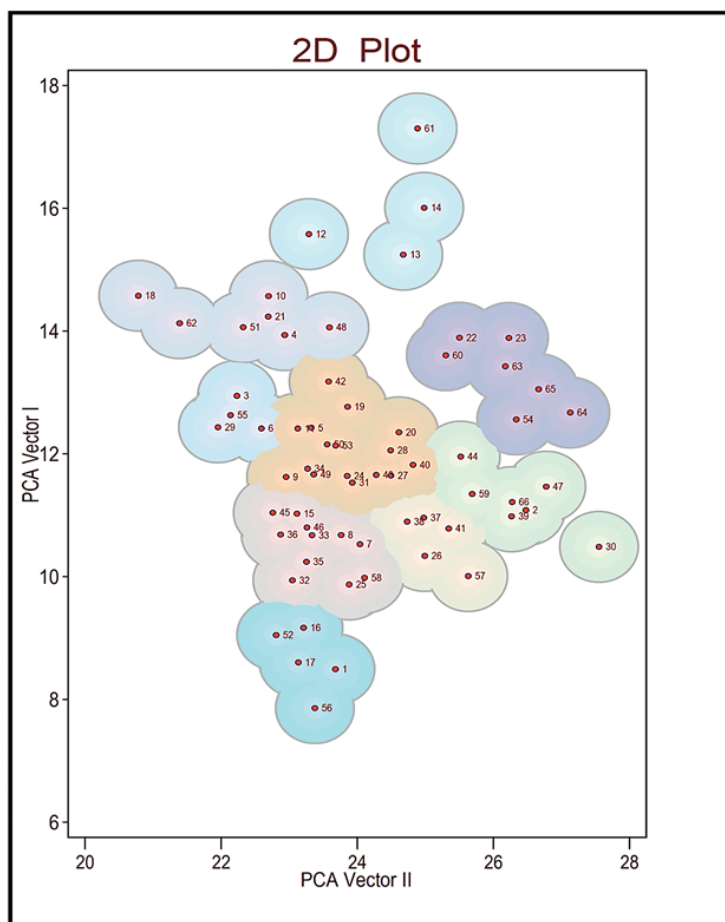| Sl. No	Genotype	PCA I X Vector	PCA II Y Vector	PCA III Z Vector
1.	AKM-9901	8.491	23.680	18.821
2.	AKM-8802	11.082	26.479	19.872
3.	AKM-9601	12.944	22.232	21.802
4.	AM-2057	13.934	22.935	19.741
5.	AB-2557	12.424	23.325	23.962
6.	ADT-14	12.413	22.593	20.257
7.	B-10-33-1	10.527	24.039	18.854
8.	CO-6	10.677	23.763	21.307
9.	CNO-3	11.623	22.955	22.091
10.	CNO-35	14.569	22.698	21.421
11.	CNO-36	12.411	23.127	21.092
12.	CNO-59	15.578	23.290	24.891
13.	COGG-902	15.244	24.676	21.998
14.	COGG-912	16.006	24.983	21.547
15.	Dhauli	11.025	23.115	18.942
16.	HUM-1	9.167	23.212	24.117
17.	HUM-6	8.602	23.136	22.625
18.	HUM-10	14.574	20.783	18.386
19.	HUM-12	12.767	23.858	20.433
20.	KM-851	12.349	24.613	22.079
21.	KM-9309	14.235	22.693	20.879
22.	LGG-460	13.892	25.502	21.049
23.	M-9-2	13.887	26.229	21.292
24.	MG-12	11.640	23.853	20.389
25.	M-9-13	9.870	23.881	20.907
26.	MGG-349	10.337	24.993	19.842
27.	ML-131	11.646	24.496	19.889
28.	ML-267	12.056	24.491	19.588
29.	ML-555	12.431	21.954	19.896
30.	ML-613	10.484	27.552	20.716
31.	ML-682	11.529	23.927	18.276
32.	ML-729	9.939	23.049	19.491
33.	ML-818	10.674	23.336	20.017
34.	ML-881	11.758	23.271	21.053
35.	ML-1108	10.241	23.255	21.196
36.	ML-1165	10.683	22.875	19.708
37.	ML-1231	10.958	24.977	20.772
38.	OUM-99-3	10.895	24.734	21.257
39.	OUM-99-6	10.981	26.268	19.910
40.	OUM-99-7	11.820	24.820	20.589
41.	OUM-11-5	10.783	25.345	19.904
42.	OUM-14-0	13.176	23.579	21.387
43.	OUM-14-2	11.655	24.279	21.485
44.	OUM-18	11.954	25.521	20.692
45.	OUM-18-4	11.041	22.759	19.059
46.	OUM-20	10.800	23.261	19.660
47.	OUM-20-1	11.465	26.778	19.064
48.	OUM-21	14.059	23.592	18.893
49.	Pusa Bold	11.664	23.360	21.617
50.	KPS-2	12.153	23.559	22.472
51.	SML-668	14.062	22.324	21.394
52.	EC-693369	9.047	22.810	23.131
53.	TARM-1	12.133	23.684	21.101
54.	PDM-154	12.559	26.337	20.203
55.	Sujata	12.629	22.139	21.133
56.	NM-92	7.857	23.378	19.261
57.	ML-1628	10.009	25.631	21.102
58.	NM-94	9.981	24.108	21.157
59.	IPM-02-14	11.346	25.690	19.878
60.	ML-1299	13.604	25.303	21.432
61.	PDM-139	17.300	24.887	22.336
62.	Nayagarh-A	14.127	21.391	20.369
63.	Nayagarh-B	13.425	26.176	21.894
64.	Nayagarh-C	11.214	26.276	18.796
65.	Khadabhanga-B	13.052	26.666	20.936
66.	Raipur Local	12.672	27.131	21.200

Fig. 1: Two dimensional scatter diagram showing relative positions of 66 greengram genotypes based on PCA scores

number of seeds pod[-1], pod length in positive direction and traits such as days to 50% flowering and days to maturity in negative direction could be used as selection criteria in that order to achieve higher productivity.The genotypes HUM-10, PDM-139 and ML-613 were shown to be distinct from other genotypes using principal component analysis, and so can be used in a hybridization program to customize agriculturally essential features and, as a result, increase seed yield in greengram.

REFERENCES

Ali, M. and Kumar, S. 2006. Greengram and urdbean: Retrospect and Prospects. In: *Advances in Greengram and Urdbean (Eds)*IIPR., Kanpur, India, pp. 1-19.

Al-Jibouri, H.A., Miller, P.A. and Rabinson, H.F. 1958. Genotypic and environmental variances and co-variances in upland cotton crosses of interspecific origin. *Agron. J.*, **50**: 633-636.

Baskaran, L., Sundaramoorthy, P., Chidambaram, A. and Ganesh, K.S. 2009. Growth and physiological activity of greengram (*Vigna radiata* L.) under effluent stress. *Int. J. Bot.,***2**(2): 107-114.

Dewey, O.R. and Lu, K.H. 1959. A correlation and path coefficient analysis of components of crested wheat grass seed production. *J. Agron.*, **57**: 515-518.

Jakhar, N.K. and Kumar, A. 2018. Principal component analysis and character association for yield components in greengram [*Vigna radiata* (L.) Wilczek] genotypes. *J. Pharmacognosy and Phytochem.,***7**(2): 3665-3669.

Kaiser, H.F. 1960. The applications of electronic computers to factor analysis. *Edu. and Psychol. Measure*, **20**: 141-151.

Mahalingam, A., Manivanan, N., Bharathi, K.K., Ramakrishnan, P. and Vadivel, K. 2020. Character association and principal component analysis for seed yield and its contributing characters in greengram (*Vigna radiata* (L.) Wilczek). *Electron. J. Plant Breed.,***11**(1): 259-262.

Marappa, N.D., Savithramma and Prabudda, H.R. 2010. Studies on correlation coefficient and path coefficient analysis in mungbean. *Env. and Ecol.*, **28**(2A): 1104-1107.

Mishra, A., Mohanty, S.K, Mishra, S., Samal, K.C. and Das, S. 2014. Character association and genetic diversity in rainfed greengram (*Vigna radiata* (L.) Wilczek). *Indian J. Dryland Agric. Res. Devel.*, **29**(1): 45-51.

Mohon, S., Sheeba, A. and Kalaimagal, T. 2019. Genetic diversity and association study in greengram [*Vigna radiata* (L.) Wilczek]. *Leg. Res.*, **1**: 1-6.

Panigrahi, K.K. and Baisakh, B. 2014. Genetic diversity assessment for yield contributing characters of greengram [*Vigna radiata* (L.) Wilczek] cultivars from Odisha. *Env. Ecol.*, **32**(1A): 294-297.

Patel, S.R., Patel, K.K. and Parmar, H.K. 2014. Genetic variability, correlation and path analysis for seed yield and its components in greengram (*Vigna radiata* (L.) Wilczek). *The Bioscan.*, **9**(4): 1847-1852.

Raturi, A., Singh, S.K., Sharma, V. and Pathak, R. 2015. Genetic variability, heritability, genetic advance and path analysis in mungbean [*Vigna radiata* (L.) Wilczek]. *Leg. Res.*, **38**(2): 157-163.

Roy, A., Hijam, Das, A., Roy S.K., Chakraborty, M., Dey, P.C. and Sahana, N. 2020. Estimation of genetic variation and characters conglomeration for yield ascribing economic traits of rice (*Oryza sativa* L.). *J. of Crop and Weed*, **16**(3): 56-61.

Singh, B., Goswami, A. and Kumar, A. 2017. Biotechnology approaches for quality improvement of vegetables. *Ann. Hort.,***10** (1): 7-14.

Singh, P. and Narayanan, S.S. 2017. Correlation analysis. *In. Biometrical Techniques in Plant Breeding(6ᵗʰ revised edition)*, Kalyani Publishers., India, pp. 80.

Tabasum, A., Saleem, M. and Aziz, I. 2010. Genetic variability, trait association and path analysis of yield and yield components in mungbean. *Pak. J. Bot.*, **42**(6): 3915-3924.

Thippani, S., Eshwari, K.B. and Bhave, M.H.B. 2017. Principal component analysis for yield components in greengram accessions (*Vigna radiata* L.). *Int. J. Pure Appl. Biosci.*, **5**(4): 246-253.