www.ThePharmaJournal.com

The Pharma Innovation

Shilpa Devi
Division of Vegetable Science,
ICAR-IARI, New Delhi, India

Arvind Nagar
KVK, Agriculture University,
Kota, Rajasthan, India

Manish Kumar
College of Horticulture,
YSPUHF, Solan, Himachal
Pradesh, India

Sandeep Kumar
Regional Research Station,
ICAR-IARI, Katrain, Himachal
Pradesh, India

Morphological Characterization of garden pea (*Pisum sativum* L.) Germplasm through regression and principal component analysis

Shilpa Devi, Arvind Nagar, Manish Kumar and Sandeep Kumar

Abstract
Pea (*Pisum sativum* L.) an extremely self pollinated, annual herb of Fabaceae family comprising each wild (*Pisum fulvum* & *Pisum elatius*) and cultivated species (*Pisum abyssinicum*). A glorious supply of dietary protein particularly essential lysine amino acid in its pods. Offseason peas holds nice economic potential however suffer on productivity front. Yield could be a complicated polygenic trait; its improvement is feasible through identification of contributory traits. With the aim to find out yield contributing traits thirty one diverse pea genotypes were morphologically evaluated. Evaluation was done in mid hill conditions of western Himalayan range during Rabi season of 2014-15. Loading of various variables supporting first principle component indicated that node at which first flower appear (NF), Number of pods per plant (NP), Number of seeds per pod (NS) and yield (Y) were the main components of divergence between 31 pea genotypes. Multiple linear regression model indicated that yield per plot might be well predicted with assistance of variables like number of pods per plant (NP), weight of 100pods (WH).

Keywords: Component, fabaceae, Himalayan, offseason, regression

Introduction
Pea a very important member of Fabaceae family placed under genus *Pisum* (2n=2x=14) and it comprise both the wild (*P. fulvum* and *P. elatius*) and cultivated species (*P. abyssinicum*) (Ellis *et al*., 2011) [6]. Mediterranean region is primary centre of its origin with Ethiopia and Near East as secondary centres (Blixt, 1970) [4]. Pea is one among the earliest food crops and is second most vital food legume next to common bean. It's predominantly a self-pollinated, annual herb having climbing or bush type growth habit. Pea holds a prominent place among vegetables due to its high nutritive value, notably proteins and various other health building substances like carbohydrates, vitamin A, vitamin C, calcium and phosphorus (Sharma, 2010) [24]. It contains an honest proportion of essential amino acids particularly lysine (Ghobary, 2010) [11] and is a cheapest source of protein in diet. Productivity per unit area of peas remains low regardless of its high economic importance; which is more prominent among offseason varieties. Peas can fix atmospheric nitrogen through symbiotic association and hence do not require nitrogenous fertilizers (Janzen *et al*. 2014) [15]. Pea fits well under intensive agriculture and livelihood enhancement due to its ability to sustain drought (Janzen *et al*., 2014) [15], and for it short duration cultivars are desirable. In Indian sub-continent pea is found to be grown since ancient times and is utilized for different purposes however with introduction of exotic collections and adoption of improved cultivars, this heritage is greatly eroded (Kumar *et al*., 2018) [18]. A small portion of diversity could preserved in gene banks (Hagenblad *et al*., 2014) [13] or among farmers occupying marginal lands (FAO, 2011) [8] who are practicing family farming (Kumar *et al*., 2018). Description and knowledge of landraces is a prerequisite for their use (Marchenay and Lagarde, 1987) [21]. Several studies conducted using different approaches are published throughout the world (Ali *et al*., 2007, Sarikamis *et al*., 2010, Ghixari *et al*., 2014, Kumar *et al*., 2018) [1, 23, 10, 18].

Yield is quite complicated trait which is highly influenced by environmental conditions and the selection on the basis of observed phenotypic variability is not effective as it depends a lot more on the nature and magnitude of heritable variations. Hence new trends of breeding need to be developed using only a couple of commercial varieties as parents within the breeding programmes again and again has also led to low genetic diversity among pea cultivars (Simioniuc *et al.,* 2002) [25].

Corresponding Author:
Shilpa Devi
Division of Vegetable Science,
ICAR-IARI, New Delhi, India

In this research pea cultivars were evaluated phenotypically for their behaviour w.r.t. early maturity so that these can be further used to develop short duration varieties.

Materials and Methods
The current investigations were carried out at Research Farm of the Department of Vegetable Science, Dr YS Parmar University of Horticulture and Forestry, Nauni, Solan (HP) in *Rabi* 2014-2015. This location lies at 30°50' N latitude, 77°11'30" E line of longitude and is 1260 m above mean sea level and represents the mid-hill zone of Himachal Pradesh. The climate of the Experimental Farm is usually characterised as sub-humid, sub-temperate with cool winters. The annual precipitation is around 1000-1300 mm, most of which is received during South West monsoon months (June-September) in India. The soil structure of the experimental farm is loam to clay loam with pH ranging between 6.85-7.04. 31 pea genotypes including 1 check variety Pb.89 were used Table 1. Each genotype was sown in a Randomized Complete Block Design (RCBD) with three replications each. The quality cultural practices as recommended in the Package of Practices of Vegetable Crops were followed to confirm a healthy crop stand. Observations were recorded on ten randomly selected plants in each replication on days to fifty per cent flowering (DF), node at which first flower appear (NF), number of pods per plant (NP), pod length (PL), pod width (PW), days to marketable maturity (DM), number of seeds or grains per pod (NS), pod yield (kg/plot) (Y), Total Soluble Solids (TSS). For shelling percentage (SP) and 100-pod weight (WP) data was recorded on 25 and 100 randomly selected pods from each genotype in each replication at the time of 2nd harvest.

Table 1: List of pea genotypes and their source of availability

Genotype/variety	Source
Solan Nirog, Pb 89*	UHF, Solan
EC-598655, EC-598615, EC-598628, EC-598677	NBPGR, Regional Research Station, Phagli
AP-2, AP-4	CSAUAT, Kanpur
Pusa Pragati, KTP-101, KTP-102	IARI Regional Research Station, Katrain
14/PEVAR-2, 14/PEVAR-3, 14/PEVAR-4, 14/PEVAR-5, 14/PEVAR-6, 14/PEVAR-7, 12/PEVAR-1, 12/PEVAR-2, 12/PEVAR-3, 12/PEVAR-4, 12/PEVAR-5, 14/PMVAR-1, 14/PMVAR-2, 14/PMVAR-3, 14/PMVAR-4, 14/PMVAR-5, 12/PMVAR-1, 12/PMVAR-2, 12/PMVAR-3, 12/PMVAR-4, 12/PMVAR-5, 12/PMPMVAR-1, 12/PMPMVAR-2, 12/PMPMVAR-3, 12/PMPMVAR-4, 12/PMPMVAR-5, Arka Ajit, VP-233, IP-3, AP-3, VRP-7	IIVR, Varanasi

*Check variety

The mean data values were subjected to analysis of variance and ANOVA was set as described by Gomez and Gomez (1983) [12] for Randomized Complete Block Design. Contribution of different characters towards the divergence was estimated with the help of principle component analysis in accordance with Lawley and Maxwell (1963) [20] and Ramchander *et al.* (1979) [22]. Mathematical approach utilized in PCA is termed eigen analysis solving the eignvalues and eigenvectors of a square symmetric matrix with sums of squares and cross products. The eigenvector associated with largest eigen value has the same direction as the first principal component. The eigenvector related to second largest eigen value determine the direction of second principal component (Esposito *et al.*, 2007) [7]. The sum of eigen values equals the trace of square matrix and the maximum number of eigenvectors equals the number of rows (or columns) of this matrix (Harris 2001) [14]. Multivariate statistical methods have been successfully used to classify quantitative and qualitative variations in many crop species like pea (Bhargava *et al.*, 2007) [3].

Results and Discussion
Genetic variability is the fundamental need for any crop breeding programme. Analysis of variance (Table 2) indicated significant differences among the genotypes for all studied traits. These differences indicated the presence of variability and opportunity for improvement in yield and quality traits of peas.

Table 2: ANOVA for different traits in pea genotypes

Traits	Source of Variation		
	Genotypes (30)ᵃ	Replication(2)	Error (60)
Days to 50% flowering (DF)	131.46*	37.43	10.93
Node at which first flower appear (NF)	4.58*	0.60	0.29
Days to marketable maturity (DM)	59.35*	22.33	14.19
Number of pods per plant (NP)	124.34*	22.91	2.71
Pod Length (PL)	3.33*	0.69	0.15
Pod Width (PW)	0.069*	0.023	0.006
Number of seeds per pod (NS)	2.55*	0.003	0.136
Weight of 100 pods (WP)	22278.49*	12989.20	535.66
Shelling percentage (SP)	432.87*	29.73	17.03
Total soluble solids (TSS)	56.65*	1.18	2.15
Yield (Y)	9.95*	1.98	0.31

*Significant at 5% level of significance
ᵃvalues in parentheses are degree of freedom

Principal Component Analysis
The principal component analysis (PCA) can scale back the dimensionality of a data set consisting of an outsized number of interrelated variables, while retaining maximum amount of the possible variation within the data set. Principal Component Analysis (PCA) involves a mathematical

approach to transform a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components (Chatfield & Collis, 1980) [5]. The primary principal component accounts for as much of the variability in the data as possible and each succeeding component accounts for as much of the remaining variability as possible.

Principal component analysis indicated that the first four components as shown in Table 4 account for the maximum estimated variation of about 74.88%. Factor analysis was applied to extract the fundamental factors underlying the observed traits of peas.

The factors were extracted singly on the basis of eigen values and it revealed the pattern and principal components of the data. The first four components having eigen values greater than 1 were retained in the analysis attributable to the substantial variations reflected in these components.

The factors corresponding to eigen values less than 1 were not considered (Fig 1). These factors were ignored due to "Guttmans lower bound principle" according to which eigen values less than unity ($\lambda < 1$) should be ignored (Kaiser 1958) [11]. The orthogonal factors were extracted. The centroid method of analysis (Lawley and Maxwell, 1963) [20] was used. The first four factors had variances of 3.65, 1.81, 1.59, 1.17 with 33.25, 16.53, 14.46 and 10.63% of total percent variance respectively aggregating to 74.88% of total variation. The first factor extracted had the positive combination of node at which first flower appear (NF), number of pods per plant (NP), Pod length (PL), Number of seeds per pod (NS), Weight of 100 pods (WP) and yield per plot (Y). The second factor extracted had the positive value of shelling percentage (SP).

The third factor extracted had the positive combination of days to fifty percent flowering (DF), Total soluble solids (TSS). The fourth factor extracted had the combination of days to marketable maturity (DM) and Pod width (PW).

The highest positive values of different characters under study in different components indicated its importance in divergence among 31 pea genotypes, whereas negative values showed the lowest contribution to the divergence (Table 4). Loading of different variables based on first principle component indicated that node at which first flower appear (NF), number of pods per plant (NP), Number of seeds per pod (NS) and yield (Y) were the main components of divergence between 31 pea genotypes (Fig 2). These findings were in line with that of Esposito *et al.* (2007) [7], Arif *et al.* (2020) [2].

Table 3: Eigen values obtained of the PCA for 2014-2015 season

Component	Initial Eigen values			Extraction Sum of Squared Loadings		
	Value	% Variance	Cumulative %	Total	% Variance	Cumulative %
1	3.658	33.256	33.256	3.658	33.256	3.658
2	1.818	16.531	49.787	1.818	16.531	49.787
3	1.591	14.460	64.247	1.591	14.460	64.247
4	1.170	10.633	74.880	1.17	10.633	74.880
5	0.828	7.528	82.408			
6	0.645	5.864	88.272			
7	0.458	4.162	92.434			
8	0.339	3.080	95.514			
9	0.310	2.818	98.332			
10	0.119	1.085	99.416			
11	0.064	0.584	100.000			

In the work of Esposito *et al.* (2007) [7] on pea genotypes, the

two first components explained 67.7% of variability in the first season of experiment and 69.8% in the second one. The study conducted by Umar *et al.* (2014) [26] on pea genotypes from different origins showed that the two parameters: Pod length and width are related to the first component which explained 40.29% of variation.

Table 4: Eigenvectors obtained of the PCA for 2014-2015 season

Principal Component*				
Variables	#PC1	PC2	PC3	PC4
Days to 50% flowering (DF)	-0.649	0.383	0.701**	-0.218
Node at which first flower appear (NF)	0.334	-0.015	-0.040	0.071
Days to marketable maturity (DM)	-0.011	0.002	-0.114	0.582
Number of pods per plant (NP)	0.530	-0.118	-0.854	0.301
Pod Length (PL)	0.738	-0.199	-0.478	-0.073
Pod Width (PW)	0.388	-0.536	0.253	0.579
Number of seeds per pod (NS)	0.772	0.340	-0.260	-0.017
Weight of 100 pods (WP)	0.625	-0.453	0.485	0.055
Shelling percentage (SP)	-0.660	0.243	-0.040	-0.386
Total soluble solids (TSS)	0.175	0.004	0.680	-0.430
Yield per plot(Y)	0.835	0.094	0.253	0.344
Eigen Values	3.65	1.81	1.59	1.17
Percentage of variance	33.25	16.53	14.46	10.63
Cumulative % of variance	33.25	49.78	64.24	74.88

*Extracted through principal component analysis
PC: Principal Component
**Bold value indicates the highest Eigen Vector for the corresponding trait amongst the five principal components

Regression Analysis

To estimate yield per plot, multiple linear regression model was applied with yield per plot as dependent variable and days to fifty per cent flowering (DF), node at which first flower appear (NF), number of pods per plant (NP), pod length (PL), pod width (PW), days to marketable maturity (DM), number of seeds (NS) or grains per pod, TSS, shelling percentage (SP) and weight of 100pods (WP) as independent variables. After eliminating the independent variables with p>0.10, a simplified multiple linear regression model was obtained.

The prediction of yield per plot with values in parentheses indicating standard errors of the regression coefficients had been given in the Table 5. Multiple linear regression model indicated that yield per plot could be well predicted with the help of number of pods per plant (NP) and weight of 100pods (WH). Both characters namely number of pods per plant and weight of 100pods had positive effects on estimation of yield per plot (Fig. 3).

The coefficient of determination (R^2) in linear regression was high (0.870) representing 87 per cent of total variation in yield per plot. Yield per plot was influenced by these two characters so linear regressions model is best suited for genetic improvement of pea germplasm.

These findings are in line with those of Kumar *et al.* (2016) [19] who followed linear regression model to predict yield in lettuce crop.

Table 5: Multiple linear regression coefficient model to predict yield per plot

Dependent Variable	Independent variables influencing yield per plot			*R²
	Intercept	Number of pods per plant (NP)	Weight of 100 pods (WP)	
Yield per plot (Y)		0.182 (0.031)	0.013 (0.003)	0.870

*Coefficient of multiple determination
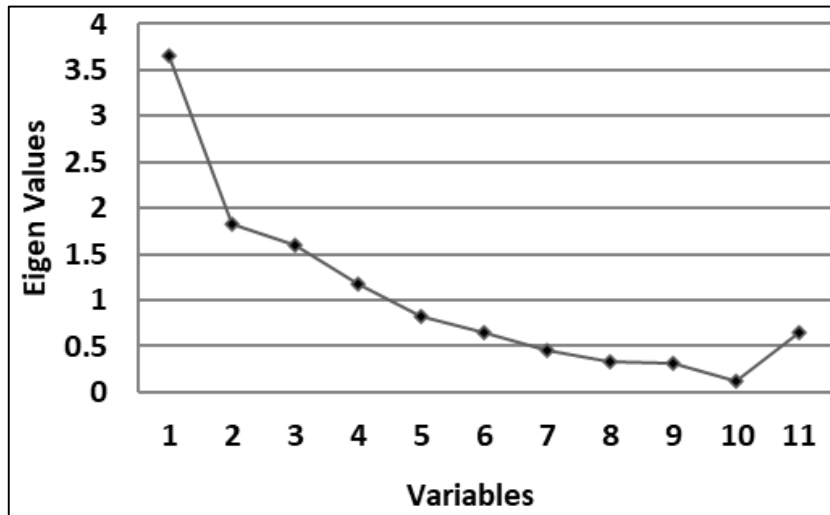
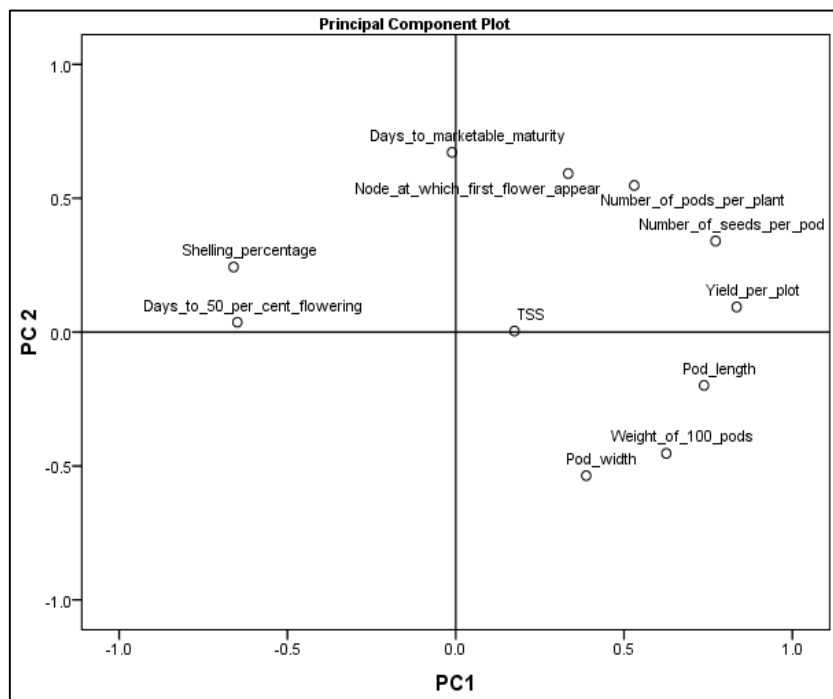**Fig 1:** Plot of Eigen values for 2014-2015 season

Fig 2: Loading of different characters based on first two principle components
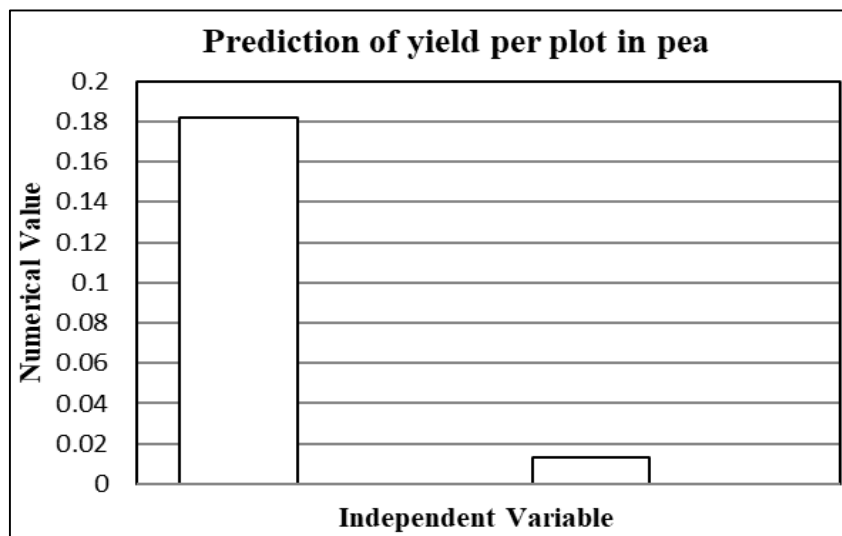

Fig 3: Regression coefficient and standard error of multiple linear regression coefficient model based on independent variables influencing yield per plot in pea

Conclusion

Principal components contributing substantially toward yield improvement revealed a considerable amount of genetic variability among studied genotypes. This variability can be tapped either through direct selection as purelines or further involvement of selected genotypes in hybridization programmes. Regression analysis gave a clear idea of traits that are directly associated with yield improvement. Yield is a complex trait and in case of garden pea it was influenced by two characters *viz* number of pods per plant and weight of 100 pods.

References

1. Ali Z, Qureshi A, Ali W, Gulzar H, Nisar M, Ghafoor A. Evaluation of genetic diversity present in pea (*Pisum sativum* L.) germplasm based on morphological traits, resistance to powdery mildew and molecular characteristics. Pak. J Bot 2007;39(7):2739-2747
2. Arif U, Ahmed MJ, Rabbani MA, Arif AA. Assessment of genetic diversity in pea (*Pisum sativum* L) landraces based on physic-chemical and nutritive quality using cluster and principal component analysis. Pak. J Bot 2020;52(2):575-580.
3. Bhargava A, Shukla S, Rajan S, Ohri D. Genetic diversity for morphological and quality traits in quinoa (*Chenopodium quinoa* Willd) germplasm. Genetic Resources and Crop Evolution 2007;54:167-173.
4. Blixt S. 1970. *Pisum.* In: Frankel OH and Besset E (*eds)*. Genetic resources in plants, their exploration and conservation. Blackwell Scientific Publications, Oxford 1970, 321-325.
5. Chatfield C, Collis A. Introduction to multivariate analysis. CRC Press Boca Raton 1980, 246.
6. Ellis THN, Hofer JI, Timmerman-Vaughan GM, Coyne CJ, Hellens RP. 2011. Mendel 150 years on. Trends Plant Science 2011;16:590-596.
7. Esposito MA, Martin EA, Cravero VP, Cointry E. Characterization of pea accessions by SRAPs markers. Scientia Horticulturae 2007;113:329-335.
8. FAO. Sauvegarder la biodiversity. Commission des ressources génétiques pour l'alimentation et l'agriculture. Organisation des Nations Unis pour l'Alimentation et l'agriculture 2011, 2.
9. FAOSTAT. Food and agriculture organization of the United Nations, statistics division. Available online: http://faostat.fao.org/ 2014.
10. Ghixari B, Vrapi H, Hobdari V. Morphological characterization of pea (Pisum sativum L.) genotypes stored in Albanian genebank. Albanian J Agric. Sci(special edition) 2014.
11. Ghobary HMM. Study of relationship between yield and some yield components in garden pea (*Pisum sativum* L.) by using correlation and path analysis. Journal of Agricultural Research 2010;36:351-360.
12. Gomez KA, Gomez AA. Statistical procedures for Agricultural Research. New York, John Wiley and Sons Inc 1983, 680.
13. Hagenblad J, Bostrom E, Nygards L, Leino MW. Genetic diversity in local cultivars of garden pea (Pisum sativum L.) conserved on farm and in historical collections. Genet. Resour. Crop Evol 2014;2(61):413-422.
14. Harris R. A multivariate analysis. Laurence Girlbaum Associates, New Jersey, USA 2001, 624.
15. Janzen J, Brester G, Smith V. Dry peas: trends in production, J Am. Stat. Assoc 2014;58:236-244.
16. Jolliffe I. Principal component analysis. Springer Series in Statistics (2nd Edn), New York, USA 2002, 487.
17. Kaiser HF. The varimax criteria for analytic rotation in factor analysis. Psychomethka 1958;23:187-200.
18. Kumar M, Jeberson MS, Singh NB, Sharma R, Patel RS. Analysis of trait association and principal component of variability in field pea (*Pisum sativum* L.) genotypes. The Pharma Innovation 2018;7(8):437-441
19. Kumar R, Kaushal S, Kumar S, Kumar D, Shukla YR, Dogra BS. Morphological characterization of newly introduced lettuce (*Lactuca sativa* L.) germplasm through principal component and regression analyses. Electronic Journal of Plant Breeding 2016;7(3):742-749
20. Lawley DN, Maxwell AE. Factor analysis as a statistical method. Butterworth, London 1963, 117.
21. Marchenay P, Lagarde MF. A la recherche de variétiés locales de plantes cultivées. Centre Natl. Rech. Scient 1987, 125.
22. Ramchander PR, Biswas SR, Singh DP, Pathan CS. Factor analysis in onion (*Allium cepa* L.). Curr. Sci 1979;48:137-141.
23. Sarikamis G, Yanmaz R, Ermis S, Bakir M, Yüksel C. Genetic characterization of pea (Pisum sativum L.) gemplasm from Turkey using morphological and SSR markers. Genet. Mol. Res 2010;9(1):591-600.
24. Sharma BB. Combining ability and gene action studies for earliness in garden pea (*Pisum sativum* L.). M Sc Thesis, GBPUAT, Pantnagar 2010, 10.
25. Simioniuc D, Uptmoor R, Friedt W, Ordon F. Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. Plant Breed 2002;121:429-435.
26. Umar HMI, Ur-rehman S, Bilal M, Atif S, Naqvi H, Manzoor SA, *et al*. Evaluation of genetic diversity in pea (Pisum sativum) based on morpho-agronomic characteristics for yield and yield associated traits. J Biodivers. Environ. Sci 2014;4(5):323-328.