INT. J. BIOL. BIOTECH., 13 (3): 351-358, 2016.

ASSESSMENT OF GENETIC DIVERSITY IN EXOTIC GERMPLASM OF WHEAT (*TRITICUM AESTIVUM* L.) BY USING PCA AND CLUSTER ANALYSIS

Shazia Parveen[1], Tasveer Zahra Bokhari[1], Muhammad Faheem Siddiqui[3], Manzoor Hussain[2]and Faheem Iqbal[2]

[1]*Institute of Pure and Applied Biology, Bahauddin Zakariya University, Multan, Pakistan.*
[2]*Regional Agricultural Research Institute, Bahawalpur, Pakistan.*
[3]*Department of Botany, University of Karachi, Karachi 75270, Pakistan.*

ABSTRACT

The objective of study was to assess genetic diversity in exotic germplasm (8th EBWYT) of wheat (*Triticum aestivum* L.) based on cluster analysis and principle component analysis (PCA). There was a significant relationship among the studied traits i.e. days to heating, plant height, spike length, canopy temperature and grain yield per plant. A positive significant correlation was observed between spike length and number of spikelet per spike while a negative significant correlation was assessed between canopy temperature and thousand kernel weight. Out of seven, three PCA's axes inhibited more than one Eigen value but the level of dissimilarity was high which indicated that the germplasm was with broad genetic base. Cluster analysis expressed high level of diversity because two main groups were observed, one with 14 and the other with 16 genotypes.

Keywords: Exotic germplasm, Genetic diversity, PCA, Cluster analysis, *Triticum aestivum*.

INTRODUCTION

Wheat (*Triticum aestivum* L.) is the most important food crop of the world which is being cultivated on 217.2 Million hectare area with 651 million tons production and average yield of 2906 kg / ha (Anonymous, 2010). With regard to its production and utilization it is one of the most important sources of sustenance for mankind since its domestication from 15000 to 10000 BC. Wheat occupies a primary position in daily food consumption because it is the cheapest source of calories and protein (Maqbool *et al.,* 2010). The contribution of wheat among total cereal intake is 84% in Pakistan. Genetic diversity of plants enhances their potential efficiency for breeding which may result in enhanced food production (Mostafa *et al.*, 2011). The necessity of performing breeding experiment for the production of resistant plant verities under different condition showed prolonged food production (Martin *et al.*, 2008). The development of new varieties with desirable genetic makeup takes great efforts by breeders to overcome the consumption pressure of ever increasing population by improving the yield potential of wheat (Memon *et al.*, 2007). To reconstruct the ideotypes of plant, available genetic resources are utilizing by wheat breeders to meet the ever increasing requirements of the population (Memon *et al.*, 2007). A measure of the genetic relationship between parents and progeny is determined by heritability value (Memon *et al.*, 2007). A lot of research work has been done to insert the desirable genes in present wheat varieties to enhance the crop productivity. The insertion of semi dwarfing genes Rht1 Rht2 are associated with development of high yielding varieties with grain yield associated traits (Sial *et al.*, 2010). Ajmal *et al.*, (2009) estimated genetic parameters and characterization associated in wheat while genetic heritability for grain yield and its related characters in spring wheat (*Triticum aestivum* L.) was studied by Shabana *et al.*, (2007). Eivazi *et al.*, (2007) suggested that identification of genetic diversity analyzed with the help of some appropriate methods i.e. PCA, factor analysis and cluster analysis. Shah *et al.*, (1999) performed genetic analysis of agronomic traits controlled by wheat chromosome 3A.

The present research was designed to sort out better performing genotypes of wheat against WG-99 (a race of stem rust fungus), under local conditions (Bahawalpur-Pakistan) and to observe diversity in valuable CIMMYT germplasm.

MATERIALS AND METHODS

A CIMMYT nursery of "8th Elite bread wheat yield trail" was selected for present study. This valuable germplasm is resistant to stem rust race i.e. WG-99. This trail was consisted of 30 wheat genotypes, sown at Regional Agricultural Research Institute (RARI), Bahawalpur- Pakistan, for performance under local environment and genetic diversity in the germplasm. Plot size was kept as 5m in length and 1.8 m in width with 6 rows. Distance between rows was 30cm. Randomized Complete Block Design (RCBD) was used with three replicates. Sowing was

done on mid November 2013-2014 cropping season. The experiment was conducted by using local recorded management practices.

During the growing season, following traits were measured by random selection and tagging 10 plants from each experimental unit. Plant height of the main tiller of each selected plant was measured in centimeters from the ground level to tip of spike excluding awn. At maturity the spike length of main tillers of nominated plant measured in centimeter, from the base to the tip of the spike excluding awn. Number of spikelet per spike, days to heading, 1000 grain weight, CTD (C), grain yield per plot (kg/ha) were measured.

Analysis of variance was done by using statistical formula given by Steel and Torrie (1980). Other statistical techniques like Principal Component Analysis (PCA) and Cluster Analysis were also used in the present study.

RESULTS AND DISCUSSION

Elite bread wheat yield trial (8^{th} EBWYT) was planted along with 3 reps and 30 entries in each. Mean grain yield was recorded about 4080 kg/hectare with a maximum of 4720.7 kg/hectare. Mean Thousand kernel weight was 36.7 grams with a maximum of 40.33 grams (Table 1).

Table 1. Descriptive statistics of studied traits in 30 genotypes of the 8^{th} EBWYT.

Sr. No.	Description	CT	DH	GY	NS	PLH	SL	TKW
1	Mean	25.575	93.701	4080	9.4893	99.823	11.622	36.7
2	SD	1.2551	2.4525	316.55	0.8566	8.5743	0.7086	2.1705
3	Variance	1.5752	6.0149	100202	0.7338	73.519	0.5021	4.7109
4	SE Mean	0.2291	0.4478	57.793	0.1564	1.5654	0.1294	0.3963
5	CV	4.9075	2.6174	7.7584	9.0269	8.5895	6.0971	5.9141
6	Minimum	22.73	90.33	3563	8.33	86.67	10.17	32.67
7	Maximum	27.27	99.67	4720.7	13	131	12.5	40.33

Key to abbreviations: CT: Canopy Temperature ºC, DH: Days to heading, GY: Grain Yield (Kg/Hectare), NS: Number of Spikelet per spike, PLH: Plant Height (cm), SL: Spike length (cm), TKW: Thousand kernel weight (g), SD = standard deviation, SE = standard error and CV = cumulative variance.

Analysis of variance of all the studied traits remained non-significant among the replicates (Table 2). By taking genotypes as factor of variability, all the studied traits found with significant variation (Table 3). This indicates possible diversity in genotypes.

Table 2.Analysis of Variance (ANOVA) of agronomic traits showing non significance among replicates.

	Canopy Temperature				Days to Heading		
SOV	DF	MS	F value	P	MS	F value	P
Reps.	2	5.22	1.75	0.1792	3.03	0.43	0.6497
Error	87	2.97			6.99		
	Grain Yield (Kg/hectare)				Number of Spikelet per spike		
Reps.	2	21014	0.14	0.8707	0.544	0.33	0.722
Error	87	151510			1.6712		
	Plant Height (cm)				Spike length (cm)		
Reps.	2	126.544	0.8	0.45	1.17	1.18	0.312
Error	87	157.219			0.99		
	Thousand kernel weight (g)						
Reps.	2	1.3	0.22	0.8			
Error	87	5.911					

Key to abbreviations: Reps. = Replicates, dF = degree of freedom, MS = mean square, P = Probability

Correlation coefficient assessed significant between spike length and number of spikelet per spike. Similar result was reported by Ilker *et al.* (2010). Thousand kernel weights were found negatively significant with canopy

temperature indicating that high canopy temperature can shrivel the seed up to ample decrease in weight. Highly significant positive correlation was found between grain yield and thousand kernel weight (Table 4). Akram *et al.,* (2008) reported that grain yield was positively correlated with thousand grain weight. Similar results were also reported by Dokuyueu and Akkaya (1999). Yucel *et al.,* (2009) shows a positive and significant correlation between spikelet number per spike, grain yield per spike, grain number per spike and spike length.

Table 3.Analysis of Variance (ANOVA) of agronomic traits among 30 genotypes.

	Canopy Temperature				Days to Heading		
SOV	DF	MS	F	P	MS	F	P
Genotypes	29	4.72	2.14	0.006648	18	11.8	0
Error	60	2.21			1.5		
	Grain Yield (Kg/hector)				Number of Spikelet per spike		
Genotypes	29	300604.5	4	0.000003	2.201	1.597	0.063375
Error	60	75097.44			1.378		
	Plant Height (cm)				Spike length (cm)		
Genotypes	29	220.6	1.756	0.033201	1.51	2.02	0.010822
Error	60	125.6			0.75		
		Thousand kernel weight (g)					
Genotypes	29	14.1	7.96	0			
Error	60	1.8					

Key to abbreviations: See Table 2

Table 4. Pearson's correlation coefficient (r) among studied traits.

Variables	DH	PLH	SL	NS	TKW	CT
PLH	0.171					
SL	-0.018	-0.002				
NS	-0.072	0.109	0.661			
TKW	0.307	0.034	0.148	0.063		
CT	-0.189	-0.177	0.071	-0.062	-0.441	
GY	0.126	0.180	-0.141	-0.090	0.525	-0.903

Values in bold are different from 0 with a significance level alpha=0.05;
Key to abbreviations: See Table 1

Principal Component Analysis, a technique to show transformation of large number of variables into correlated small number of components was carried out. Factors (F_1 and F_2) explained the variability of about 58.83% with the Eigen value 2.413 and 1.705 respectively. Seven factors explained 100 percent variability of the whole parameters estimated (Table 5).

Table 5. Principal Component Analysis.

	F1	F2	F3	F4	F5	F6	F7
Eigen value	2.413	1.705	1.039	0.957	0.511	0.306	0.069
Variability (%)	34.470	24.363	14.842	13.670	7.301	4.373	0.980
Cumulative %	34.470	58.833	73.675	87.346	94.647	99.020	100.000

Scree plot (Fig. 1) among Eigen values, Factors and cumulative variability indicating effective factors showing maximum variability. In Fig. 2, spike length and number of spikelet per spike has been shown with same pattern of variation explained by F_2. Plant height, thousand kernel weight, days to heading and yield is briefly explained with respect to their variation with F_1. Canopy temperature is negative to all of the components.

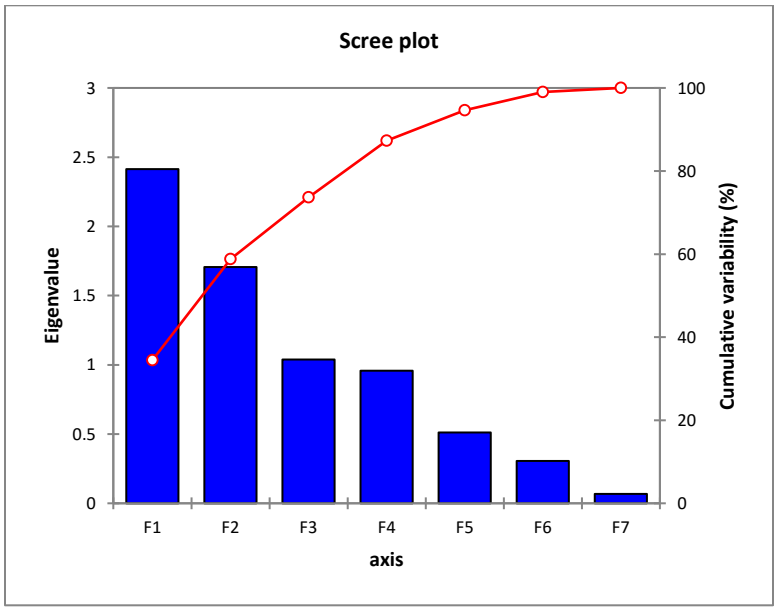


Fig. 1.Scree plot indicting factors distribution along with Eigen values.Cumulative variability (Red line) indicating F1 and F2 as the most effective components of variability.

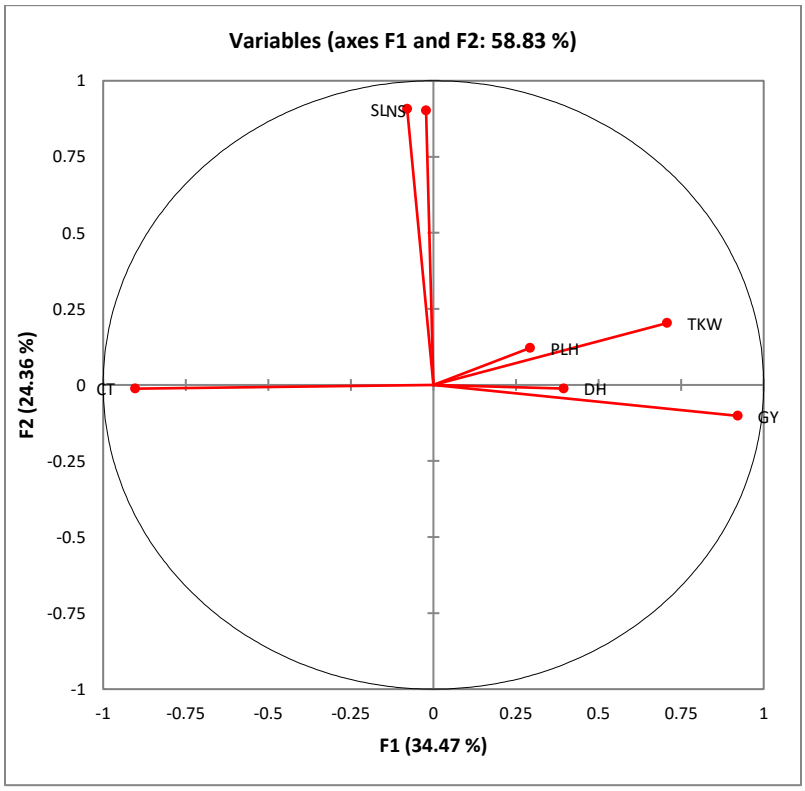


Fig. 2. Percent variability (58.83) explained by F1 (X-axis) and F2 (Y-axis).

Table 6, showed percent contribution of each studied trait toward the appropriate factor. Few genotypes such as 21, 30, 16 and 29 showed high diversity as compared to the others (Fig.3).

Table 6.Contribution of the variables (%).

	F1	F2	F3	F4	F5	F6	F7
DH	6.442	0.008	44.428	30.530	17.608	0.011	0.974
PLH	3.560	0.870	44.350	39.967	10.767	0.279	0.207
SL	0.262	48.297	0.226	1.610	0.309	49.198	0.098
NS	0.021	47.784	0.181	2.256	7.160	40.125	2.473
TKW	20.755	2.426	0.483	18.928	48.719	7.048	1.641
CT	33.762	0.009	4.308	3.640	14.674	0.298	43.309
GY	35.198	0.606	6.024	3.070	0.763	3.040	51.299

Key to abbreviations: See Table 1

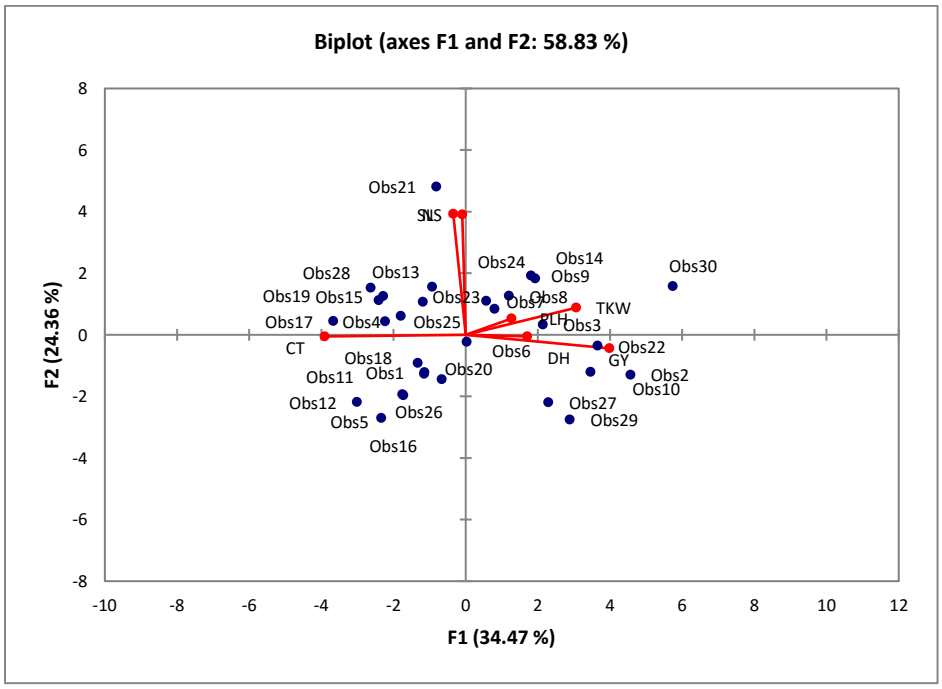


Fig. 3.Biplot of genotypes and the studied traits, showing the variation explained by F1 and F2.

In Fig. 4, structural equation of multiple regression model is explained. Grain yield as a dependent Y-variable, along with error variance of 12798.02 depicts relationship with all other studied traits taken as X-variables.

As long as diversity among the genotypes is concerned, a profile shows the pattern of variation (Fig. 5) and the distance among similarities of 30 genotypes is shown in Fig. 6.
Cluster analysis also expresses high level of diversity because two main groups have been observed (Fig. 6.) i.e. the one with 14 and the other with 16 genotypes.

Histogram of the data of 30 genotypes shows that thousand kernel weights above 40 is concerned with only three genotypes (Fig. 7). Grain yield is not normally distributed as in Fig. 8, hence above 4600 kg ha^{-1} is found in two genotypes only.

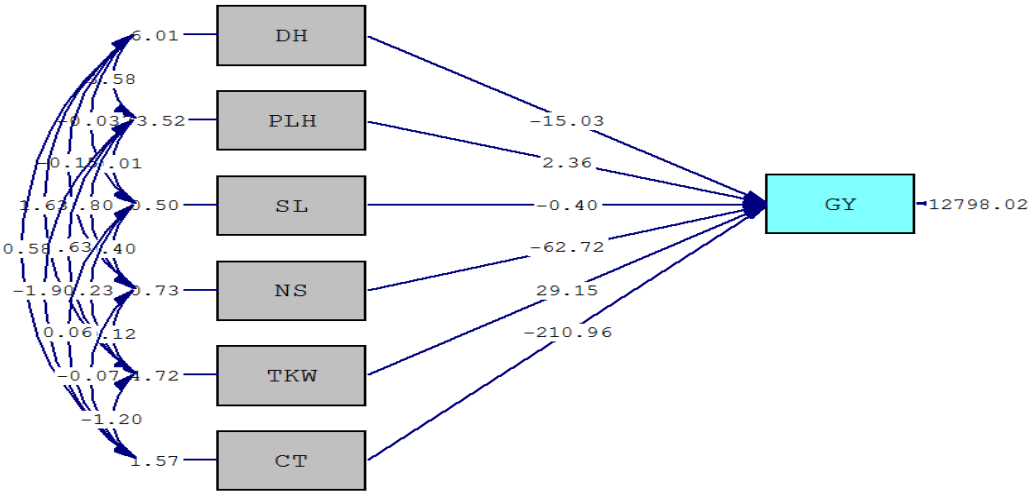


Fig. 4. Structural equation of multiple regression model showing Grain Yield (GY) asY-variable being influenced by multiple X variables (DH, PLH, SL, NS, TKW, CT).

Note: GY = - 15.034*DH + 2.357*PLH - 0.399*SL - 62.724*NS + 29.152*TKW - 210.964*CTError var. = 12798.022, R^2 = 0.872

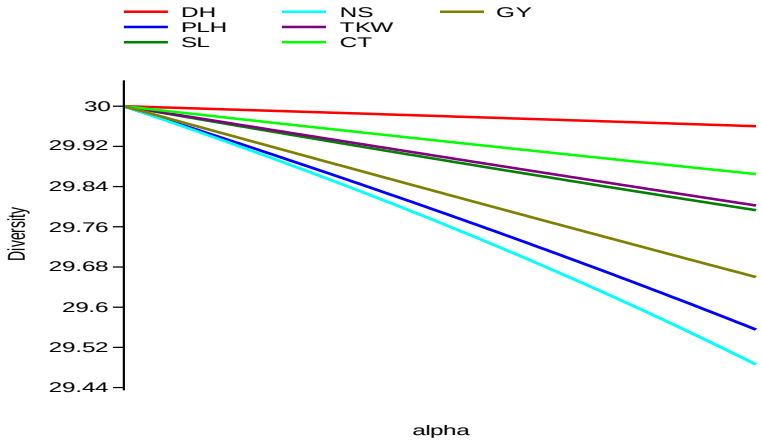


Fig. 5. Depending upon the studied traits a diversity profile of 30 genotypes made by using a software PAST

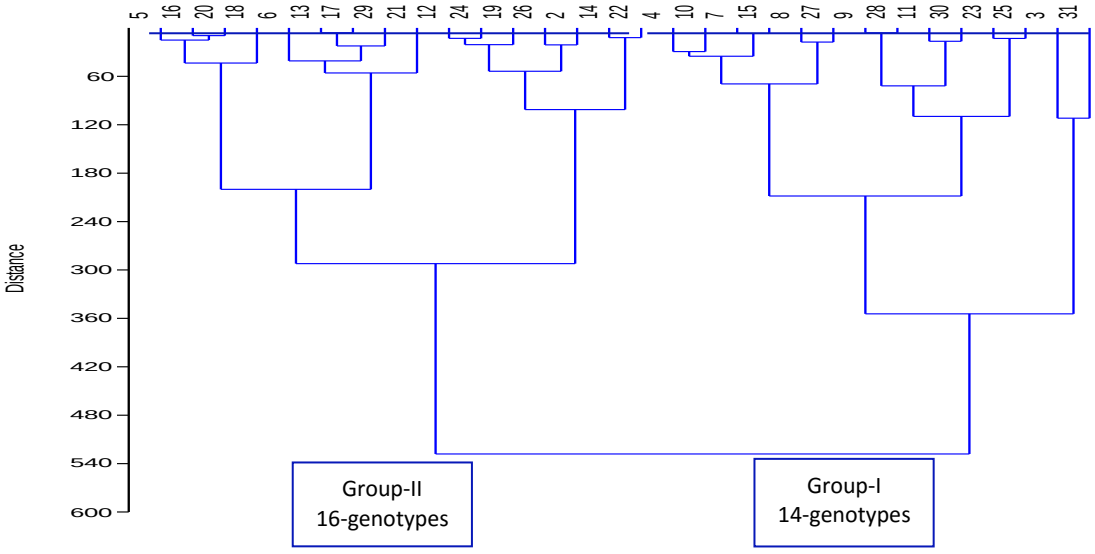
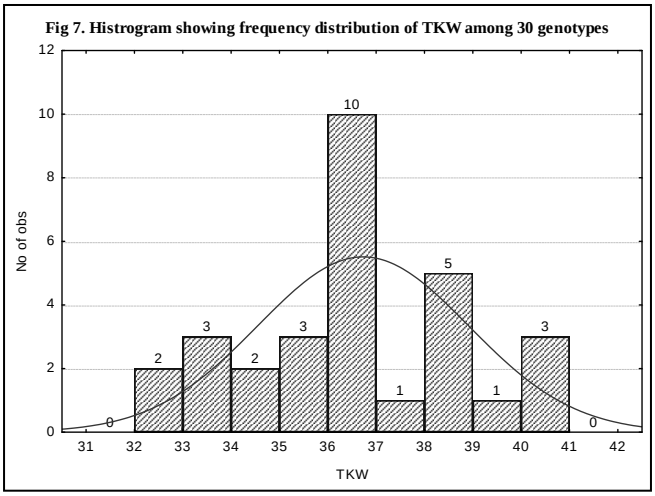
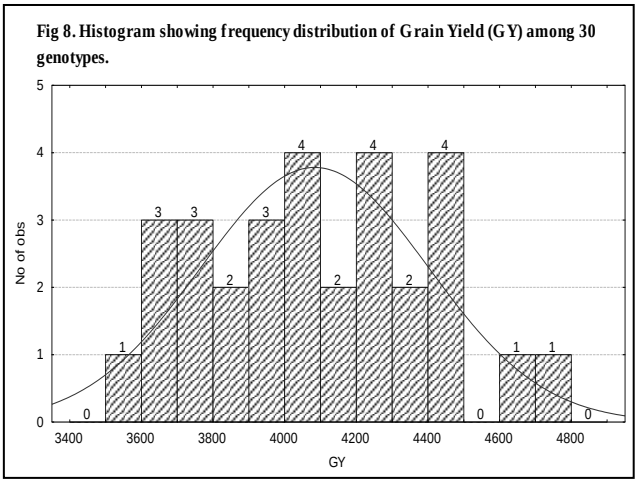


Fig. 6. Cluster Analysis showing Genetic Diversity among 30 Genotypes





The Least Significant Difference (LSD)showing significant groups among genotypes as genotype number 02, 10 and 30 is depicted in (Table 7) and also for thousand kernel weight genotype 14, 29 and 30 expressed significant.

Table 7. Least significant difference (LSD) of studied traits among 30 genotypes.

Gen.	DH Mean		GY Mean		PLH Mean		SL Mean		TKW Mean		CT Mean	
1	90.333	K	4048	CDEFGHI	90.667	CDE	11.333	ABCDE	38.333	ABCDE	26.633	ABC
2	92.667	FGHIJ	4720.7	A	103.67	BCDE	11	BCDE	38.333	ABCDE	22.733	G
3	95	CDE	4243.7	BCDEF	131	A	11.5	ABCDE	35.667	FGHI	24.633	BCDEFG
4	97.667	AB	3613.3	IJ	90	CDE	11.833	ABC	36.667	DEFG	27.1	A
5	93	EFGHI	3753.3	HIJ	104	BCDE	10.667	CDE	37	CDEF	27	AB
6	94.333	DEF	4212.7	BCDEFG	101.67	BCDE	11.667	ABCD	35.667	FGHI	25.833	ABCDE
7	93	EFGHI	4157.3	CDEFGH	86.667	E	12.333	AB	38.667	ABCD	25.1	ABCDEFG
8	91.667	HIJK	4417	ABCD	89.667	CDE	12.333	AB	37	CDEF	24.367	CDEFG
9	93.667	DEFGH	4244	BCDEF	102	BCDE	12.5	A	39.333	AB	24.9	ABCDEFG
10	94	DEFG	4494.3	ABC	101.67	BCDE	11	BCDE	39	ABC	23.833	EFG
11	92	GHIJK	4002.7	DEFGHIJ	105.33	BCD	11	BCDE	32.667	J	25.8	ABCDE
12	91	IJK	3785	GHIJ	102.33	BCDE	10.667	CDE	34.667	GHIJ	27.267	A
13	95	CDE	3921	EFGHIJ	103	BCDE	12.333	AB	36.667	DEFG	26.5	ABC
14	92.667	FGHIJ	4217	BCDEFG	103	BCDE	12.333	AB	40.333	A	24.933	ABCDEFG
15	95.667	BCD	3601.7	IJ	103	BCDE	12.333	AB	36.333	EFGH	27.067	A
16	95.667	BCD	3788.3	GHIJ	101.33	BCDE	10.667	CDE	33.333	J	26.767	ABC
17	95	CDE	3563	J	90.667	CDE	12	ABC	34	IJ	27.067	A
18	93	EFGHI	4000	DEFGHIJ	87.333	DE	11.333	ABCDE	34.333	HIJ	25.3	ABCDEF
19	94	DEFG	3602	IJ	93.667	BCDE	12	ABC	37	CDEF	26.767	ABC
20	92.333	FGHIJK	3839	FGHIJ	96.667	BCDE	11	BCDE	37	CDEF	25.967	ABCDE
21	92.333	FGHIJK	3931.7	EFGHIJ	106.67	BC	12.333	AB	36.333	EFGH	25.7	ABCDE
22	99.667	A	4346.7	ABCDE	101.33	BCDE	11.667	ABCD	38.667	ABCD	23.967	DEFG
23	90.333	K	4014.7	DEFGHI	104.67	BCDE	12.5	A	36.667	DEFG	26.267	ABCD
24	92.667	FGHIJ	4337	ABCDE	96.667	BCDE	12.167	AB	37.333	BCDEF	25.333	ABCDEF
25	91	IJK	4068	CDEFGH	93	BCDE	12	ABC	33	J	25.3	ABCDEF
26	93.333	EFGH	4168.7	BCDEFGH	99	BCDE	11.167	ABCDE	33.667	IJ	25.433	ABCDEF
27	92.667	FGHIJ	4414	ABCD	94.333	BCDE	10.167	E	37	CDEF	23.667	EFG
28	90.667	JK	3808.3	FGHIJ	104	BCDE	12.333	AB	35.667	FGHI	27	AB
29	97	BC	4479	ABC	97	BCDE	10.333	DE	40.333	A	25.733	ABCDE
30	99.667	A	4609.3	AB	110.67	B	12.167	AB	40.333	A	23.267	FG
LSD Value	**2.0224**		**447.57**		**18.302**		**1.4118**		**2.1776**		**2.4283**	

REFERENCES

Ajmal S. U., N. Zakir andM.Y. Mujahid (2009). Estimation of Genetic Parameters and Characterization Association in Wheat. *J. Agric. Biol. Sci.,* 1 (1):15-18.

Akram Z., S. Ajmal, M. Munir(2008). Estimation of correlation coefficient among some yield parameters of wheat under rain-fed conditions. *Pak. J. Bot.,*40 (4): 1777-1781.

Anonymous (2010). World Wheat Crop to Be Third Largest Ever." Farmers Weekly. 152.13: 134. Academic Search Premier. Web. 13 Mar. 2013.

Dokuyueu T. and A.Akkaya (1999). Path coefficient analysis and correlation of grain yield and yield components of wheat genotypes. *RACHIS*. 18: 17-20.

Eivazi A.R., M.R.Naghavi, M. Hajheidari, *et al.,*(2007). Assessing Wheat (*Triticumaestivum* L.) Genetic diversity using quality traits, amplified fragments length polymorphism, simple sequence repeats and proteome analysis. *Ann Appl Boil.,* 152:81-91.

Ilker, E., F. A. Tonk and M. Tosun (2010). Heterosis for yield and its components in bread wheat crosses among powdery mildew resistant and susceptible genotypes. *Pak. J. Bot.,* 42(1): 513-522.

Maqbool, R., M. Sajjad, I. Khaliq, Aziz-ur-Rehman, A.S. Khan, S.H. Khan(2010). Morphological diversity and traits association in bread wheat (*Triticum aestivum* L.). *American-Eurasian J. Agric. Environ. Sci.,* 8: 216–224.

Martin E., V. Cravero, A. Esposito, F. Lopez, L. Milanesi and E.Cointry(2008). Identification of markers linked to agronomic traits in globe artichoke. *Aust. J Crop Sci*., 1:43-46.

Memon, S., M. D. Qureshi, B. A. Ansari and M. A. Sial (2007). Genetic heritability for grain yield and its related characters in spring wheat (*Triticum aestivum* L.). *Pak. J. Bot.,* 39(5): 1503-1509.

Mostafa K., H.F. Mohammad and M.Mohammad(2011). Genetic diversity of wheat genotypes based on cluster and principal component analyses for breeding strategies. *Aus. J. Crop Sci.,* 5(1): 17-24.

Shabana M., D.Q. Mueen, A.A. Bashir and A.S. Mahboob(2007). Genetic heritability for grain yield and its related characters in spring wheat (*Triticum aestivum* L.). *Pak. J. Bot.,* 39(5): 1503-1509.

Shah M. M., P.S. Baenziger, Y. Yen, K.S. Gill, B.M. Silva and K.Halilogu (1999). Genetic analysis of agronomic traits controlled by wheat chromosome 3A. *Crop Sci.,* 39:96-102.

Sial M.A., M.U. Dohat, K.A. Laghari, M.A. Arain, S.M. Mangrio and A.J. Pirzada(2010). Agronomic Performance of Semi-dwarf and dwarf Wheat genotypes. Nuclear Institute of Agriculture, Tandojam. *World Appl. Sci. J.,* 8: 30-33.

Steel, R.G.D. and J.H. Torrie (1980). *Principles and Procedures of Statistics.* McG raw Hill Book Inc., New York.

Yucel, C., F.S. Baloch, and H. Ozkan (2009). Genetic analysis of some physical of bread wheat grain (*Triticum aestivum* L.). *Turk. J. Agric. For.,* 33: 525-535.

(*Accepted for publication April 2016*)