INT. J. BIOL. BIOTECH., 17 (4): 693-699, 2020.

GENETIC VARIATION AND ASSOCIATION AMONG UPLAND COTTON GENOTYPES UNDER SEMI ARID CONDITIONS

Muhammad Husnain Bhatti[1], Muhammad Irfan Yousaf[1*], Muhammad Munir[2], Muhammad Naeem Khan[3], Dilbar Hussain[1], Waseem Akbar[1], Muhammad Asad Hafeez[1], Shoaib Anwar Kohli[4], Muhammad Umer Khalid[5] and Muhammad Abdullah[6]

[1]*Maize and Millets Research Institute (MMRI), Yusafwala, Sahiwal. husnaincrss@gmail.com, irfanpbg.uaf@gmail.com, dilbarhussainmmri@gmail.com , wahussaini@yahoo.com ,asadhafeezpbg@gmail.com,*
[2]*Entomological Research Institute, Ayub Agricultural Research Institute, Faisalabad 38000, Punjab Pakistan. mmunir.sri@gmail.com.*
[3]*Agronomic Research Institute, Ayub Agricultural Research Institute, Faisalabad 38000. neaamfibercrops@gmail.com*
[4]*Fodder Research Institute (FRI), Sargodha, shoaibkohli2@gmail.com*
[5]*Department of Plant Breeding and Genetics, University of Sargodha.omercheema36@yahoo.com*
[6]*Wheat Research Institute AARI, Faisalabad, abdmsc@yahoo.com*
***Corresponding author's email:** irfanpbg.uaf@gmail.com

ABSTRACT

Assessment of available germplasms is the most vital step in developing a crop variety or improving an existing one for a particular character. In the current research, fifteen upland cotton genotypes were evaluated for their genetic potential under semi-arid conditions through two different multivariate analyses i.e. Principal component analysis (PCA) and Agglomerative *hierarchical* clustering (AHC) to characterize cotton genotypes on the basis of their performance. A significant variation was observed among all the investigated attributes. The correlation matrix revealed the positive association of sympodial branches per plant (SBP), number of bolls per plant (BPP) and ginning out turn (GOT) with seed cotton yield (SCY), whereas staple length (S/L) revealed positive interaction with ginning out turn (GOT) but indicated negative relationship with other yield related traits. First four PCs contributed eighty-five (85%) variation and hence emphasized the role of plant height (PH), days to first flower (DFF), sympodial branches per plant (SBP) and monopodial branches per plant (MBP) for maximum variation. In PC_1 genotypes possess maximum yield relating traits, LCA revealed two groups of genotypes, findings of PCA agreed with that of LCA. Information obtained through correlation, PCA and LCA will be used to devise breeding strategy for development of genotypes having batter yield and lint quality.

Key words: Upland Cotton, Genetic distance, PCA, Correlation, Genetic variability, Biplot analysis

INTRODUCTION

Industrial horizon of Pakistan is dominated by textile industry and its affiliated sectors and cotton crop holds central position to this huge industry. Importance of cotton can better be understood with the fact that Pakistan is the largest exporter of yarn, third major consumer and ranked 4th in production. It provides livelihood to millions of farmers and provide work to associated labor market, so enhanced production, with better pricing and improved quality can help to reduce the poverty in cotton belt of Pakistan. At present it covers an area of more than 2.37 million hectares with a production of 9.86 million bales and accounts for 0.8 % to GDP and 4.5 percent to value addition (Anonymous, 2018-19). A gradual decrease in cotton production in Pakistan is being observed from last few years, from 13.96 million bales in 2014-15 to 9.86 million bales in 2018-19%. This is mainly due to decrease in the area of cotton cultivation, which is occupied by the maize and sugar cane crops. Furthermore, lint yield per hectare has also decreased from 802 kg ha^{-1} in 2014-15 to 707 kg ha^{-1} in 2018-19. The key factors responsible for the reduction in lint yield are high temperature coupled with drought stress, poor crop management, high input rates, attack of different insects/pest like whitefly and pink bollworm etc., diseases (cotton leaf curl virus, bacterial leaf blight and root rot etc.). This situation could be overcome by developing highly productive, climate smart cotton genotypes. The key component in the development of such genotypes is the exploration of existing cultivated and wild germplasm as genetic variation provides the foundation for any breeding program.

Existing domesticated cotton genotypes possessed a narrow genetic base (Shakeel *et al.*, 2015). The genetic divergence among available germplasms is prerequisite to broad the genetic base through hybrid breeding program. Researchers emphasized the need to evaluate the genetic variation among cotton genotypes for lint yield and associated traits (Ahmad *et. al*., 2012). Genetic variations among various cotton genotypes for yield and fibre quality traits were examined for their improvement (Baloch *et al.*, 2014 and Bhatti *et al.*, 2020). The degree of genetic variation can be assessed using different biometrical techniques especially multivariate analysis-based techniques such as principal component analysis (PCA), agglomerative hierarchical clustering (AHC) and Principal coordinate analysis (PCoA). Among above stated biometrical procedures, principal component analysis and cluster analysis (LCA) was adapted by the researchers to classify and characterize genotypes on the basis of their similarities/dissimilarities in different crop species like cotton, maize and sorghum etc. (Yousaf *et al*., 2017; Yousaf *et al*., 2018 and Saeed *et al*., 2018). In view of the above information, the study was planned to characterize cotton genotypes on the basis of their lint yield and related traits by using two multivariate approaches viz., PCA and LCA. Principal component analysis and agglomerative hierarchical clustering will enable us to select parent genotypes that could be used in a breeding program to develop highly productive cotton genotypes. Therefore, the purpose of current study was to determine the similarities and differences among genotypes for various quality and yield traits in upland cotton.

MATERIALS AND METHODS

The current experiment was executed at the farm of Cotton Research Sub-Station Raiwind during kharif-2019. Fifteen cotton genotypes were planted with inter-row distance of 2.5 feet while 1 foot within row, thinning was done to accomplish the equal spacing between plants (Table 1). Randomized complete block design (RCBD) was used to lay-out the experiment in three replications. All standard agronomic practices were performed uniformly to the experiment and data was recorded by selecting ten fully guarded, representative plants in each replication of each genotype**.** Data was recorded for the yield related attributes and fiber quality traits viz., plant height (cm), days to first flower, monopodial branches per plant, sympodial branches per plant, bolls per plant, ginning out turn (%), seed cotton yield (kg ha^{-1}), staple length (mm), fiber fineness (µg/inch) and fiber strength (tppsi). The data were subjected to analysis of variance (ANOVA) according to Steel *et al.* (1997) to calculate the statistical significance among cotton genotypes. Data was further subjected to principal components analysis (PCA) and agglomerative hierarchical clustering (AHC) to access the magnitude of variance related to each component contribution to the total variation and to measure the genetic distance among genotypes as devised by Sneath and Sokal (1973). XLSTAT-2016 software was employed to statistically analyze the data and Microsoft Excel (v2019) was used for graphical explanation.

RESULTS AND DISCUSSION

Analysis of variance delineated that all traits studied have significant means squares values among genotypes (Table 2), thus indicating the presence of variation in the plant material, Sufficient variation among genotypes was also discussed by many researchers for the yield and quality traits in upland cotton, (Baloch *et al.*, 2014).

Mean performance of genotypes

The mean values indicated that maximum height (cm) was achieved by the cultivar FH-142 (133 cm) while minimum height was expressed by IUB-222 (101 cm) (Table 3). With reference to days taken to first flower MNH-886 and FH-142 showed the minimum range (52 days) whereas NIAB BT1 and CIM-599 revealed the maximum number (64 days). With regarding to Monopodial branches per plant data range varied between 1.1 (IUB-222) to 2.0 (VH-282) and for Sympodial branches per plant NIAB BT-1 revealed the minimum Sympodial branches (15) whereas cultivar FH-142 expresses maximum number (19) for this trait. The range for number of bolls $plant^{-1}$ varied between 20 and 28 for cultivars CIM 602 and MNH-886 respectively. Ginning Out Turn (G.O.T%) scaled between 35% to 42% among genotypes and VH-282 and NIAB BT-1 fall for minimum value while FH-142 and AA-904 Exhibited the maximum range for this attribute. As for as staple length is concerned AA-904 showed maximum Fiber length (30mm) whereas NIAB BT-1 scored minimum (27 mm) among genotypes. Fiber Fineness graded from 4.5 to 5.5 (µg/inch) and cultivar IUB-222 fall for minimum range and FH-172 expressed maximum value (5.5). Fiber strength graded from 93.30 to 104.00 (tppsi) whereas genotype FH-142 is a top scorer and NIAB BT-1 exhibited minimum strength 93.30 (tppsi.).

Correlation among traits

Correlation matrix obtained through primary statistics indicated important relationships between studied characters. Plant height (PH) displayed negative association with days to first flower (DFF) and for all the fiber quality traits but showed positive correlation with Monopodial (MBP) and Sympodial branches (SBP) per plant (Table 4). Salahuddin *et al* (2010) reported the negative interconnection of plant height with seed cotton yield, but narrated positive association with earliness. Days to first flower (DFF) attained its importance as an indicator for early maturing genotypes, Baloch and Baloch (2004) also delineated that the first flower can be used as prime indicator for early maturing genotypes, in correlation matrix it revealed affirmative significant value with Monopodial branches per plant and Fiber Fineness, but displayed negative association to all other traits.

Sympodial branches per plant (SBP) exhibited significant positive interaction with bolls per plant (BPP), Ginning out turn (G.O.T%) and Fiber Strength (STR), whereas Monopodial branches per plant displayed positive significant association with Days to first flower (DFF) but revealed negative interconnection with rest of the traits. Findings of Latif *et al.* (2015) described the positive interdependence of bolls per plant with seed cotton yield. For Fiber Quality Traits, Staple Length (S/L) revealed significant positive association with ginning out turn (G.O.T%) but displayed negative association with yield related attributes. These findings are in synchronization with Yaqoob *et al.* (2016). For Fiber Strength (STR) positive significant association was displayed with bolls per plant (BPP) and Ginning out Turn (G.O.T%) Sympodial Branches per plant (SBP) and Seed cotton yield (SCY) but for rest of the traits, its association is negative.

Table 1. List of cotton genotypes used in the study.

Sr. No	Genotype Name	Sr. No	Genotype Name
1	AA-904	9	MM-58
2	AA-905	10	NIBGEE-4
3	CIM-599	11	NIBGEE-5
4	CIM-602	12	VH-282
5	FH-118	13	VH-303
6	FH-142	14	NIAB BT-1
7	FH-172	15	MNH-886
8	IUB-222		

Table 2. Mean Square values of traits under study.

SOV	DF	PH	DFF	MBP	SBP	BPP	G.O.T	SCY	S/L	FF	STR
Replication	2	7.22	0.62	0.03	0.20	40.0	0.13	29659	0.02	0.04	0.52
Genotype	14	311.1**	40.1**	0.25**	7.0**	15.5**	13.5**	179662**	1.1**	0.2**	29.9**
Error	28	3.19	0.33	0.035	0.93	5.73	0.14	5640	0.07	0.02	0.16

*** Significant at 1% level*

PH= Plant height, DFF =days to 1st flower, MBP=Monopodial branches/plant, SBP=Sympodial branches/plant, BPP=Bolls/plant, G.O.T = Ginning out turn, SCY=Seed cotton yield, S/L= Staple Length, FF=Fiber fineness, STR= Fiber Strength

Table 3. Range, Mean values and Standard deviation of studied traits.

Traits	Minimum	Maximum	Mean	Std. deviation
PH	101.0	133.0	113.13	9.78
DFF	52.0	64.0	57.73	3.78
MBP	1.10	2.0	1.63	0.282
SBP	15.50	19.7	17.54	1.54
BPP	20.0	28.0	24.73	2.25
G.O.T%	35.20	42.20	39.48	2.12
Yield	1923.0	2660.0	2185.60	244.72
S/L	27.6	30.2	28.86	0.62
FF	4.50	5.50	4.77	0.27
STR	93.30	104.0	99.27	3.07

PH= Plant height, DFF =days to 1st flower, MBP=Monopodial branches/plant, SBP=Sympodial branches/plant, BPP=Bolls/plant, G.O.T = Ginning out turn, SCY=Seed cotton yield, S/L= Staple Length, FF=Fiber fineness, STR= Fiber Strength

Table 4. Correlation coefficients between yield and yield related traits in cotton genotypes.

Variables	PH	DFF	MBP	SBP	BPP	G.O.T%	Yield	S/L	FF
DFF	-0.1127								
MBP	0.0526	**0.8781**							
SBP	0.0609	**-0.8904**	**-0.8660**						
BPP	0.1769	**-0.5283**	**-0.5364**	**0.5695**					
G.O.T%	-0.0319	**-0.6413**	**-0.6884**	**0.7180**	0.3532				
Yield	0.1090	**-0.8711**	**-0.9111**	**0.9286**	**0.7322**	**0.6632**			
S/L	-0.2809	-0.0467	-0.0668	0.0740	-0.1400	0.3106	0.0000		
FF	-0.1533	0.0340	0.1049	0.0655	-0.0355	-0.1652	0.0126	-0.0898	
STR	-0.0208	-0.0197	-0.3818	0.1780	0.3342	0.2187	0.2991	-0.1842	-0.1756

Values in bold form are significantly correlated

PH= Plant height, DFF =days to 1st flower, MBP=Monopodial branches/plant, SBP=Sympodial branches/plant, BPP=Bolls/plant, G.O.T = Ginning out turn, SCY=Seed cotton yield, S/L= Staple Length, FF=Fiber fineness, STR= Fiber Strength

Table 5. Principal Components (PCs) for ten plant characters in selected cotton genotypes.

	PC1	PC2	PC3	PC4
Eigen value	4.755	1.488	1.182	1.046
Variability (%)	47.552	14.887	11.824	10.464
Cumulative %	47.552	62.441	74.264	84.729

Factor Loading/ Correlation of PCs with Plant traits

Plant Traits	PC1	PC2	PC3	PC4
PH	0.0368	-0.5189	0.0407	-0.6431
DFF	-0.4153	-0.0626	-0.1780	0.2035
MBP	-0.4322	-0.0526	0.0595	-0.1041
SBP	0.4326	0.0677	0.1640	-0.0372
BPP	0.3205	-0.2960	0.0334	0.0965
G.O.T%	0.3543	0.2601	-0.1924	-0.0918
Yield	0.4462	-0.0580	0.0987	0.0257
S/L	0.0294	0.6757	-0.2148	-0.1737
FF	-0.0347	0.0767	0.7768	0.3833
STR	0.1513	-0.3135	-0.4887	0.5807

Table 6. Clustering of genotypes on linkage distance basis.

Cluster Names	Cotton genotypes falling under specific cluster
Cluster 1	FH142, MNH-886, IUB-222, FH-172, AA-904
Cluster 2	VH-303, AA-905, CIM-602, FH-118, NIBGE-4, NIBGE-5, MM-58, VH-282, CIM-599, NIAB-Bt-1

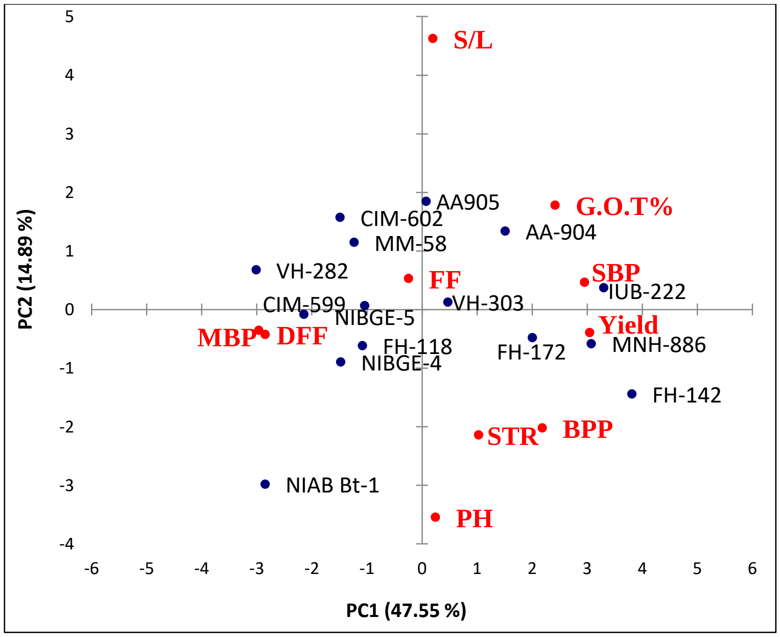

Fig. 1. PC1/PC2 Biplot analysis amongst cotton genotypes and studied traits.

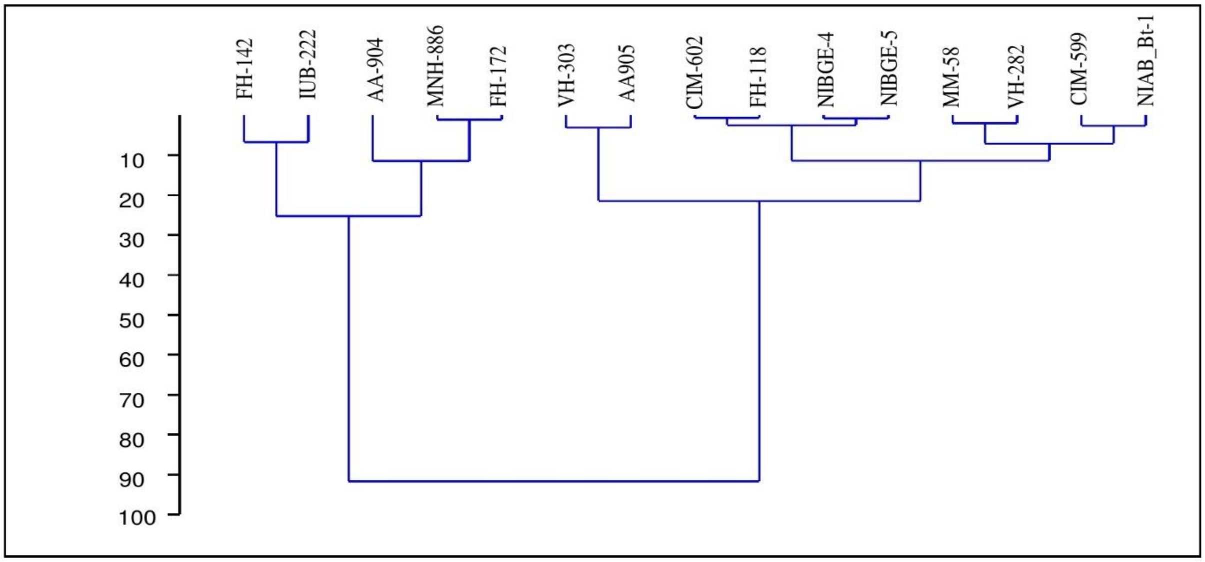

Fig. 2. Grouping of genotypes on linkage distance basis (Ward's Method).

Principal Component Analysis

Principal components analysis was applied on all attributes concurrently to get the in depth understanding of variation and its sequence, thus obtained Eigen values described clearly the proportion of variation for ach principal component, whereas direction and interaction among genotypes are well depicted by Eigen vectors. By all of ten principal components only four PCs revealed the Eigen value >1. The first Principal component counted for 47.55% variability (Table 5) and the seed cotton yield (SCY) contributed more to the variation followed by sympodial branches per plant (SBP), Ginning out turn (G.O.T%) and bolls per plant (BPP), respectively.

Whereas days to First flower, monopodial branches per plant and Fiber Fineness participated negatively to the first principal component. Khan *et al.* (2015) narrated the positive contribution of bolls/plant (BPP), sympodial branches per plant (SBP) to the greater variation in principal component. The PC2 which accounted for 14.88% variation anchored by staple length, followed by Ginning out Turn, but negatively affected by almost all other traits.

The principal component three (PC3) which represents 11.82% variation attained by Fiber Fineness (FF), sympodial branches per plant (SBP) with minimum contribution of Monopodial branches per plant (MBP), Seed Cotton Yield (SCY) and Bolls/plant (BP) but negative effects from rest of the traits. PC4 contained 10.46% variation and characters attributed positively to this component are fiber fineness (FF), Fiber Strength (STR), and days to First Flower (DFF) but gained negative participation from all other studied traits. The first principal component revealed maximum number of traits for variation and so these attributes can take part major role to create more diverse genotypes and can be utilized in varietal development programs.

Biplot Analysis

Ultra-imposition of studied variables is exhibited by the bi-plot of principal component analysis as vector described the direction and extent of variability in each variable (Fig.1) The genotypes which are close to the origin depicted the more similarity with one another but those scale away from the origin representing maximum variation and less similarity. It is also pertinent to mention that fiber quality traits are not in positive direction and opposite to yield related attributes except for sympodial branches per plant.

Agglomerative Hierarchical Clustering (AHC)

Dendrogram obtained through Ward’s method placed the genotypes in two main groups (Fig. 2). Group (A) comprised of five genotypes *viz.*, FH-142, IUB-222, AA-905, FJ-172 and MNH-886, while GROUP (B) comprised rest of the ten Genotypes *viz.*, VH-303, AA-905, CIM-602, FH-118, NIBGE-4, NIBGE-5, MM-58, VH-282, CIM-599, NIAB-Bt-1 (Table 6). Baxevanos *et al.* (2008) studied the interdependence of different traits among cotton genotypes and arranged them in seven clusters. As Linkage Cluster Analysis, placed genotypes on the basis of genetic distance and those have more allelic differences stationed in different groups.

REFERENCES

Ahmad, M.Q., S.H. Khan and F.M. Azhar (2012). Decreasing level of genetic diversity in germplasm and cultivars of upland cotton (*Gossypium hirsutum* L) in Pakistan. *J. Agric. Social Sci.*, 8: 92-96.

Anonymous (2018-19). Economic Survey of Pakistan, Ministry of Finance, Govt. of Pakistan.

Baloch, M.J. and Q.B. Baloch (2004). Plant characters in relation to earliness in cotton (*Gossypium hirsutum* L.) *Proc. Pakistan Acad. Sci.*, 41: 103-108.

Baloch, M.J., C. Kumar, W.A. Jatoi and I.H. Rind (2014). Phenotypic correlation and regression analysis of yield and fibre traits in upland cotton (*Gossypium hirsutum* L.). *Pak. J. Agri. Agril. Engg. Vet. Sci.*, 30(2): 135-146 (2014).

Baxevanos, D., C. Goulas, S. Tzortzios and A. Mavromatis (2008). Interrelationship among and repeatability of seven stability indices estimated from commercial cotton (*Gossypium hirsutum* L.) variety evaluation trials in three Mediterranean countries. *Euphytica*, 161: 371-382.

Bhatti, M. H., M.I. Yousaf, A. Ghani, M. Arshad, A. Shehzad, A. Mumtaz, M.U. Khalid, M.Z. Khalid, M.Z. Mushtaq and S.A.S. Shah (2020). Assessment of genetic variability and traits association in upland cotton (*Gossypium hirsutum L.*). *Int. J. Botany Stud.*, 5 (2): 148-151.

Khan, F. Z., S. U. Rehman, M. A. Abid, W. Malik, C.M. Hanif, M. Bilal, Q. Qanmber, A. Latif, J. Ashraf and U. Farhan (2015). Exploitation of germplasm for plant yield improvement in cotton (*Gossypium hirsutum* L.). *J. Green Physiol. Genet. Genom.*, 1: 1-10.

Latif, A., M. Bilal, S.B. Hussain and F. Ahmad (2015). Estimation of genetic divergence, association, direct and indirect effects of yield with other attributes in cotton (*Gossypium hirsutum* L.) using biplot correlation and path coefficient analysis. *Tropical Plant Res.*, 2: 120–126.

Saeed, M., A. Mumtaz, D. Hussain, M. Arshad, M.I. Yousaf and M.S. Ahmad (2018). Multivariate analysis-based evaluation of maize genotypes under high temperature stress. *I3 Biodiversity*, 1, 105: 1-12

Salahuddin, S., S. Abro, M.M. Kandhro, L. Salahuddin and S. Laghari (2010). Correlation and path coefficient analysis of yield components of upland cotton (*Gossypium hirsutum* L.). *World Appl. Sci. J.*, 8: 71- 75.

Shakeel, A., I. Talib, M. Rashid, A. Saeed, K. Ziaf and M.F. Saleem (2015). Genetic diversity among upland cotton genotypes for quality and yield related traits. *Pak. J. Agric. Sci.*, 52: 73-77.

Sneath, P.H.A. and R.R. Sokal (1973). *Numerical Taxonomy: The Principles and practice of numerical classification.* Free-Man WF and Co, San Francisco, USA.

Steel, R.G.D., J.H. Torrie and D. Dickey (1997). *Principles and procedures of statistics. A Biochemical Approach.*3rd Ed. McGraw Hill book Co., New York P.172-177.

Yaqoob, M., S. Fiaz and B. Ijaz (2016). Correlation analysis for yield and fiber quality traits in upland cotton. *Commun. Plant Sci.*, 6: 55-60.

Yousaf, M. I., K. Hussain, S. Hussain, A. Ghani, M. Arshad, A. Mumtaz and R.A. Hameed (2018) Characterization of indigenous and exotic maize hybrids for grain yield and quality traits under heat stress. *Int. J. Agric. Biol.*, 20: 333-337.

Yousaf, M. I., K. Hussain, S. Hussain, R. Shahzad, A. Ghani, M. Arshad, A. Mumtaz and N. Akhter (2017). Morphometric and phonological characterization of maize (*Zea mays* L.) germplasm under heat stress. *Int. J. Biol. Biotech*., 14 (2): 271-278.

(*Accepted for publication June 2020*)