International Journal of Multidisciplinary Research and Development
www.allsubjectjournal.com
Online ISSN: 2349-4182, Print ISSN: 2349-5979
Received: 31-07-2021, Accepted: 16-08-2021, Published: 01-09-2021
Volume 8, Issue 9, 2021, Page No. 16-21

Determination of genomic diversity of upland cotton *(Gossipium hirsutum)* genotypes using ssr markers

Samreen Sarwar[1], Syed Bilal Hussain[1*], Syed Aun Muhammad[1], Muhammad Zubair[2], Tahir Naqqash[1]
[1] Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Multan, Pakistan
[2] Department of Forestry and Range Management, Faculty of Agricultural Sciences and Technology, Bahauddin Zakariya University, Multan, Pakistan

Abstract
Genetic diversity is prerequisite for any breeding plan. Genetic diversity along with phenotypic diversity provides the chance to select the genotypes with multiple good traits simultaneously. To find out the genetic diversity among 45 accessions of *Gossypium hirstum*, 44 Simple sequences repeats (SSR) and expressed sequence tag SSRs (EST-SSR) were used (from BNL, DPL, JESPER and NAU series). Total of 222 alleles were produced with an average of 4.9 alleles per primer. Polymorphism Information Content (PIC) value was ranged from 0.16 to 0.19 with a mean value of 0.76. Pair-wise genetic estimation based on Nei 1973 ranged from 0.50 to 0.96, with an average 0.79. In bootstrap cluster analysis maximum accession falls into a single cluster indicating low genetic diversity. Positive correlation among phenotypic traits ranged from 0.001 to 0.99 while, negative correlation from -0.004 to -026. Principle components analysis (PCA) analyses generate the 7 PCs components contributing 71.1%. Maximum variation contributed by PC1 is 14.3. Cluster analysis grouped the 45 accession into 4 clusters. Twenty nine out of 45 genotypes fall in a single cluster. Phenotipycally CIM-707 showed maximum diversity while, genotypically CIM-707 showed little resemblance with CIM-600, CIM-132 and CIM-622. Low genetic diversity was observed among both genetic and phenotypic as maximum accessions shears the same group. This study helps for selecting diverse accessions with multiple phenotypic traits simultaneously. This also suggests furthering elaborating the molecular genetic diversity by using more SSR marker to improve the resolution of upland cotton for good breeding plans.

Keywords: cotton, genetic diversity, phenotypic diversity, SSRS marker

Introduction
Cotton is a warm-climate shrub of *Malvaceae* family and the genus *Gossypium* that grows naturally as a perennial (Carrière *et al*., 2016) [5]. The genus *Gossypium* is a critical oilseed and spinnable fiber material crops (Karishma *et al*., 2016) [14], grown in more than 100 tropical and subtropical areas (Parekh *et al*.,2016) [19]. It is believed that from 2000 to 2030 global demand for cotton and cotton product will increase up to 102 % (Xiao *et al*., 2009) [27]. The remarkable outcomes of cotton crop positions eight among significant edible oil producing crops in all over the world. Dehydrated cotton seed contain roughly 28.3–4 4.1% oil which comprised of immersed unsaturated and few unsaturated fatty acids. Cotton seed oil consist of more than 50 % linoleic acid, major products of cotton crop are lint, cotton seed, seed supper and cotton cake (Keshamma *et al*., 2009) [15]. *Malvaceae* family has broad phenotypic variation among more or less 46 diploid (26) and the 5 allotetraploid (52) species (3). Among fifty one species merely 4 genotypes are mostly used to cultivate, counting two diploids, *Gossypium arboreum L* and the *Gossypium herbaceum L.* gives two percent of the world's cotton while two tetraploids species. *Gossypium hirsutum* (*G. hirsutum*) that contribute about ninety percent, *Gossypium barbadense*, yield eight percent (Tyagi *et al*., 2006).
Genetic diversity provides the building blocks for the crop improvement. For the description of the genetic similarity between the cultivars is an important front to choose best parent combinations, in favor of genetic diversity in a breeding program, consequently it is a key for breeders to have a decent knowledge of genetic diversity of upland cotton, its structure and connections between cultivars (Ehsan *et al*., 2013) [8]. The genetic diversity of upland cotton and its related species has been characterized using pedigree and morphological data (May *et al*., 1995) [17], biochemical markers like isozyme (Wendel *et al*., 1992) [26] and DNA-based molecular markers (Fang *et al*., 2013) [9]. Among the DNA based method SSR for study the genetic diversity SSR loci are especially helpful (Ehsan *et al*., 2013) [8] that provide population arrangement of domesticated species due to their higher rank for allelic diversity.
Various authors designate genetic diversity as the base for Polymorphism and species diversity and play an important role in cotton breeding program across the globe (Chen and Du., 2006) [6]. Pereira *et al*., (2015) [20] emphasized increase of diversity by using wild genotypes in breeding programmes. Korir et al., (2013) [16] described the use of molecular markers as cost effective and rapid technique for varietal identification. Since cotton is an important cash and export crop of Pakistan, therefore, objective of present research is to identify genotypes with multiple good traits that could provide a broad base for an effective breeding program.

Material and Methods
Plant material
We obtained seeds of 50 accessions of *G. hirsutum* from Central Cotton Research Institute (CCRI) Multan, Pakistan.

Accessions were selected based on genetic diversity. These seeds were grown in field of the Faculty of Agriculture Sciences and Technology Bahauddin Zakariya University, Multan. Fresh leaves are selected to extract the DNA by using CTAB method. 0.8% agarose gel used to quantify the DNA. Good quality DNA sample selected for the PCR by using 44 SSR markers were used for the present study (Table 1). SSR markers were used from BNL (Huntsville, AL, USA, http://www.resgen.com), DPL, NAU (EST SSR) (Han *et al*., 2006) [12] and JESPER series. (Reddy *et al*., -2001; Akhtar *et al*., 2010) [22, 1] The chosen primers were randomly distributed across the whole cotton genome, and their sequence was taken from the cotton marker database (http://www.cottonmarker.org/*)*. These primers were designed by Oligo Humanizing Genomics Macrogen, PCR product was analyzed by using DNA polyacrylamide gel electrophoresis (PAGE).

Table 1: Primers, annealin temperature, alleles and PIC value

	Primer	Annealing	Total No of allele	PIC
1	BNL2449	55ºc	5	0.90
2	BNL3171	55ºc	8	0.85
3	BNL7948	55ºc	3	0.68
4	BNL2662	55ºc	7	0.84
5	BNL827	55ºc	4	0.85
6	BNL2448	55ºc	3	0.65
7	BNL3052	55ºc	8	0.92
8	BNL3651	55ºc	4	0.69
9	BNL1066	55ºc	2	0.57
10	BNL4096	55ºc	3	0.69
11	BNL3410	55ºc	8	0.90
12	BNL786	55ºc	3	0.62
13	DPL0323	55ºc	8	0.94
14	DPL0519	55ºc	7	0.86
15	DPL0079	55ºc	4	0.75
16	Jesper274	55ºc	7	0.91
17	Jesper134	55ºc	5	0.87
18	Jesper292	55ºc	5	0.70
19	Jesper101	55ºc	2	0.49
20	Jesper36	55ºc	5	0.84
21	Jesper110	55ºc	5	0.77
22	Jesper153	55ºc	7	0.88
23	NAU5170	55ºc	4	0.73
24	NAU8786	55ºc	4	0.85
25	NAU2095	55ºc	7	0.68
26	NAU2954	55ºc	4	0.83
27	NAU5046	55ºc	5	0.83
28	NAU2868	55ºc	5	0.74
29	NAU2161	55ºc	2	0.16
30	NAU2715	55ºc	4	0.76
31	NAU988	55ºc	8	0.92
32	NAU5418	55ºc	8	0.83
33	NAU2714	55ºc	8	0.85
34	NAU5269	55ºc	3	0.54
35	NAU1070	55ºc	7	0.87
36	NAU3911	55ºc	5	0.86
37	NAU2836	55ºc	6	0.86
38	NAU1672	55ºc	5	0.90
39	NAU980	55ºc	6	0.73
40	NAU2838	55ºc	4	0.46
41	NAU3948	55ºc	4	0.81
42	NAU5121	55ºc	2	0.57
43	NAU1366	55ºc	4	0.64
44	NAU2437	55ºc	4	0.78

Statistical analysis

Power marker v 3.25 was used to find out the genetic diversity (http://www.powermarker.net.). To calculate the genetic relationships among different genotypes a dendrogram was constructed using bootstrap neighbor joining method based on 1972 model (Nei., 1973) was used in which similar genotypes were grouped together in clusters by using statistical software power marker version 3.25. (It also measured the gene diversity, alleles number and PIC). The mean data of all the 16 morphological traits were analyzed by using basic statistic correlation analysis by using statistical software Minitab v 16. Principle coordination analysis and cluster analysis was done by using statistical software SPSS v 19 (Sneath *et al*., 1973) [24].

Results and Discussion

Phylogenic tree

Phylogenic tree was generated by using bootstrap neighbor joining (Nj) method based on Nei (14) similarity coefficient which cluster all the accession into cluster, sub-cluster and sub-sub cluster. Main three clusters are formed named as A, B and C (Fig.1). Cluster A containing 7 accessions formed an isolated group from the rest of cluster and sub divided into two sub cluster named as A1 and A 2. Sub group A1 further divided into sub-sub groups as (A1-a) containing isolated CRIS-134 and (A1-b) having three varieties CRSM-38, BH-60 and AA-803. Sub group A2 also differentiate into two sub-sub group (A2-a) containing separate BZU-75 while sub-sub group (A2-b) comprising NIBGE-2 and FH-142. Cluster B also formed to somewhat independent group of 5 accessions by dividing into two sub groups B1 and B2. Sub group B1 further divided into (B1-a) having CIM-600 and (B1-b) CIM-622 and CIM-132. B2 sub group contained CIM-707 and AGC-333. Most of accession used in this study are belongs to cluster C, having sub and sub-sub groups which finely differentiating the accession according to their ancestry. Cluster C divided into two major groups C1 and C2.C1 group furthermore divided into two sub-sub groups (C1-a) and (C1-b). Additionally (C1-a) also divided into (C1a-a) and (C1a-b). (C1a-a) contain CIM-496, AGC-337, AGC-333, AGC-335, AGC-334, CIM-446, CIM-336 and a separate group of NIBGE-3 and MNH-886. On the other hand (C1a-b) consist of CIM-506, IUB-222, FH-14, N121, ASG, SITARA-008 and CIM-448.(C1-b) comprise 1R-3701, FH-113, CYTO-178, CYTO-171, AA-802, CIM-473, CIM-534 and CIM-482.C2 cluster divided into two main sub cluster (C2-a) further divided into (C2a) having CIM-602, CIM-599, N-111, and CEMR-33. (C2-b) consist of CIM-616, FH-901 an AA-703. Isolated (C2-b) cluster contain CIM-598 and CIM-573.

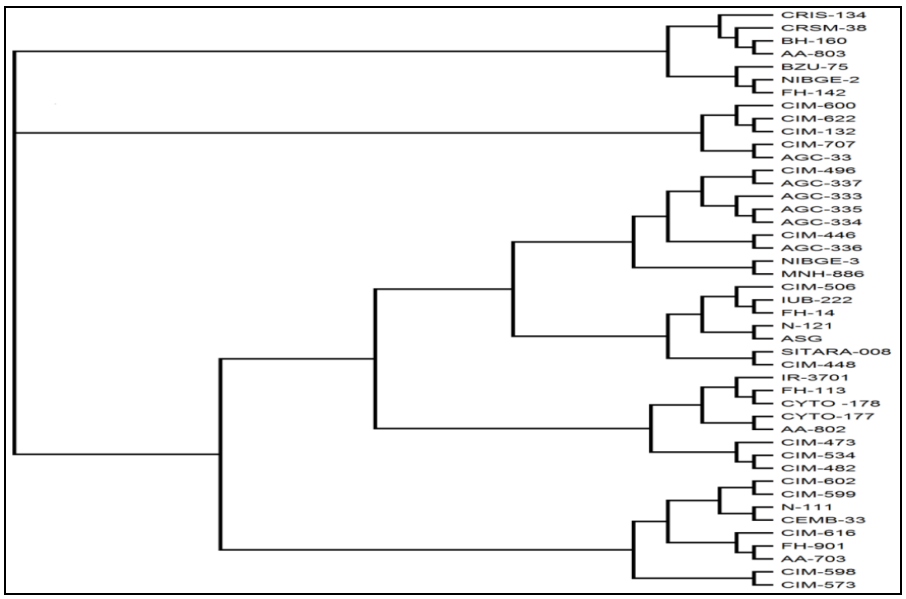

Fig 1: Phylogenic tree was generated by using bootstrap neighbor joining (Nj) method based on Nei 1973 similarity coefficient which cluster all the accession into cluster, sub-cluster and sub-sub cluster*).*

Correlation Analysis

Simple correlation coefficients were used to find out the association among 16 studied morphological traits (Table 2). Association among the different traits is very useful to find out the variations. It is also important for new breeding plans because correlation provide the opportunity to find a genotype with multiple desirable traits simultaneously (Ali *et al*., 2009) [2, 3]. In this study correlation analysis was performed by using Minitab v.16 software. Bolls weight showed maximum significant positive correlation (0.99) with locks per plant. Minimum positive correlation was found among yield per plant and bolls per plant with a correlation value (0.001). Maximum negative correlation was found among chlorophyll percentage and sympodial number (-0.004). It also showed minimum negative correlation (-026) with monopodial node number per plant and positive correlation with plant height (0.36) and1st monopodial number (0.12).

Table 2: Correlation analysis among morphological traits of upland Cotton

	PH	1stMN	Mon/p	Sym/P	Ch%	Bolls/p	S,index	Locks/B	B.Wt	Seed/B	GOT	Yield/p	Uni.i	Micronair	strength
1stMN	0.20														
Mon/p	-0.25	0.13													
Sym/P	0.31	0.08	0.08												
Ch%	0.32	0.12	-0.26	0.00											
Bolls/p	-0.03	0.14	-0.06	0.07	-0.19										
S,index	0.02	0.15	0.36	0.05	0.15	0.41									
Locks/B	0.28	-0.05	-0.04	0.25	0.37	0.06	0.40								
B.Wt	-0.06	-0.17	0.20	-0.03	0.02	0.87	0.17	1.00							
Seed/B	-0.02	-0.07	-0.01	-0.11	0.16	0.46	0.28	0.17	0.07						
GOT	0.28	0.01	0.01	-0.21	0.25	0.20	0.79	0.27	0.12	0.67					
Yield/p	-0.20	-0.12	0.10	0.30	0.05	0.00	0.58	0.52	0.26	0.79	0.31				
Uni.i	-0.01	-0.02	0.03	0.28	-0.01	0.60	0.09	0.83	0.63	0.81	0.85	0.67			
Micronar	-0.11	-0.26	0.06	-0.17	-0.17	0.94	0.83	0.34	0.79	0.39	0.65	0.58	0.84		
strength	-0.17	0.15	0.25	-0.25	-0.08	0.86	0.51	0.16	0.08	0.39	0.88	0.76	0.77	0.90	
St.length	0.22	0.06	0.21	-0.11	0.01	0.70	0.07	0.02	0.17	0.09	0.58	0.43	0.13	0.19	0.46

Principle Component Analysis (PCA)

Principle component analysis deals with the division of the total variance into its components. These components are helpful for employment and conservation of genetic resources. It also makes possible the utilization of satisfactory germplasm for the betterment of crop for a specific trait. For effective breeding plans PCA is efficient tool to select the parental lines (Amna *et al*., 2009) [4]. In this case out of 16 principle component (PCs), 7 components were obtained having Eigen value greater than 1. Theses seven character contributed about 71.1% diversity among 45 accessed genotypes (Table 3).

Table 2. Correlation analysis among morphological traits plant height (PH), 1st monopodial number (1stM N), Monopodial number (Mon/p), Sympodial node (Sym/P), Chlorophyll percentage (Ch%), Bolls per plant (Bolls/p), Seed index (S,index), Locks per Boll (Locks/B), Boll weight (g) (B. Wt), Seed per ball (Seed/B), Ginning out turn (GOT), Yield per plant (g) (Yield/p), Uniformity index (%) (Uni.i), Micronaire (Micronair), Staple strength (g/tex), (strength) Staple length (mm)

Table 3: Principle Component Analysis (PCA)

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.3	14.25	14.3
2	2.00	12.48	26.7
3	1.79	11.21	37.9
4	1.60	10.00	47.9
5	1.37	8.57	56.5
6	1.24	7.77	64.3
7	1.09	6.82	71.1
8	0.89	5.59	76.7
9	0.79	4.91	81.6
10	0.75	4.71	86.3
11	0.63	3.94	90.2
12	0.48	3.01	93.2
13	0.42	2.63	95.9
14	0.30	1.85	97.7
15	0.19	1.21	98.9
16	0.17	1.08	100.0

Cluster analysis

Cluster analysis of phenotypic traits divided the genotypes in to different clusters having desirable traits into one cluster. Forty five (45) genotypes grouped into 4 clusters based on various traits Cluster analysis indicates that cluster 1 contained 7 genotypes cluster 2 consist of 29, cluster3 comprised on 8 while, cluster 4 consist of only 1 genotype (Table 4). Genotype in cluster 1 showed the maximum responsible value of yield per plant, Uni.i plant height. Cluster 2 shows the genotypes with maximum proportion of plant height, Uni.i, Chlorophyll percentage and GOT. Cluster 3 consists of genotypes with plant height, GOT and Uni.i. Same results wad also reported by (Saeed *et al*., 2014) [23].

Table 4: Cluster memberships of different genotypes

Clusters	Member genotypes
Cluster 1	ASG-336, FH-113, SITARA-008, FH-14, CEMB-33, N-111, N-121
Cluster 2	ASG, CIM-506, CIM-622, CIM-132, CIM-600, NIBGE-3, MNH-886, CIM-448, AGC-334, AGC-335, CIM-446, CIM-482, CIM-534,CIM-573,CIM-598,CIM-599, BZU-75, NIBGE-2, CIM-616, FH-901, AA-703, AA-802, CYTO-177, CYTO-178, FH-142, IR-3701, BH-160, AA-803, CRIS-134
Cluster 3	AGC-333, AGC-337, CIM496, CIM-602, CIM-473, CRSM-38, AGC-33
Cluster 4	CIM-707

In our study 222 alleles were amplified ranged from 8 to 2 alleles with a mean of 4.9 alleles per primer, which are cumulatively 76% informative. The most informative primer was DPL 0323 which showed the polymorphism up to 94%. NAU 1070 produced 7 bands and responsible for 87% polymorphism. Almost 5 primers NAU 988, JESPER274, BNL2449, BNL 3052 and DPL 0323 are informative more than 90%, but none of them able to differentiate all the varieties. According to (Powell, Morgante et al. 1996) [21] (19) SSR makers can reveal prominent genetic diversity among closely related genotypes even when few loci are use. SSR markers expose huge amount of genetic variations among sample genotypes, however average polymorphism observed from these primers is considerably low. By using SSR marker observation of narrow genetic base in upland cotton is discussed in many studies (Kalivas, Xanthopoulos et al. 2011) [13] (20-21), (Zhang, Wang et al. 2011) [28].

A phylogenic bootstrap neighbor joining tree based on Nei (1972) [18] similarity index was constructed. This tree divided the 45 genotypes under study into different clusters according to phylogenic similarity. Three major groups are formed based on ancestry. Our results revealed that CIM482, CIM 534, CIM506 and CIM473 have strong genetic association and fall in the same group. Such results were also reported by Saeed *et al*., (2013). In our results AA703, N121, FH133 and AA802 sheared the same cluster.

The selection of genotypes with multiple traits of interest is important for any breeding program. Association among different traits is important for this aim, as it may be helpful to recognize a variety with a combination of different desired traits simultaneously (Ali et al., 2009) [2, 3]. Simple correlation among morphological traits ranged from (0.001) to (0.999). Our results indicate maximum correlation among locks per plant and boll weight. Yield per plant reveled negative correlation with plant height (-0.19) and 1st monopodial number (-0.12). Plant height showed negative correlation with lint strength, micronaire, yield seed per boll, boll weight, bolls per plant, and monopodial nodes per plant. Staple length revealed positive correlations with plant height which indicated its positive behavior toward plant yield. Such results propose that these genotypes could be of better choice at the time of selection of high yielding genotypes. Same results were also reported by (Farooq *et al*., 2011; Farooq *et al*., 2014) [10, 23], (Saeed *et al*., 2014, Farooq *et al*., 2014) [23, 23].

The dissolution of variances into their components is helpful for the profit and stability of the genetic resources. It also gives the opportunity to select the genotypes with a particular trait, indirectly important for crop improvement. In this study we obtained seven PCs component that are responsible for 71.1% variations among our selected morphological character. PC1 contributes maximum variation with (14.25%) toward the total diversity. Importance of PC1 also discussed by (Mujaju and Chakauya 2008) PC2 contribute (12.5%), followed by PC3 (11.2) %, PC4 (9.99%), PC5 (9.570), PC6 (7.8%) and PC 7 (6.81).

In our case 1st PC was majorly because of variations in monopodial number, bolls per plant, GOT and seed index. 2nd PC was due to diversity of plant height, sympodial node, S index and yield per plant. PC 3 was related to diversity among staple length, chlorophyll percentage and GOT. PC4 formed because of the diversity of ball weight, seed per ball, yield per plant and sympodial node. PC5 contained the diversity of1st monopodial number, seed per ball and staple

strength. PC6 contained the variation among bolls per plant, yield per plant, GOT. PC7 is mainly because of the variations among Chlorophyll percentage, S.index and yield per plant. PC analysis is useful to confirm the diversity because of traits in material studied, which could be helpful for future studies.

To find out a genotype with combination of different traits simultaneously cluster analysis was used. Cluster analysis of 16 traits distributed the 45 accession in to 4 clusters. Maximum accessions are fall in cluster 2 while cluster 4 contains only a single genotype CIM-707. Maximum variations are because of cluster 3 components. GOT, Uni.i and plant height reveled the most variations. Same results also concluded by (Farooq *et al*., 2014) [23].

Cluster analysis of morphological traits kept the genotype CIM-707 in a separate group. Cluster 4 contains only a single accession indicating CIM-707 more diverse then other 44 selected genotypes. In addition, cluster 4 is most distant to cluster 3 (AGC-333, AGC-337, CIM496, CIM-602, CIM-473, CRSM-38, and AGC-33). All these results are confirmed by genotypic analysis of theses accessions. Dendrogram based on genotypic data also showed more or less same pattern of clusters. CIM-707 fall in an isolated group indicating genotypic divergence then other accessions, but genotypic analysis revealed some genetic connection of CIM-707 with CIM-600, CIM-132 and CIM-622.

Conclusion

The current study showed efficiency of SSR markers in identifying diversity in cotton as in line with Dahab et al., 2013. It is further concluded that phonotypical CIM-707 showed more diversity then other traits while, genotypically CIM-707 sheared the same genetic background as CIM-600, CIM-132 and CIM-622.

References

1. Akhtar K, Haidar S, Khan M, Ahmad M, Sarwar N. Evaluation of Gossypium species for resistance to cotton leaf curl Burewala virus. Annals of Appl Bio,2010:157:135-147.
2. Ali MA, Khan IA, Nawab NN. Estimation of genetic divergence and linkage for fibre quality traits in upland cotton. J Agric Res,2009:47(3):229-236.
3. Ali MA, Nawab NN, Abbas A, Zulkiffal M, Sajjad M. Evaluation of selection criteria in Cicer arietinum L. using correlation coefficients and path analysis. Austr J Crop Sci,2009:3:65-70.
4. Amna N, Jehanzeb F, Abid M, Muhammad S, Muhammad R. Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of Gossypium hirsutum L. *Вестник Орловского государственного аграрного университета*,2009:43:2241-2267.
5. Carrière Y, Fabrick JA, Tabashnik BE. Can pyramids & seed mixtures delay resistance to Bt crops? Trends Biotech,2016:34:291-302.
6. Chen G, DU, XM. Genetic diversity of source germplasm of upland cotton in China as determined by SSR marker analysis. Acta Genetica Sinica,2006:33:733-745.
7. Dahab AA, Saeem M, Mohamed BB, Ashraf MA, Puspito AN, Bajwa KS *et al*. Genetic diversity assessment of cotton (Gossypium hirsutum L.) genotypes from Pakistan using simple sequence repeat markers. Aust J Crop Sci,2013:7(2):261-267.
8. Ehsan B, Haque A, Younas M, Shaheen T, Huma T. Assessment of genomic diversity of cotton (Gossypium hirsutum) genotypes using simple sequence repeats markers through genetic analysis software. Int. J. Agric. Bio,2013:15:968-972.
9. Fang DD, Hinze LL, Percy RG, LI P, Deng DA. Microsatellite-based genome-wide analysis of genetic diversity & linkage disequilibrium in Upland cotton (Gossypium hirsutum L.) cultivars from major cotton-growing countries. *Euphytica*,2013:191:391-401.
10. Farooq A, Farooq J, Mahmood A, Shakeel A, Rehman KA, Batool A *et al*. An overview of cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity. Aust J Crop Sci,2011:5(13):1823.
11. Farooq J, Anwar M, Riaz M, Farooq A, Mahmood A, Shahid MTH *et al*. Correlation and path coefficient analysis of earliness, fiber quality and yield contributing traits in cotton (Gossypium hirsutum L.). J of Ani Plant Sci,2014:24(3):781-790.
12. Han Z, Wang C, Song X, Guo W, Gou J. Characteristics, development & mapping of Gossypium hirsutum derived EST-SSRs in allotetraploid cotton. Theor App Genet,2006:112:430-439.
13. Kalivas A, Xanthopoulos F, Kehagia O, Tsaftaris AS. Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. Genet Mol Res,2011:10(1):208-217.
14. Karishma R, Lakshmi, Sahithya U, Suneetha P, Krishna M. Determination of Total Gossypol & Free Gossypol Content in different varieties of Bt and Non Bt Cotton seed extracts by High-Performance Liquid Chromatography (HPLC). Res J Biotech,2016:11(2):70-74.
15. Keshamma E, Rohini S, Rao K, Madhusudhan B, Kumar MU. Tissue culture-independent in planta transformation strategy: an Agrobacterium tumefaciens-mediated gene transfer method to overcome recalcitrance in cotton (Gossypium hirsutum L.). J. Cotton Sci,2009:12:264-272.
16. Korir NK, Han J, Shangguan LF, Wang C, Kayesh E, Zhang YY *et al*. Plant variety and cultivar identification: advances and prospects. Critical Reviews Biotech,2013:33:111-125.
17. May OL, Bowman DT, Calhoun DS. Genetic diversity of US upland cotton cultivars released between 1980 & 1990. Crop Sci,1995:35:1570-1574.
18. Nei M. Genetic distance between populations, Am Nat,1972:106:283-291.
19. Parekh MJ, Kumar S, Zala HN, Fougat RS, Patel CB. Development & validation of novel fiber relevant dbEST–SSR markers and their utility in revealing genetic diversity in diploid cotton (Gossypium herbaceum & G. arboreum). *Ind. Crops Prod*,2016:83:620-629.
20. Pereira GD, Cazé ALR, Michelle GD, Almeida VC, Magalhães FODC, Silva FLD *et al*. Optimal use of SSR markers for varietal identification of upland cotton. *Pesquisa Agropecuária Brasileira*,2015:50(7):571-581.
21. Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S *et al*. The comparison of RFLP, RAPD, AFLP

and SSR (microsatellite) markers for germplasm analysis. Mole. breed,1996:2(3):225-238.
22. Reddy OUK, Pepper AE, Abdurakhmonov I, Saha S, Jenkins JN. New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. J Cotton Sci,2001:5:103-113.
23. Saeed F, Farooq J, Mahmood A, Riaz M, Hussain T. Assessment of genetic diversity for Cotton leaf curl virus (CLCuD), fiber quality & some morphological traits using different statistical procedures in 'Gossypium hirsutum 'L. Aust J of Crop Sci,2014:8(3):442-447.
24. Sneath PHA, Sokal RR. Numerical taxonomy. Freeman, San Francisco, 1973.
25. Tyagi P, Gore MA, Bowman DT, Campbell BT, Udall JA. Genetic diversity and structure in the US Upland cotton (Gossypium hirsutum L.). Theor. App. Genet,2014:127:283-295.
26. Wendel JF, Brubaker CL, Perciva AE. Genetic diversity in Gossypium hirsutum & the origin of upland cotton. Am J Bot,1992:79:1291-1310.
27. Xiao J, Wu K, Fang DD, Stelly DM, Yu J. New SSR markers for use in cotton (Gossypium spp.) improvement. J Cotton Sci,2009:13:75-157.
28. Zhang Y, Wang XF, Li ZK, Zhang GY, MA ZY. Assessing genetic diversity of cotton cultivars using genomic and newly developed expressed sequence tag-derived microsatellite markers. Genet Mole Res,2011:10(3):1462-1470.