**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. 25-30**

# Molecular and phylogenetic study of CLCuD tolerant cotton cultivars by using microsatellite markers

**Anum Shaheen[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 & Technology, Bahauddin Zakariya University, Multan, Pakistan

**Abstract**
In any breeding plan genetic diversity is significant importance. Genotypic and phenotypic diversity allows a breeder to select varieties having manifold traits concurrently. In current study, to found out the genetic diversity among 66 accessions of *Gossipium hirsutum*, we used thirty-one simple sequence repeat (SSR) markers. These markers produced total 138 loci and an average locus per primer was 4.4. Polymorphism information content (PIC) ranged from 0.27 to 0.92 with an average value 0.59. BNL-2709, NAU-5418, J-274 SSR markers gave maximum polymorphism. Gene diversity ranged from 0.32 to 0.92 and average was 0.64. BNL-2709, NAU-5418, J-274 gave maximum gene diversity. The range of pair-wise genetic estimation based on Nei 1973 was from 0.09 to 0.93. A dendrogram was produced by using the UPGMA analysis which grouped the 66 accessions into two main clusters. One cluster comprises 40 accessions while other cluster constitutes 26 accessions. We also check diversity on the basis of phenotypic traits that allows a breeder to select varieties with interesting genotypic phenotypic characters at the same time. Correlation analysis among phenotypic traits gave positive correlation that ranged from 0.00 to 0.78 while negative correlation ranged from -0.001 to -0.363. Principle component analysis (PCA) was performed which showed variation on the basis of traits among varieties and also show the major contributor in the variation. PCA analysis produced 6PCs component contributing to 86% variability, PC1 is responsible for maximum variation with 18.22. Minimum variation was contributed by PC14 that was 1.103. Cluster analysis divided the 66 accessions in to 5 clusters on basis of various phenotypic traits. Tarzan-5 genotype falls in separate cluster representing its diversity from other accessions. In cluster analysis, genotypes in cluster 1 showed maximum value of plant height (97.1), boll per plant (37.0). Cluster 2 manifested maximum value of plant height (274.3), boll per plant (28.1). Cluster 3 showed maximum value of plant height (151.6), boll per plant (32.0). Cluster 4 manifested maximum value of plant height (101.6), seed per boll (26.3). Cluster 5 showed maximum value of plant height (127.0), seed per boll (24.2).

**Keywords:** *Gossipium hirsutum*, SSR marker; cotton leaf curl disease, polymorphism information content

## Introduction
Cotton (*Gossypium hirsutum L.*: Genus: *Hirsutum*; Family: *Malvaecae*) has got importance as one of the most significant non-food fiber crop in Pakistan. It has got importance in national economy because it offers considerable foreign exchange earnings. Apart from earning foreign revenues, it offers butter and bread to the people in amount of millions (Farooq *et al;* 2018b) [6]. *Gossypium* comprises additional 50 familiar species that belong to group of 8 eight genome. Only 4 species, *G. herbaceum* (A1), *G. hirsutum* (AD1), *G. barbadense* (AD2), and *G. arboreum* (A2), had been domesticated and cultivated. The specie which gives more than 95% of cotton and cultivated globally is *G. hirsutum L.* also called upland cotton whose genome size is 2.5 Gb and chromosome number is 2n = 4x = 52. China includes in one of the biggest cotton-consuming and producing states around the globe. The world's total production of cotton has decreased in previous limited years nevertheless the feeding has amplified yearly (Huang *et al;* 2017). Cotton and cotton related products provide us fifty five percent of the distant exchange wages of country and 10 percent of GDP (gross domestic products). The area of cotton cultivation has enlarged expressively around 7.86 million acres in 2015-2016 in the past 30 years. After USA, India and China Pakistan rank 5$^{th}$ number in cotton production (Farooq *et al;* 2018a) [5]. Many factors limits the cotton production like erratic rainfall, increasing biotic and abiotic stresses decline the yield (Boopathi *et al;* 2015) [4]. Yield reducing abiotic factors are flooding, drought, scold and heat stress (Khalil *et al;* 2017) [8]. Insect pests are the main one in the factors which lowers the yield. One sixty two insect pest species attack cotton globally, that feeds on cotton throughout diverse growth stages. Around 20-40% losses in cotton are informed owing to attack of insect pest in Pakistan. *G. hirsutum L.* is further prone to outbreak of bollworm complex as well as sucking insect pest than *Gossypium arboreum*. Bollworms and sucking pest are two types of cotton pests. Cotton's most important pests are sucking pest like aphids (*Aphis gossypii, Glover*), thrips (*Thrips tabaci Lindeman*), jassids (*Amrasca biguttula, Ishida*) and whiteflies (*Bemisia tabaci, Gennadius*) (Farooq *et al;* 2018b) [6]. The crucial factor in reducing profitable cotton cultivation is *Thrips tabaci Lindemann* (Farooq *et al;* 2018a) [5].

In addition to the abiotic stresses, cotton leaf curl disease solely confines cotton production in Pakistan by 20-30% yearly that is dependent on disease severity. Throughout the previous rare years, the disease had been continuously

informed in many nations across South Asia, Africa and many precisely in northwestern areas of India and Pakistan and also in China. Characteristic disease symptoms are curling of leaf, darkening of vein, swelling of vein, enation and greening of infested plants throughout the infection's initial phase; cup-shaped leaf-like structures were seen on leaves' undersides in severe infection (Fig.1). In harshly infected plants, internodal length decreases, that stunted the plant growth. The virus responsible for the disease belongs to the genus *Begomovirus* of family *Geminiviridae*, entirely spread by *Bemasia tabaci*, a whitefly vector (Abbas *et al;* 2015) [1].

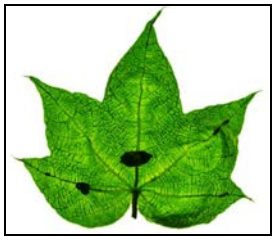


**Fig 1:** Greening on leaf, thickening on vein and enations on the base of CLCuD infected leaf

This study was conducted to do molecular and phylogenetic study of CLCuV tolerant cotton cultivars by using microsatellite markers. SSR markers are polymorphic nature bands in the DNA comprising repeating entity of 1-6 base pair. These repetitive sequences are present in the entire genome. The repeats of nucleotides in mono, di and tri series are called microsatellites. SSRs are mostly used in cotton due to raised likelihood for phylogenetics and genetic diversity study (Bolek *et al;* 2016) [3].

## Materials and Methods

Sixty six cotton varieties were obtained from Central Cotton Research Institute (CCRI) Multan. Their leaves were stored at -4°C to carry out further procedure. Plant's phenotypic data was also recorded like plant height, leaf length, leaf width, boll weight etc for our research. DNA extraction was done by Doyle and Doyle method with few modifications and PCR was performed with SSR markers.

For PCR, 31 SSRs primers were used. The PCR reaction was processed with PCR water (10.25 µl), dNTPs, PCR thermopol buffer, $MgCl_2$ (2.5 µl each), forward and reverse primers (1 µl), Taq DNA enzyme (0.25 µl), DNA sample (5 µl). Total reaction mixture was of 25 µl. The reaction was incubated at 94°C following 35 amplification cycles (94°C for 8 min, 55°C for 1 min, 72°C for 8 min). The ultimate PCR products were observed on polyacrylamide gel electrophoresis (PAGE).

### Statistical Analysis

Molecular and phylogenetic study of cotton was done on the basis of genetic data and morphological traits. Power marker v 3.25 software was applied on genetic data. Phylogenetic tree was generated by using UPGMA method based on Nei's 1973 similarity coefficient. Correlation analysis was performed to find out the association of morphological traits by using SPSS software. Principle component analysis (PCA), cluster analysis was also done by using same software.

## Results and Discussion

### Allele number evaluation

Thirty one SSR markers were used in order to do molecular study on 66 genotypes of *G. hirsutum.* Total 138 and average 4.4 loci were amplified by the markers. Maximum 8 loci were amplified by a single primer.

### Calculating Polymorphism, Gene diversity, Frequency of allele

The range of gene diversity was from 0.32 to 0.92. NAU-5418 and BNL-2709 (0.92) marker showed maximum value of gene diversity. NAU-5270 (0.32) showed minimum value of gene diversity.

Polymorphism information content (PIC) values represent polymorphism for each SSR marker. The range of PIC was from 0.27-0.92. BNL 2709 (0.92) marker showed maximum value of polymorphism while NAU 5270 (0.27) marker showed minimum value of polymorphism.

**Table 1:** Major frequency of allele, Polymorphism information content and genetic diversity

| SR. NO. | Primer | Major Allele Frequency | Allele NO | Gene Diversity | PIC |
|---|---|---|---|---|---|
| 1 | NAU 2954 | 0.45 | 4.00 | 0.59 | 0.51 |
| 2 | NAU 1070 | 0.52 | 7.00 | 0.63 | 0.58 |
| 3 | NAU 4042 | 0.41 | 12.00 | 0.77 | 0.74 |
| 4 | NAU 3414 | 0.39 | 7.00 | 0.72 | 0.67 |
| 5 | J 134 | 0.79 | 2.00 | 0.33 | 0.28 |
| 6 | NAU 3911 | 0.65 | 3.00 | 0.51 | 0.44 |
| 7 | NAU 5172 | 0.55 | 4.00 | 0.58 | 0.50 |
| 8 | NAU 5046 | 0.36 | 8.00 | 0.74 | 0.70 |
| 9 | NAU 5465 | 0.68 | 2.00 | 0.43 | 0.34 |
| 10 | NAU 4105 | 0.32 | 12.00 | 0.82 | 0.80 |
| 11 | NAU 5418 | 0.15 | 23.00 | 0.92 | 0.91 |
| 12 | NAU 998 | 0.73 | 5.00 | 0.44 | 0.40 |
| 13 | NAU 2437 | 0.64 | 7.00 | 0.54 | 0.49 |
| 14 | NAU 2714 | 0.70 | 3.00 | 0.46 | 0.41 |
| 15 | NAU 2838 | 0.18 | 25.00 | 0.90 | 0.89 |
| 16 | NAU 2868 | 0.27 | 11.00 | 0.79 | 0.76 |
| 17 | NAU 2095 | 0.26 | 7.00 | 0.80 | 0.77 |
| 18 | NAU 1366 | 0.45 | 8.00 | 0.71 | 0.67 |
| 19 | NAU 980 | 0.29 | 19.00 | 0.86 | 0.85 |
| 20 | NAU 5270 | 0.80 | 2.00 | 0.32 | 0.27 |
| 21 | DPL 0323 | 0.55 | 9.00 | 0.61 | 0.56 |

| | | | | | |
|---|---|---|---|---|---|
| 22 | BNL 827 | 0.30 | 20.00 | 0.86 | 0.85 |
| 23 | BNL 786 | 0.41 | 6.00 | 0.72 | 0.67 |
| 24 | BNL 2709 | 0.17 | 26.00 | 0.92 | 0.92 |
| 25 | J 274 | 0.18 | 21.00 | 0.91 | 0.90 |
| 26 | BNL 2449 | 0.62 | 9.00 | 0.58 | 0.55 |
| 27 | BNL 1672 | 0.79 | 12.00 | 0.37 | 0.37 |
| 28 | BNL 4096 | 0.65 | 2.00 | 0.45 | 0.35 |
| 29 | J 292 | 0.68 | 2.00 | 0.43 | 0.34 |
| 30 | J 110 | 0.73 | 6.00 | 0.45 | 0.43 |
| 31 | J 153 | 0.61 | 4.00 | 0.55 | 0.49 |
| | Mean | 0.49 | 9.29 | 0.64 | 0.59 |

**Phylogenetic Tree**

UPGMA tree method based on Nei' s 1973 similarity co-efficient was utilized in order to generate the phylogenetic tree which alienated all accessions into clusters, sub-clusters and sub sub-clusters. Two chief clusters A and C were formed further isolating into sub-clusters and sub sub-clusters.

40 accessions were included in cluster A forming a distinct group while 26 accessions were included in cluster C making a distinct group.

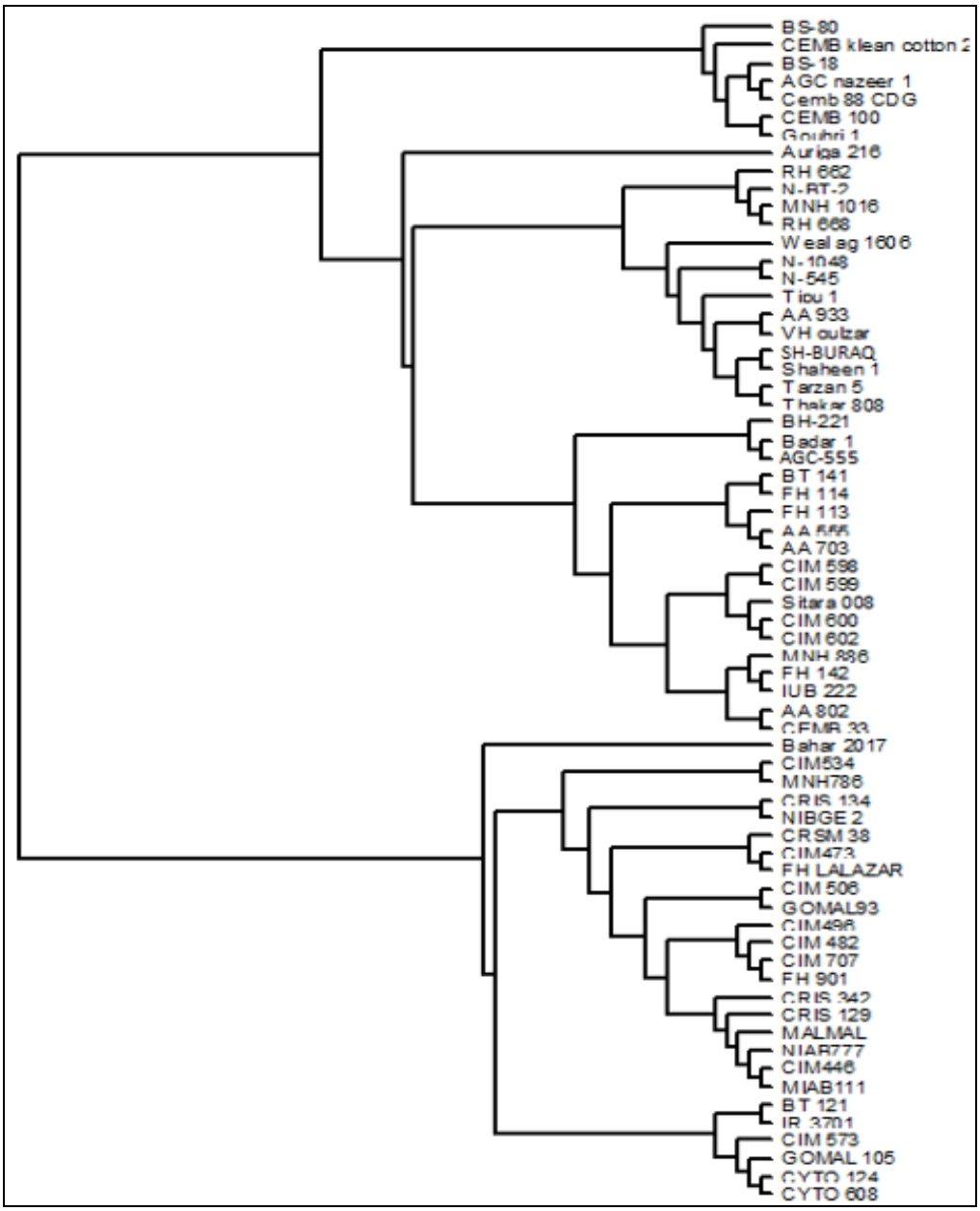



**Fig 2:** Bootstrap NJ rooted tree

**Correlation Analysis**

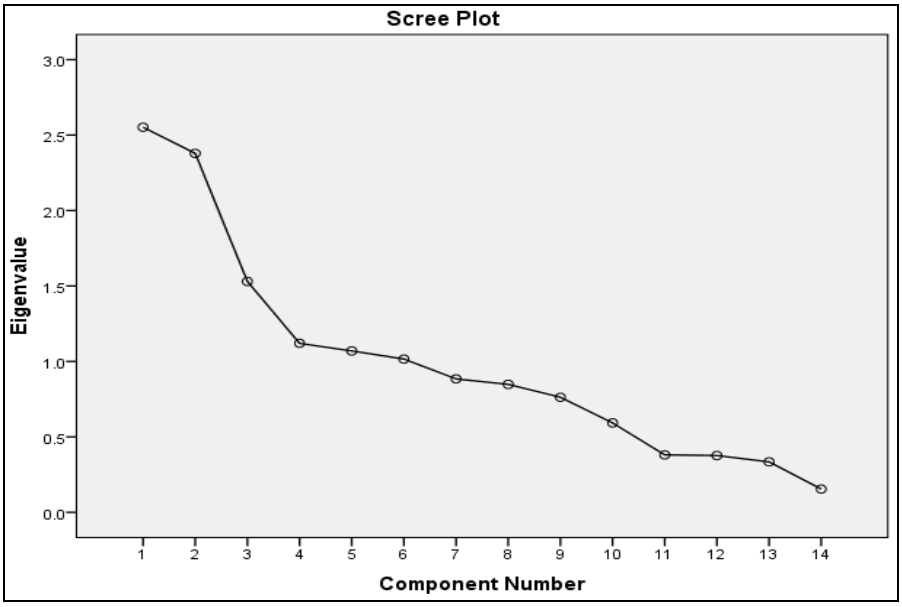
Simple correlation coefficients were utilized to study the relationship of 14 morphological traits. SPSS software was utilized to perform correlation analysis. In correlation analysis, nodes to first monopodia (FMP) and 1st sympodial node number (SN) showed maximum correlation with value of 0.783. Monopodia per plant (MP) and CLCuD showed minimum correlation with a value of -0.363.

**Table 2:** Simple correlation of 14 morphological traits of 66 accessed cotton genotypes

| | | PH | MP | FMP | SP | SN | LF | LW | PL | FP | LB | BP | SB | BW | CLCuD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PH | Pearson Correlation | 1 | .038 | .334 | .182 | .212 | -259 | -.267 | -.125 | -.054 | -246 | -.065 | -.037 | -.039 | .048 |
| | Sig. (2- tailed) | | .764 | .006 | .144 | .088 | .036 | .030 | .318 | .664 | .046 | .606 | .769 | .758 | .703 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| MP | Pearson Correlation | .038 | 1 | 241 | .037 | .298 | .107 | .010 | .140 | .237 | 245 | .128 | -.055 | .001 | -.363 |
| | Sig. (2-tailed) | .764 | | .051 | .768 | .015 | .394 | .939 | .262 | .055 | .047 | .308 | .662 | .997 | .003 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| FMP | Pearson Correlation | .334 | .241 | 1 | .297 | .783 | -.064 | -.091 | -.065 | .139 | .157 | -.247 | -.095 | -.210 | -.232 |
| | Sig. (2-tailed) | .006 | .051 | | .015 | .000 | .611 | .466 | .602 | .266 | .208 | .046 | .446 | .091 | .061 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| SP | Pearson Correlation | .182 | .037 | .297 | 1 | .130 | -.111 | -.264 | -.052 | .212 | -.186 | -.008 | -.068 | -.094 | .131 |
| | Sig. (2-tailed) | .144 | .144 | .768 | .015 | .298 | .357 | .032 | .677 | .088 | .134 | .951 | .587 | .455 | .294 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| SN | Pearson Correlation | .212 | 298 | .783 | .130 | 1 | .123 | .099 | -.135 | .106 | 288 | -241 | -.106 | -.125 | -275 |
| | Sig. (2-tailed) | .088 | .015 | .000 | .298 | | .325 | .428 | .279 | .398 | .019 | .051 | .399 | .319 | .025-. |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| LF | Pearson Correlation | -.259 | .107 | -.064 | -.111 | .123 | 1 | .584 | .377 | .086 | 281 | .161 | -.023 | .224 | .004 |
| | Sig. (2-tailed) | .036 | .394 | .611 | .375 | .325 | | .000 | .002 | .495 | .022 | .197 | .857 | .070 | .974 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| LW | Pearson Correlation | -.267 | .010 | -.091 | -.264 | .099 | .584 | 1 | .269 | .017 | .186 | .116 | .007 | .060 | -.044 |
| | Sig. (2-tailed) | .030 | .939 | .466 | .032 | .428 | .000 | | .029 | .890 | .136 | .354 | .957 | .635 | .727 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| PL | Pearson Correlation | -.125 | .140 | -.065 | -.052 | -.135 | .377 | .269 | 1 | .194 | .175 | -.011 | .000 | .057 | .043 |
| | Sig. (2-tailed) | .318 | .262 | .602 | ,677 | .279 | .002 | .029 | | .119 | .160 | .929 | 1.000 | .648 | .731 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| FP | Pearson Correlation | -.054 | .237 | .139 | .212 | .106 | .086 | .017 | .194 | 1 | .069 | .069 | -.084 | .237 | -.001 |
| | Sig. (2-tailed) | .664 | .055 | .266 | .088 | .398 | .495 | .890 | .119 | | .580 | .583 | .504 | .056 | .991 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| LB | Pearson Correlation | -.246 | .245 | .157 | -.186 | .288 | .281 | .186 | .175 | .069 | 1 | -.056 | .064 | .119 | -.101 |
| | Sig. (2-tailed) | .046 | .047 | .208 | .134 | .019 | .022 | .136 | .160 | .580 | | .654 | .608 | .342 | .420 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| BP | Pearson Correlation | -.065 | .128 | -.247 | -.008 | -.241 | .161 | .116 | -.011 | .069 | -.056 | 1 | -.099 | .173 | .277 |
| | Sig. (2-tailed) | .606 | .308 | .046 | .951 | .051 | .197 | .354 | .929 | .583 | .654 | | .430 | .164 | .024 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| SB | Pearson Correlation | -.037 | -.055 | -.095 | -.068 | -.106 | -.023 | .007 | .000 | -.084 | .064 | -.099 | 1 | .017 | .034 |
| | Sig. (2-tailed) | .769 | .662 | .446 | .587 | .399 | .857 | .957 | 1.000 | .504 | .608 | .430 | | .890 | .785 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| BW | Pearson Correlation | -.039 | .001 | -.210 | -.094 | -.125 | .224 | .060 | .057 | .237 | .119 | .173 | .017 | 1 | .160 |
| | Sig. (2-tailed) | .758 | .997 | .091 | .455 | .319 | .070 | .635 | .648 | .056 | .342 | .164 | .890 | | .200 |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |
| CLCuD | Pearson Correlation | .048 | -.363 | -.232 | .131 | -.275 | .004 | -.044 | .043 | -.001 | -.101 | .277 | .034 | .160 | 1 |
| | Sig. (2-tailed) | .703 | .003 | .061 | .294 | .025 | .974 | .727 | .731 | .991 | .420 | .024 | .785 | .200 | |
| | N | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 | 66 |

**Principle Component Analysis (PCA)**

PCA was done using SPSS software. Total 14 principle components were obtained out of which 4 have Eigen value greater than 1 while other 10 gave Eigen values lower than 1.




**Graph 1:** Eigen values of all traits

In Principle Component analysis, maximum variation was observed on the basis of plant height (PH) 18.23% in cotton genotypes. While minimum variation was on the basis of CLCuD 1.10% in cotton genotypes.

**Table 3:** Cumulative percentages and percentage of varience of principle components

| Component | Initial Eigenvalues | | |
|---|---|---|---|
| | Total | % of variance | Cumulative % |
| PH | 2.55 | 18.23 | 18.23 |
| MP | 2.38 | 16.99 | 35.22 |
| NFM | 1.53 | 10.93 | 46.15 |
| SPP | 1.12 | 8.00 | 54.15 |
| FSN | 1.07 | 7.64 | 61.79 |
| LL | 1.02 | 7.25 | 69.04 |
| LW | 0.88 | 6.32 | 75.36 |
| PL | 0.85 | 6.06 | 81.42 |
| FP | 0.76 | 5.45 | 86.86 |
| LB | 0.59 | 4.23 | 91.10 |
| BP | 0.38 | 2.72 | 93.82 |
| SB | 0.38 | 2.69 | 96.51 |
| BW | 0.33 | 2.39 | 98.90 |
| CLCuD | 0.15 | 1.10 | 100.00 |

**Cluster Analysis**

Cluster analysis was done by using SPSS software. Phenotypic traits divided the genotypes of cotton into different clusters. Genotypes having desirable phenotypic traits were separated into one cluster. 66 cotton genotypes were separated into 5 clusters on the basis of different phenotypic traits.

**Table 4:** Cluster analysis of 14 traits of cotton

| Traits | Initial Cluster Centres | | | | |
|---|---|---|---|---|---|
| | Cluster | | | | |
| | 1 | 2 | 3 | 4 | 5 |
| PH | 97.1 | 274.3 | 151.6 | 101.6 | 127.0 |
| MP | 2.6 | 2.2 | 2.7 | 2.6 | 2.3 |
| NFM | 3.4 | 8.8 | 11.6 | 10.2 | 7.3 |
| SPP | 19.3 | 15.7 | 19.0 | 11.7 | 17.8 |
| FSN | 6.9 | 9.7 | 12.8 | 11.3 | 9.0 |
| LL | 8.3 | 7.3 | 7.4 | 10.3 | 7.0 |
| LW | 9.1 | 8.0 | 7.9 | 11.7 | 7.8 |
| PL | 7.1 | 7.4 | 7.1 | 9.2 | 6.8 |
| FP | 2.7 | 2.0 | 2.0 | 2.4 | 2.7 |
| LB | 4.1 | 3.6 | 3.4 | 4.2 | 3.8 |
| BP | 37.0 | 28.1 | 32.0 | 19.4 | 22.8 |
| SB | 28.9 | 27.2 | 27.4 | 26.3 | 24.2 |
| BW | 2.8 | 2.9 | 2.6 | 2.8 | 2.7 |
| CLCuD | 4.1 | 5.0 | 3.2 | 3.4 | 2.7 |

Cluster analysis designated that three genotypes were included in cluster 1. Only one genotype was included in cluster 2 making it diverse of all. Cluster 3 confined nine genotypes. Ten genotypes were included in cluster 4. Forty-three genotypes were included in final fifth cluster.

**Table 5:** Cluster membership of various genotypes

| Clusters | Member genotypes |
|---|---|
| Cluster:1 | CEMB-33, A-555 and IR-3701, |
| Cluster:2 | Tarzan-5 |
| Cluster:3 | CIM-602, CIM-573, Cyto-124, Cyto-608, Gomal-105, CIM-342, Malmal, NIAB-777 and CIM-496 |
| Cluster:4 | Bahar-2017, N-BT-2, CIM-600, Sitara-008, IUB-222, AA-802, MNH-886, AA-703, FH-113 and CIM-446 |
| Cluster:5 | CEMB-88 (DG), AGC-Nazeer-1, Ghauri-1, CEMB-Klean cotton-2, CEMB-100, BS-80, BS-18, BH-221, Badar-1, AGC-555, Auriga-216, AA-933, Wealag-1606, VH-gulzar, Tipu-1, Thakkar-808, SH-Buraq, Shaheen-1, RH-662, RH-668, N-1048, N-545, MNH-1016, CIM-598, CIM-599, FH-142, FH-114, BT-141, BT-121, FH Lalazar, NIAB-111, CRIS-129, CIM-707, FH-901, CIM-482, CIM-473, CIM-534, MNH-786, CIM-506, CRSM-38, Gomal-93, CRIS-134 and NIBGE-2 |

We applied Power marker v 3.25 in our study. 31 SSR markers amplified total 138 alleles. Average number of alleles was 4.4. NAU-2838, NAU-980, BNL- 827 markers gave maximum 8 alleles which represented high polymorphism. JESPER-292, JESPER-134, NAU-5465, NAU-5270, BNL-4096 markers gave minimum 1 allele. BNL-2709, NAU-5418 marker showed maximum gene diversity of 0.92. NAU-5270, J-134 gave minimum gene diversity of 0.32, 0.33. Gene diversity ranged from 0.32-0.92. Average gene diversity was 0.64. PIC represent level of polymorphism shown by each SSR marker. BNL-2709, NAU-5418 marker represented maximum PIC of values 0.92, 0.91. J-134, NAU-5270 marker represented minimum PIC of values 0.28, 0.27. Range of polymorphism was 0.27-0.92. Average polymorphism was 0.59. Similar type of results was showed by previous studies (Kuang *et al;* 2014; Wang *et al;* 2011) [9, 12]

Phylogenetic relationship was observed using Power marker v 3.25. Similarity index was determined by Bootstrap neighbour joining method and UPGMA based on Nei 1973

similarity index. This tree divided the 66 accessions into two major clusters according to their phylogenetic similarity. Our results revealed that 40 accessions fall in a group while 26 accession fall in another group. This means that those 40 have genetic association with each other and 26 have genetic association with each other. This type of relationship was showed in previous studies by (Abbas *et al;* 2015) [1]

Before starting any breeding programme the information about association among various traits is a prerequisites as it gives an opportunity for selection of genotypes having desirable traits simultaneously. Phylogenetic study also involved SPSS analysis. Correlation, PCA, cluster analysis was done. In correlation analysis, association among some traits were found. Nodes to first monopodia (FMP) and 1st sympodial node number (SN) showed maximum correlation with value of 0.783. Monopodia per plant (MP) and CLCuD showed minimum correlation with a value of -0.363. Similar results were showed by (Saeed *et al;* 2013) [11]. In PCA, total 14 principle components were obtained out of which 4 have Eigen value greater than 1 while other 10 gave Eigen values lower than 1. In PCA, maximum variation was observed on the basis of plant height (PH) 18.23% in cotton genotypes. While minimum variation was on the basis of CLCuD (1.10%) in cotton genotypes. This type of relationship was showed in previous studies by (Saeed *et al;* 2014) [10]. PCA is an important tool for identifying parental lines for successful breeding programmes. In cluster analysis, phenotypic traits divided the genotypes of cotton into different clusters. Genotypes having desirable phenotypic traits were separated into one cluster. 66 cotton genotypes were separated into 5 clusters on the basis of different phenotypic traits. Cluster analysis designated that three genotypes were included in cluster 1. Only one genotype was included in cluster 2 making it diverse of all. Cluster 3 confined nine genotypes. Ten genotypes were included in cluster 4. Forty-three genotypes were included in final fifth cluster. Such type of results were also showed in previous studies carried out by (Amna *et al;* 2013)

## Conclusion

In our study, gene diversity ranged from 0.32-0.92 and PIC ranged from 0.27-0.92. Genotypically, Bahar 2017 variety was diverse and phenotypically Tarzan 05 was diverse from all. Nodes to first monopodia (FMP) and 1st sympodial node number (SN) showed maximum correlation. Maximum variation was observed on the basis of plant height while minimum variation was on the basis of CLCuD in cotton genotypes. We could use these diverse varieties to bring diversity in other varieties because diversity is the main demand for breeding purpose and in agriculture. We could choose parents for bringing diversity through crosses. These informative markers could also be used for evaluation of diversity in other varieties.

## References


1. Abbas A, Iqbal MA, Rahman MU, Paterson AH. Estimating genetic diversity among selected cotton genotypes and the identificationof DNA markers associated with resistance to cotton leaf curl disease. Turkish Journal of Botany,2015:39:1033-1041.
2. Nazir A, Farooq J, Mahmood A, Shahid M, Riaz M. Estimation of genetic diversity for CLCuV, earliness and fiber quality traits using various statistical procedures in different crosses of Gossypium hirsutum L. Вестник аграрной науки,2013:43(4):2-9.
3. Bolek, Yuksel, Bhatti, Khezir, Bardak, Adem, Azhar *et al*. Insight in the utilization of Marker Assisted Selection in Cotton (A Review). Molecular Plant Breeding. 10.5376/mpb,2016, 07(0010).
4. Boopathi N, Sathish SKD, Ravikesavan R Molecular Breeding for Genetic Improvement of Cotton (Gossypium spp.), 2015. 10.1007/978-3-319-22521-0_21.
5. Farooq M, Shaheen T, Rizwan M, Shakeel Q, Shahid MR, Kainth RA *et al*. Screening of different advanced genotypes of transgenic cotton (Gossypium hirsutum L.) against thrips (Thrips tabaci Lindemann.) under Faisalabad Agro-ecosystem. Journal of entomology and zoology studies,2018a:6:1275-1279.
6. Farooq M, Shaheen T, Shahid MR, Abbas HG, Rizwan M, Kainth RA *et al*. Response of different advanced genotypes of transgenic cotton (Gossypium hirsutum L.) against Amrasca biguttula biguttula (ishida) (Homoptera: Cicadellidae). Journal of Entomology and Zoology Studies,2018b:6:818-822.
7. Huang C, Nie X, Shen C, You C, Li W, Zhao W *et al*. Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNP s. Plant biotechnology journal,2017:15:1374-1386.
8. Khalil H, Raza ABM, Afzal M, Aqueel MA, Khalil MS, Mansoor MM. Effects of plant morphology on the incidence of sucking insect pests complex in few genotypes of cotton. Journal of the Saudi Society of Agricultural Sciences,2017:16:344-349.
9. Kuang M, Yang W, Wang F, Xu H, Wang Y, Zhou D *et al*. Screening of highly informative and representative microsatellite markers for genotyping of major cultivated cotton varieties. Genetics and Molecular Research,2014:13:9777-9786.
10. Saeed F, Farooq J, Mahmood A, Hussain T, Riaz M, Saghir A. Genetic Diversity in upland cotton for cotton leaf curl virus disease earliness and fiber quality. Pakistan Journal of Agricultural Research,2014:27:226-236
11. Saeed F, Farooq J, Mahmood A, Riaz M, Hussain T, Majid A. Estimation of genetic variability for Cotton leaf curl virus disease, fiber quality and some morphological traits using multivariate analysis in exotic Gossypium (Diploid and tetraploid) species. Journal of Agrobiology,2013:30:107-115.
12. Wang, X, Feng C, Lin Z, Zhang X. Genetic diversity of sea-island cotton (Gossypium barbadense) revealed by mapped SSRs. Genet Mol Res,2011:10:3620-3631.