Integrative
Plant Biotechnology
ISSN: 3007-2735 (Online), 3007-2727 (Print)

Check for updates

ROOTS PRESS

Research Article

Genomic Diversity Analysis and Association Mapping of Cotton Germplasm for Cotton Leaf Curl Disease (CLCuD) Resistance Using SSR Markers in Two Hotspots of Punjab, Pakistan

Asia Batool[1], Jehanzeb Farooq[2]*, Muhammad Umar Shahbaz[3], Amjad Farooq[2], Ghulam Sarwar[4]

[1] *Pulses Research Institute, Ayub Agricultural Research Institute (AARI), Faisalabad, Pakistan.*
[2] *Cotton Research Station, Ayub Agricultural Research Institute (AARI), Faisalabad, Pakistan.*
[3] *Plant Pathology Research Institute, AARI, Faisalabad, Pakistan.*
[4] *Cotton Research Institute, Multan, Pakistan.*

ABSTRACT

The cotton crop is the major contributor to Pakistan's economy. In the recent past, cotton crops faced many issues including leaf curl virus disease which resulted in decreased cotton production in Pakistan. The present research work was designed to discover tolerance/resistance sources through extensive experimentation using conventional and biotechnological tools. For this purpose, 71 genotypes from all over Pakistan and some lines from India and Turkey were sown at two hotspots, i.e., Faisalabad and Vehari in the month of June to find superior genotypes that can tolerate/resist this threatening disease. The sick plot method of screening was adopted for screening genotypes against CLCuD. The variety FH-118 was used as a spreader line in this experiment. For genomic analysis of cotton germplasm for cotton leaf curl disease (CLCuD) resistance using SSR markers in two hotspot regions of Punjab, 25 SSR primers were utilized to estimate the genetic variation among cotton genotypes and 12 were identified as highly polymorphic. Primers BNL-1551, JESPER-247, NAU-1103, NAU-1230, NAU-1233, NAU-1369, NAU-2152, NAU-2317, NAU-2336, NAU-2355, NAU-2503, NAU-2691 provided PIC values of 0.51, 0.57, 0.52, 0.64, 0.62, 0.55, 0.50, 0.63, 0.57, 0.56, 0.64 and 0.55 respectively. The maximum genetic similarity coefficient value was found to be 6.85 between FH-479 and NS-161 which showed that these genotypes might have common ancestors or same morphological traits while lowest coefficient value was found 2.449 between genotypes FH-315 and Tarzan-5 which indicated that they have different parental origin or distinct morphological characteristics. Based on similarity of genotypes dendrogram clustered all genotypes into 37 clusters. Most diverse genotype identified was FH-315 which can be used as a parent in future breeding programs. Association mapping studies revealed that five markers NAU3254-1600, dPL0526-260, NAU5163-200, CIR094-700, and cgr6356-150 are associated with CLCuD resistance at CRS, Faisalabad. While At CRS, Vehari the marker NAU3385-500 is different while 4 others are common at both locations. The study demonstrates that DNA-based molecular markers like SSR associated with CLCuD resistance could be used for screening resistant/tolerant cotton lines.

Keywords: Cotton, CLCuD, SSR markers, Genomic diversity, Sick plot technique

OPEN ACCESS

***Correspondence**
Jehanzeb Farroq
drjehanzeb.pbg@gmail.com

Article History
Received: November 03, 2024
Accepted: November 16, 2024
Published: December 01, 2024

Cite this article
Batool, A., Farooq, J., Shahbaz, M. U., Farooq, A., Sarwar, G. (2024). Genomic Diversity Analysis and Association Mapping of Cotton Germplasm for Cotton Leaf Curl Disease (CLCuD) Resistance Using SSR Markers in Two Hotspots of Punjab, Pakistan Integrative Plant Biotechnology, 02, 01-16.

INTRODUCTION

Cotton is the mainstay of Pakistan’s economy and is often stated as white gold. Cotton is a vital cash crop and exported cotton products contribute largely to Pakistan's foreign exchange earnings (Ali et al., 2019). During 2023-24, the area under cotton planting showed a growth rate of 13.1 %, and its production showed a remarkable boost of 108.2 percent to 10.2 million bales (Pakistan Economic Survey, 2023-24).

It contributed 0.7 % to GDP and 2.9% to agriculture value addition. Pakistan's textile industry depends entirely on cotton crops due to its fiber (Ali and Khan, 2007). Many people are engaged in the cotton industry and approximately 1.5 million people are associated with the cultivation of cotton (Hussain *et al*., 2010).

There are about 75 important single-stranded DNA viruses which belong to cotton diseases (Watkins, 1981), and among these, CLCuD has been the most threatening, especially in the Indo-Pak region, for about the last 50 years (Sattar *et al*., 2013). This viral disease has incurred enormous losses to Pakistan's economy (Farooq *et al*., 2011; Ali *et al*., 2019). It is caused by a complex of single-stranded DNA viruses belonging to the genus Begomovirus and family Geminiviridae (Mathews, 1987).

Genetic differences are valuable for planning any tolerance enhancement program aiming at the development of varieties against disease. Some scientists like Li et al. (2008) recommended that exploitation of older germplasm and accumulation of new genotypes will allow developing genetic changeability which may be subjected to selection pressure for probing disease tolerance against CLCuD. These variations in genotypes not only improve the chances of developing resistance but also can develop desirable plant types that can be manifested in the establishment of broad-spectrum breeding programs (Van Esbroeck and Bowman, 1998). Genetic variations provide the basis for crop improvement (Ali *et al*., 2009). DNA markers are mainly used to find true resistant/tolerant cotton germplasm against different diseases. The genetic distance between a DNA marker and a trait is measured through linkage mapping (Rahman *et al*., 2017). The documented DNA markers can be utilized for starting marker-assisted selection for disease resistance in plants (Ali *et al*., 2019; Jabran *et al*., 2023). For many reasons, utilization of DNA markers for determining development of cotton cultivars tolerant against CLCuD and its causal agents is very important.

Firstly, field screening under natural conditions is time consuming and mainly depends upon the vector population which varies due to variable environmental conditions (Jabran *et al*., 2023). Secondly, the inoculum of viruses including alternative hosts, weeds and many other factors may also influence the genotypic responses of cotton toward disease and virus strains. Humidity and temperature and response of border plants are different as compared to plants inside (non-border plants) the cotton field (Rahman *et al*., 2017). Likewise, if cotton is planted near to the fruit orchards, it creates humidity which may result in severe disease symptoms. Disease imposition in greenhouse is extremely difficult because of variations in the prevailing microclimate. So, the above-mentioned factors forced the researchers to exploit DNA markers as they are more desirable and less time consuming and also helpful in finding genotypes that are resistant to or tolerant to viral diseases. DNA markers provide good results in short time with more precision than field screening (Rahman *et al*., 2017). Different kinds of DNA markers are extensively used in cotton for genetic diversity calculation including RFLP, AFLP, RAPD, SSRs and ISSRs (Tyagi et al., 2014). Among all these DNA markers, SSRs are mainly used markers in molecular studies. Mostly SSR markers are used due to high polymorphism.

Simple Sequence Repeats (SSRs) are also called microsatellites and are unique in nature. These are short tandem repeats containing two to six base units and are robust because of their high reproducibility and co-dominance nature. SSRs are used to associate different quantitative trait loci (QTLs) in segregating populations in cotton (Yu *et al*., 2013). Other molecular markers, e.g., ISSR and RAPD, co-dominant in nature, do not show a higher level of polymorphism and reproducibility for genetic diversity (Dongre and Parkhi, 2005). The present study aims to assess genetic diversity in 71 exotics as well as local cotton germplasm accessions using SSR markers and the association of these SSR markers with seed cotton yield and CLCuD resistance.

MATERIAL AND METHODS

Plant cultivation

The cultivation of 71 germplasm accessions at two hot spots (CRS, Faisalabad, and Vehari) was done using a sick plot approach following Batool *et al*. (2024). This is the same set of germplasm accessions that were used in our previous study (Batool *et al*., 2024), which was used here to study the genetic diversity and association mapping of **CLCuD resistance and seed cotton yield with SSR markers.**

Genetic diversity of local and exotic cotton germplasm related with begomoviral tolerance using SSR markers

Sample collection

In this experiment, the diversity of all 71 collected genotypes was studied. Leaf samples were collected from all genotypes 70 days after sowing (DAS) from the field. Samples were placed in the ice and shifted to the laboratory for DNA isolation to perform polymerase chain reaction (PCR).

DNA isolation

Isolation of high-quality DNA from research samples is the basic material for the identification of different elements in crops based on Polymerase Chain Reaction (PCR). An array of methods is used to separate genomic DNA from plant material using commercially available kits. All the methodologies used follow the same basic concept of extraction of DNA which includes the removal of proteins, polysaccharides, pigments, RNA, and subsequent purification of DNA from PCR inhibitors. A series of steps typically obtain the separation of genomic DNA from plant materials. Total genomic DNA from the leaves of the respective varieties was extracted from each variety grown in the field. With the aid of the grinder, leaf sample samples were ground. As described below, 0.1 gram of each ground sample for DNA isolation was taken using the CTAB process. DNA was extracted from each grinded leaf sample using modified CTAB (CetyleTrimethyl Ammonium Bromide) method (Doyle and Doyle, 1987). The water bath was switched on and set at 65 °C to pre-heat 2XCTAB with 1% β-mercaptoethanol. A grinded sample weighing 0.1 was taken and 750 uL of pre-warmed CIAB was added and incubated in water bath for 1 hour. An equal volume of phenol-chloroform / isoamyl alcohol (25:24:1) was added and gently mixed by inverting the tube into an emulsion form. The centrifugation was performed at 13,500 rpm for 15 minutes. The supernatant solution (the top aqueous layer) was moved to a new Eppendorf 1.5mL tube, and the remaining chloroform layer was discarded. Added equivalent amounts of chloroform/isoamyl alcohol (24:1). It was gently combined with inverting the tube to create an emulsion. The centrifugation was done at 13,500 rpm for 10 minutes. The second supernatant (the top aqueous layer) solution was moved to a new 1.5 mL Eppendorf tube, and the remaining chloroform liquid was discarded. Chilled 2-propanol was used to precipitate the DNA (0.6 volumes). Centrifugation was performed for 10 minutes at 13,000 rpm; the supernatant was discarded, and the pellet was washed twice with 70 percent ethanol. The pellet was dried air and re-suspended in TAE or d2H2O buffer.

DNA quantification

The isolated DNA was measured with a spectrophotometer (Nano drop, Bio-Rad). The agarose gel was used to determine further to determine Isolated DNA consistency further. DNA content was evaluated by running samples of DNA on the electrophoresis of 1 percent agarose gel. Under UV light the gel was visualized for DNA content estimation. Sharp and clear bands that exhibited good DNA quality were seen. The samples that didn't have good bands were turned down.

Polymerase chain reaction (PCR)

Dilution was made prior to PCR. DNA samples were diluted to 30ng / µl in line with formula C1V1 = C2V2. The different volume of DNA was taken from DNA preps, and the volume for the necessary concentrations was made up to 100µl with the double distilled ionized water (d_3H_2O). Based on their abundance in the cotton genome and polymorphism material, the Primer pairs were selected from CMD in the primordial sequencing. The PCR conditions were designed for template DNA, 10X PCR $MgCl_2$, buffer, dNTPs, reverse and forward primers, and Taq DNA polymerase concentrations. Gradient PCRs have also optimized the annealing temperatures for all the SSR primers. The 50ng / µL working concentration was prepared by diluting the stock DNA for PCR and optimizing the protocol by using a single set of primers (Table 1) and DNA for temperature annealing.

The PCR reaction mixture was prepared with a final volume of 25µL using the following reagents. Forward and reverse primers (0.4 µM) 0.5µl each, DNA template (250 ng) 1µl, DNA Taq polymerase (0.5 U/µl) 0.5µl, 10X PCR buffer 2.5µl, $MgCl_2$ (2 mM) 2µl, dNTPs (0.2 mM) 0.5µl and d_3H_2O 17.5µl. Qcycler Satellite Gradient 1 Thermal Cycler was used to carry out the PCR, and the following protocol was used. Initial denaturation was carried out at 95^{o}C for 5 min followed by 35 cycles of Denaturation at 94^{o}C for 1 min, Annealing at 52^{o}C for 1 min, and extension at 72^{o}C for 1 min. The final extension was done at 72^{o}C for 5min.

Gel electrophoresis

Horizontal gel electrophoresis with agarose (2%) was used to analyze the polymerase chain reaction amplicons. The gel reporting system was used to imagine the amplified sample, compared to the 50bp DNA ladder in the gel for the PCR test. The PCR products were analyzed using an ethidium bromide staining solution and polyacrylamide gel electrophoresis. The gel tool was washed with tap water and then rinsed with double distilled deionized water. Gel tiles, spacers, and combs with newspaper were also washed and dried. Gel apparatus was mounted, taking note that both gel plates were even at the bottom. In the inner layer, notched plates were set, and the plates were secured firmly from the layer screws. Four liters of TBE buffer were packed, and this buffer filled 1/3 of the tank. In a sterile beaker, 52.5 ml 0.5X TBE buffer was taken: 29% acrylamide stock solution: 1% till acrylamide was prepared in a mixture form. Acrylamide / bisacrylamide mixture 22.5ml was removed from stock solution and applied to the 0.5X

TBE buffer. To it is added 700µl 10 per cent APS and 50µl TEMED. Gel was gently mixed and poured into the apparatus by tilting the tool in a slant position. Combs were carefully placed into the gel, and they allowed the gel to solidify for 15-20 minutes. The gel was raised by losing the side screws and put within the tank of the apparatus. The 2-4 µl PCR drug was poured into the wells and was tested at 14 watts for 1, 5-2, 5 hours roughly. Gel was extracted slightly and placed for 20 minutes in the fixative solution. After that the gel was moved for 10 minutes to staining solution. Then it was put for 4-10 minutes depending on the presence of the band in developing solution. When electrophoresis was done, UV Trans illuminator observed gels to analyze banding pattern under UV light. Gels have been shot using documentation device Syne Gene Gel. Based on the known ladder size the size of each sample was recorded.

Assessment of CLCuD incidence percentage

All the 71 genotypes were screened using sick plot technique in June (Iqbal *et al*., 2014) at two locations in Punjab i.e., Faisalabad and Vehari. The 71 genotypes were the experimental material that was planted in a randomized complete block design. The variety FH-118 was used as a spreader line in this experiment. The disease (%) of cultivars was measured using the disease scale described by Akhtar *et al*. (2010) and Farooq *et al*. (2011). Then %age of CLCuD incidence was calculated by using the following formula.

$$\text{CLCuD disease incidence (\%)} = \frac{\text{Sum of all disease ratings}}{\text{Total number of plants}} \times 16.16$$

The data for CLCuD% are reported in our previous work (Batool *et al*., 2021) and were used here to associate with SSR markers.

Statistical analysis

Based on the presence (1) and absence (0) of bands, a typical statistical approach has been used to measure the allele frequencies, heterozygosity and the adequate number of alleles for all the genotypes (Liu *et al*., 2000). Ambiguous bands which could not be clearly differentiated were not scored. Analysis of the clusters was carried out using NTSYS-PC software (Exeter Software, NY, USA; Rohlf, 2000). Dendrograms were constructed using the Arithmetic Averages (UPGMA) method using the Unweighted Pair Group method. In addition, the study of genetic variation and clustering was performed using NTSYS-PC software (Exeter Software, NY, USA; Rohlf, 2000). The analysis of association mapping revealed that to fit the data to the simple linear regression model this equation was used, y = b0 + b1 x + e. For this purpose, single marker analysis was done (Wang et al. 2010). The results obtained gave the estimates for b0, b1 and the F statistics for each marker. We were interested in whether the marker is linked to QTL. We tested this idea by determining if b1 is significantly different from zero. The F statistics compare the hypothesis H0: b1 = 0 to an alternative H1: b1 not 0. The pr (F) is a measure of how much support there is for H0. A smaller pr (F) indicates less support for H0 and, thus, more support for H1. Significance at the 5%, 1% and 0.1%, levels are indicated by * (Significant) , ** (Highly Significant) , and *** (Very highly Significant), respectively.

Table 1. Combine variance components for morphological and physiological traits among advanced wheat lines under normal and heat stress states.

Sr. No.	Marker Name	Amplicon Size	Chromosome	Primers Sequence
1	NAU-2691 F	225 bp	Chr # 3	5'TCACATCTTGCAAGCTCATT3'
2	NAU-2691 R			5'AGTTAAAACCGGGCTGAGAT3'
3	NAU-2697 F	158 bp	Chr # 13	5'ATTTCCCATGGTCATAGCAG3'
4	NAU-2697 R			5'GAAAGGAGTCGGAAATGAGA3'
5	NAU-2503 F	217 bp	Chr # 19	$^{5'}$GATCGAAATCAAAGGGCTTA$^{3'}$
6	NAU-2503 R			$^{5'}$CGTTGGATGAAGTTGATGAT$^{3'}$
7	NAU-1369 F	242 bp	Chr # 2	$^{5'}$TGGCAGAGATGAATGTAAGC$^{3'}$
8	NAU-1369 R			$^{5'}$GGTAACGGATGGAAAATCAC$^{3'}$
9	NAU-2083 F	168 bp	Chr # 1	$^{5'}$AGAAGAGGTTGACGGTGAAG$^{3'}$
10	NAU-2083 R			$^{5'}$TGAGTGAAGAACCTGCACAT$^{3'}$
11	NAU-2265 F	224 bp	Chr # 2	$^{5'}$CAATCACATTGATGCCAACT$^{3'}$
12	NAU-2265 R			$^{5'}$CGGTTAAGCTTCCAGACATT$^{3'}$
13	NAU-2162 F		Chr # 4	$^{5'}$ACACAAAAACCCAAAGGAAA$^{3'}$

14	NAU-2162 R			$^{5'}$CACGAGTGTCCTTGCTACAG$^{3'}$
15	NAU-2169 F	244 bp	Chr # 24	$^{5'}$AGCAAATCCAATCACTTGGT$^{3'}$
16	NAU-2169 R			$^{5'}$AAAAGTAGCGGGATGAGATG$^{3'}$
17	NAU-2355 F	249 bp		$^{5'}$ACAAACAAAACGCCTTCTTC$^{3'}$
18	NAU-2355 R			$^{5'}$AACACAAAAACGGTTCCAGT$^{3'}$
19	NAU-2336 F	196 bp	Chr # 14	$^{5'}$TGGAAAAGGAAGAGGAGAGA$^{3'}$
20	NAU-2336 R			$^{5'}$CCCTGAAGTTGTCAAGCTCT$^{3'}$
21	NAU-2317 F	181 bp	Chr # 10	$^{5'}$GACTCCAGCCTTCACACAT$^{3'}$
22	NAU-2317 R			$^{5'}$TGGAAGAGTATAACGGCAGA$^{3'}$
23	BNL-1551 F	184 bp	Chr # 21	$^{5'}$CGCAAGCCACCTGTAAAAC$^{3'}$
24	BNL-1551 R			$^{5'}$TCGAATTTTCTCTCTCTCTCTCTCT$^{3'}$
25	NAU-1103 F	181 bp	Chr # 11	$^{5'}$GGAGCCAGAAGTTGAGAAAA$^{3'}$
26	NAU-1103 R			$^{5'}$TTCGGCTTCTGCTTTTACTT$^{3'}$
27	NAU-1141 F	210 bp	Chr # 13	$^{5'}$CCCCTCTCTCTGTTTCTCAA$^{3'}$
28	NAU-1141 R			$^{5'}$AAGGGGTTTGAAGGGTTATC$^{3'}$
29	NAU-1233 F	245 bp	Chr # 11	$^{5'}$TTCGGGAAAGTTAGAGGAGA$^{3'}$
30	NAU-1233 R			$^{5'}$TCCTCAGAGCTCGGAATAGT$^{3'}$
31	NAU-1366 F	222 bp	Chr # 21	$^{5'}$CATGAAGCTTTTCCCACTTT$^{3'}$
32	NAU-1366 R			$^{5'}$CAGCTTATCCACCCCTAATG$^{3'}$
33	NAU-1230 F	247 bp	Chr # 5	$^{5'}$CATGCAAATCCATGCTAGAG$^{3'}$
34	NAU-1230 R			$^{5'}$TCAAAAGGTTCTTTGGTGGT$^{3'}$
35	NAU-5091 F	174 bp	Chr # 21	$^{5'}$AGTGCAGACAAATGCAAAGA$^{3'}$
36	NAU-5091 R			$^{5'}$CTGGCTATGATGATGAGGAG$^{3'}$
37	JESPR-247 F	157 bp	Chr # 6	$^{5'}$GCTTCTTCCATTTTATTCAAG$^{3'}$
38	JESPR-247 R			$^{5'}$CAGCGGCAACCAAAAAG$^{3'}$
39	NAU-2139 F	165 bp		$^{5'}$ATCCTACCCCTACCCTGAGA$^{3'}$
40	NAU-2139 R			$^{5'}$GATGGAGGAGGACATCAGAC$^{3'}$
41	NAU-3608 F	228 bp	Chr # 16	$^{5'}$GCTATTTCTGATTCTGGGTCTC$^{3'}$
42	NAU-3608 R			$^{5'}$AAAACCCAGTCTTTTCTTTCC$^{3'}$
43	NAU-3654 F		Chr # 7	$^{5'}$TTACCAGCAGCCAACACTAA$^{3'}$
44	NAU-3654 R		Chr # 7	$^{5'}$TCCCCTTCAACATCTTCTTC$^{3'}$
45	NAU-5005 F	262 bp	Chr # 19	$^{5'}$AAGGTAGGAAGCAATGCAAC$^{3'}$
46	NAU-5005 R			$^{5'}$AAAACATGTAGGAACGAGCA$^{3'}$
47	NAU 2152 F	198 bp	Chr # 21	$^{5'}$GAAAACAAGGCACTTGAACC$^{3'}$
48	NAU 2152 R			$^{5'}$GATCGATCAAAACCGGTAAC$^{3'}$
49	NAU 3529 F	224 bp	Chr # 19	$^{5'}$TAAACAATCCATGGTTCACG$^{3'}$
50	NAU 3529 R			$^{5'}$CCGATGAAGATTCACTGACA$^{3'}$

RESULTS

Diversity of cotton genotypes using SSR markers

Gel electrophoresis

The present study was conducted to analyze the genetic diversity among various cotton genotypes through SSR markers. 25 SSR primers were used to analyze the 71 cotton genotypes using PCR. Results were obtained by polyacrylamide gel electrophoresis, as shown in Figure 1. The primers with their respective PIC values are given in Table 2.

Table 2. Primers on the basis of PIC values.

Sr.No.	PIC=0-0.3	PIC=0.31-0.6	PIC=0.61>
1	NAU-1366	BNL-1551	NAU-1230
2	NAU-2139	JESPER-247	NAU-1233
3	NAU-2162	NAU-2083	NAU-2317
4	NAU-2169	NAU-1141	NAU-2503
5	NAU-2265	NAU-1369	
6	NAU-5005	NAU-2152	
7		NAU-2336	
8		NAU-2355	
9		NAU-2691	
10		NAU-2697	
11		NAU-3529	
12		NAU-3608	
13		NAU-3654	

A total of 75 alleles were identified among 71 genotypes and allele sizes ranged from 90 to 450. However, the polymorphic information content (PIC) value varied from 0.1 to 0.6. A maximum number of primers proved to be efficient in determining genetic variability as they have PIC values of 0.5 and greater than 0.5. Out of 25 primers, 12 were identified as highly polymorphic. Primers with PIC values of more than 0.5 are considered polymorphic and with less than 0.5 are considered as monomorphic.

Monomorphic are not used in genetic diversity estimation in genotypes. For instance, primers BNL-1551, JESPER-247, NAU-1103, NAU-1230, NAU-1233, NAU-1369, NAU-2152, NAU-2317, NAU-2336, NAU-2355, NAU-2503, NAU-2691 provided PIC values of 0.51, 0.57, 0.52, 0.64, 0.62, 0.55, 0.50, 0.63, 0.57, 0.56, 0.64 and 0.55 respectively. It is stated that makers with a PIC value of 0.5 or greater than it has a high potential to determine genetic diversity while markers with a PIC value between 0.3 and 0.4 have moderate efficiency in the determination of genetic variations.

Band scoring

After PCR products run on a gel, bands on the gel were scored as 1 for presence and 0 for the absence of band. Scoring was conducted only by minimum and maximum allele size in a genotype amplified by a specific primer. Scoring is given in tables.

The table represents that maximum genetic similarity coefficient value was found 6.85 between genotypes 57 (FH-479) and 1 (NS-161) which shows that may have common ancestors or same morphological traits while lowest similarity coefficient value was found 2.449 between genotypes 30 (FH-315) and 31 (Tarzan-5) which indicate that they have different parental origin or distinct morphological characteristics.

Dendrogram

Based on scoring the following dendrogram was constructed by MSVP. This shows two main groups of all genotypes, and one genotype is quite different from the rest. These groups are further divided into subgroups. This dendrogram contains 8 clusters. The amplified size by each primer, their allelic frequency in population, population size in which allele was found, and polymorphic information content are given in Table 3. The dendrogram constructed by using the Arithmetic Averages (UPGMA) method using the Unweighted Pair Group method showing genetic diversity is given in Figure 2. The dendrogram demonstrates that genotype number 47 is much different from each other fall separate in cluster which means quite different genetic makeup then 24 and 25 of same cluster, 24 and 25 are quite similar to each other. While the genotype 44 is quite dissimilar form all three genotypes and do not fall in the cluster. In second cluster there are only 4 genotypes 63, 64, 71 and 5 which means there are much similar to each other and different from genotypes of other clusters. Genotypes 5 and 64 have more similarities than genotype 53 of same cluster. Similarly, 53, 66, 65, 39 and 26 fall in cluster 3. In this cluster genotype 5 have much genetic differences from all other genotypes of same cluster while all other have many similarities. Cluster 4 has 11 genotypes which are 70, 67, 3, 19, 35, 32, 33, 20, 12, 11 and 9 these are all genetically much cluster to each other and fall in the same cluster. Cluster 5 has 7 genotypes as 29, 4, 68, 57, 58, 56 and 48. In this cluster all are similar to each other while genotype 48 is quite different from the rest of genotypes of the same cluster has separated itself from other genotypes of same cluster.

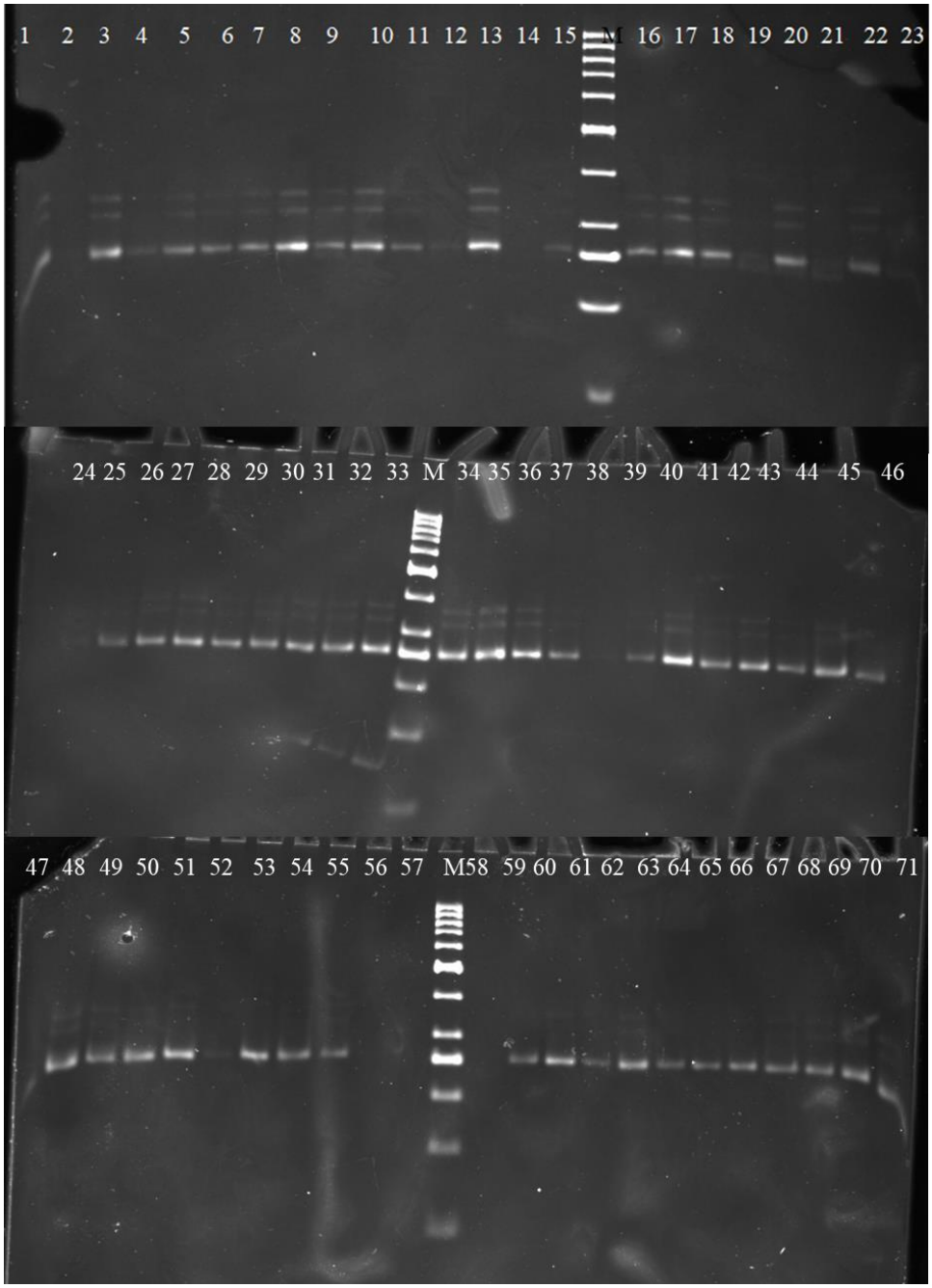

Figure 1. PCR Product analysis amplified by SSR primer pair NAU-1233 amplifying product sizes of 260bp, 340bp and 360bp M=50bp. The lane sequence (1-23) from left to right on gel represents the cotton genotypes listed in table 4.

Table 3. Band size amplified by each primer, their allelic frequency in population, population size in which allele was found and polymorphic information content.

Primer Name	Band Size	Population Size	Allele Freq
BNL-1551	160	50	0.71
	170	52	0.74
	270	46	0.66
	280	47	0.67
JESPER-247	90	54	0.77
	150	50	0.71
	160	30	0.43

NAU-2083	180	56	0.80
	350	55	0.79
NAU-1103	200	49	0.70
	260	48	0.69
NAU-1141	160	36	0.51
	170	62	0.89
	310	61	0.87
	330	62	0.89
NAU-1230	230	55	0.79
	240	32	0.46
	350	36	0.51
NAU-1233	260	55	0.79
	340	36	0.51
	360	36	0.51
NAU-1366	220	63	0.90
	230	62	0.89
	340	60	0.86
NAU-1369	250	54	0.77
	410	48	0.69
	430	38	0.54
NAU-2139	170	60	0.86
	180	60	0.86
NAU-2152	170	44	0.63
	205	52	0.74
	220	51	0.73
	230	51	0.73
NAU-2162	100	66	0.94
	290	61	0.87
	310	56	0.80
NAU-2169	240	68	0.97
	420	63	0.90
	450	64	0.91
NAU-2265	230	61	0.87
	240	55	0.79
	360	60	0.86
	400	60	0.86
NAU-2317	240	38	0.54
	250	47	0.67
NAU-2336	190	39	0.56
	200	52	0.74
NAU-2355	100	43	0.61
	120	40	0.57

	250	56	0.80
	265	46	0.66
NAU-2503	240	34	0.49
	250	37	0.53
	320	46	0.66
	340	37	0.53
NAU-2691	180	53	0.76
	200	47	0.67
	350	43	0.61
	355	44	0.63
NAU-2697	180	61	0.87
	190	60	0.86
	245	38	0.54
	160	58	0.83
NAU-3529	250	57	0.81
	270	54	0.77
NAU-3608	180	57	0.81
	200	58	0.83
	310	48	0.69
NAU-3654	170	62	0.89
	180	63	0.90
	300	47	0.67
	310	50	0.71
NAU-5005	234	62	0.89
	250	57	0.81
	320	57	0.81
	Total= 75 alleles		

Table 4. Maximum, minimum and average distance of each genotype with all other genotypes.

Sr No.	Genotypes	Minimum	Maximum	Average
1	NS 161	4.796	6.856	5.826
2	CIM 602	3.742	6.403	5.44925
3	SILKEE	3.873	6.403	5.10525
4	VH -GULZAR	3.742	6	5.0045
5	FH-152	4.359	6.325	5.1065
6	SITARA-14	3.606	6.633	5.23075
7	FH-444	4.123	6.083	5.11125
8	FH-490	3.606	6.245	5.01425
9	FH-442	4	6.481	5.083

10	SITARA-15	3.606	6.403	5.1225
11	FH- KAHKSHAN	4.472	6	5.12025
12	BS -15	4.69	6.245	5.35175
13	ZAKARIYA-1	4.796	6.325	5.514
14	IUB-65	4.69	6.782	5.64825
15	CEMB-55	3.606	6.403	5.37025
16	BS-80	3.464	6	4.86825
17	RH-622	3.606	6.325	4.84875
18	NIAB-874	4.123	6.481	5.13375
19	NIBGE-9	4.123	6.325	5.263
20	FH-119	3.742	6.164	5.0885
21	CIM-622	4.243	6.403	5.138
22	MNH-1016	4.123	6	5.19225
23	NW-142	4.472	6.164	5.18975
24	WEAL-AG-6012	4.69	6.557	5.47075
25	NIBGE-7	4.796	6.245	5.572
26	FH-326	4.583	6.164	5.447
27	FH-450	3.606	6.164	5.12925
28	FH-114	3.317	5.831	4.7295
29	PASCO-1000	4.243	6	4.84775
30	FH-315	2.449	5.745	4.60925
31	TARZAN-5	3.162	5.657	4.25325
32	FH-142	4.123	6	4.7355
33	FH-2073	4	6.164	5.07175
34	FH-242	3.742	5.745	4.91275
35	IR-NIBGE-8	4.583	5.916	4.9965
36	VH-363	4	5.916	5.10375
37	RH-688	3.742	6.164	4.9555
38	MNH-988	4	6	4.9765
39	BH-201	4.472	6.164	5.159
40	FH-458	3.606	6	5.0605
41	CYTO-179	4	5.916	4.8805
42	FH-471	4	5.916	4.958
43	FH-91	4	6	4.979
44	NS-181	4.899	6.164	5.26575
45	FH-466	4.359	6	5.3555
46	VH-327	4	6.164	5.13075
47	FH-313	4.243	6.481	5.222

48	FH-472	4.472	5.916	5.278
49	FH-473	4	5.831	5.05475
50	FH-NOOR	3.162	5.657	4.6625
51	NIAB-545	3.606	5.657	4.5205
52	Rajat	4.359	5.477	4.77475
53	WEAL-AG-SHAHKAR	4.243	5.745	4.956
54	FH-Lalazar	3.464	5.657	4.77725
55	MNH-886	4.243	5.831	4.79875
56	NIAB-878	3.162	6.481	4.92925
57	FH-479	3.873	6.083	4.89975
58	FH-477	4.243	6.164	5.09075
59	FH-476	3.742	5.657	4.9515
60	FH-478	4	5.568	4.74175
61	FH-312	4.123	5.385	4.769
62	Weal -AG-1606	4.359	5.385	4.813
63	NIAB-BT-2	4.472	5.831	5.01175
64	Turkey-1	5	5.745	5.262
65	IUB-63	3.873	5	4.9045
66	FH-342	4.69	4.796	4.58975
67	FH-474	4.123	4.472	4.52025
68	NIAB-1011	4.583	4.583	4.44025
69	Turkey-2	4.899	4.899	4.741
70	Turkey-3	4.10	6.235	5.167
71	FH-118	4.342	5.893	5.1175

Cluster 6 is a smaller one with only 4 genotypes which are 2, 14, 13 and 1. Genotype 1 is dissimilar to the other genotypes of the same cluster. Starting from 5 Eculidean then genotype 1 is dissimilar from rest of the genotypes of all clusters. Cluster 7 has 7 genotypes which are 49, 55, 54, 43, 52, 62 and 59. All genotypes fall in same cluster closer to each-others which means all are similar. However, genotype 21 is different from the other genotypes and does not fall in any cluster and it is separated from other. Cluster 8 has 27 genotypes. It is a bigger cluster and divided into 3 sub-clusters. Genotypes 45, 51, 50, 60, 27, 36 and 7 are similar to each other and fall in the main cluster while two genotypes 18 and 22 are different from the other genotypes in main cluster. Genotypes 36 and 7 are also much dissimilar from the other genotypes of main cluster. Sub-cluster 1 has 7 genotypes which are 41, 31, 34, 30, 28, 42.

and 17 these falls in the same sub-cluster means these are much similar to each other while different from rest of genotypes of same cluster. Sub-cluster 2 has 7 genotypes 61, 10, 37, 16, 6, 8 and 3 while sub-cluster 3 has only four genotypes which are 69, 40, 38 and 15 which have different genetic make-up from the other genotypes. Dendrogram constructed with the help of polymorphic loci created through SSR markers. The maximum genetic similarity coefficient value was found 6.85 between genotypes FH-479 and NS-161 which shows that they may have common ancestors or the same morphological traits while the lowest similarity coefficient value was found 2.449 between genotypes FH-315 and Tarzan-5 which indicate that they have different parental origin or distinct morphological characteristics. On the basis of similarity of genotypes dendrogram clustered all genotypes into 37 clusters. Most diverse genotype was 30 which can be used as parent in future breeding programs.

Association mapping of CLCuD resistance at CRS, Faisalabad and CRS, Vehari

For association mapping 109 markers were utilized to study different traits. For CLCuD studies it was revealed that five markers NAU3254-1600, dPL0526-260, NAU5163-200, CIR094-700, cgr6356-150 are associated at CRS,

Faisalabad. While At CRS, Vehari the marker NAU3385-500 is different while 4 others are common at both locations. Regarding the positioning of markers at CRS, Faisalabad it was found that NAU3254-1600 is located at chromosome 11 , dPL0526-260 at 6,8 and 11, NAU5163-200 at 4, 7,8 11,19,24 , CIR094-700 at 3 and 7 while cgr6356-150 at chromosome number 10 and 20. While at Vehari NAU3254-1600 is located at chromosome 8, 11 and 12, CIR-094-700 at 2 and 7, dPL0526-260 at 8 and 11, NAU3385-500 at chromosome No-2 and cgr6356-150 at chromosome at 7, 18 and 20. The results of association mapping for CLCuD are given in Table 5 and 6.

Table 5. Association mapping of CLCuD% at CRS, Faisalabad, single marker analysis using WinQTL deviance F-test with linkage group (LG), marker, intercept (b0), slope (b1), hypothesis test (-2ln(L0/L1), F value and probability.

Chrome	Marker	b0	b1	2ln(L0/L1)	F(1, n2)	pr (F)	R2
11	NAU3254-1600	22.937	2.508	15.609	16.966	0.000104047***	0.1921
08	dPL0526-260	23.017	2.012	9.562	9.947	0.002386006**	0.1187
26	NAU5163-200	22.257	-1.899	8.294	8.55	0.004670435**	0.0909
07	CIR094-700	23.207	1.935	7.47	7.656	0.0072577**	0.1012
20	cgr6356-150	22.775	1.573	5.535	5.594	0.020833833*	0.075
24	NAU5163-200	21.071	-2.164	4.989	5.023	0.028233456*	0.0679
04	NAU5163-200	22.768	1.46	4.742	4.766	0.032427894*	0.0742
11	NAU5163-200	22.746	1.405	4.596	4.614	0.035225522*	0.0627
19	NAU5163-200	21.255	-2.033	4.584	4.602	0.035462211*	0.0632
07	NAU5163-200	23.184	1.484	4.466	4.479	0.037918406*	0.0653
10	cgr6356-150	22.705	1.356	4.406	4.418	0.039223212*	0.0602
11	dPL0526-260	22.764	1.373	4.308	4.316	0.041468961*	0.0589
06	dPL0526-260	22.782	1.381	4.282	4.289	0.042101338*	0.0515
03	CIR094-700	22.822	1.333	4.068	4.069	0.047560889*	0.0444

Table 6. Association mapping of CLCuD% at CRS, Vehari, , single marker analysis using WinQTL deviance F-test with linkage group (LG), marker, intercept (b0), slope (b1), hypothesis test (-2ln(L0/L1), F value and probability.

Chrome	Marker	b0	b1	2ln(L0/L1)	F(1, n2)	pr (F)	R2
11	NAU3254-1600	44.353	-4.248	8.006	8.236	0.005447968 **	0.1004
07	CIR094-700	43.657	-4.169	6.484	6.598	0.012377566 *	0.0899
11	dPL0526-260	44.574	-3.63	5.725	5.794	0.018760248 *	0.0775
04	NAU3254-1600	43.614	-3.756	5.714	5.783	0.018873829 *	0.0773
02	NAU3385-500	43.934	3.691	5.672	5.738	0.019315002 *	0.0616
12	NAU3254-1600	44.161	-3.652	5.638	5.703	0.019679497 *	0.0822
02	CIR094-700	44.128	3.552	5.375	5.426	0.022771350 *	0.0729
04	NAU5163-200	44.579	-3.536	5.255	5.301	0.024341531 *	0.0849
08	NAU3254-1600	43.063	-3.923	5.247	5.292	0.024452625 *	0.068
07	cgr6356-150	43.262	-3.711	5.022	5.057	0.027715531 *	0.0599

20	cgr6356-150	44.589	-3.366	4.747	4.771	0.032351389 *	0.0647
08	dPL0526-260	44.245	-3.244	4.517	4.532	0.036832151 *	0.0592
18	cgr6356-150	44.793	3.245	4.331	4.34	0.040936587 *	0.05

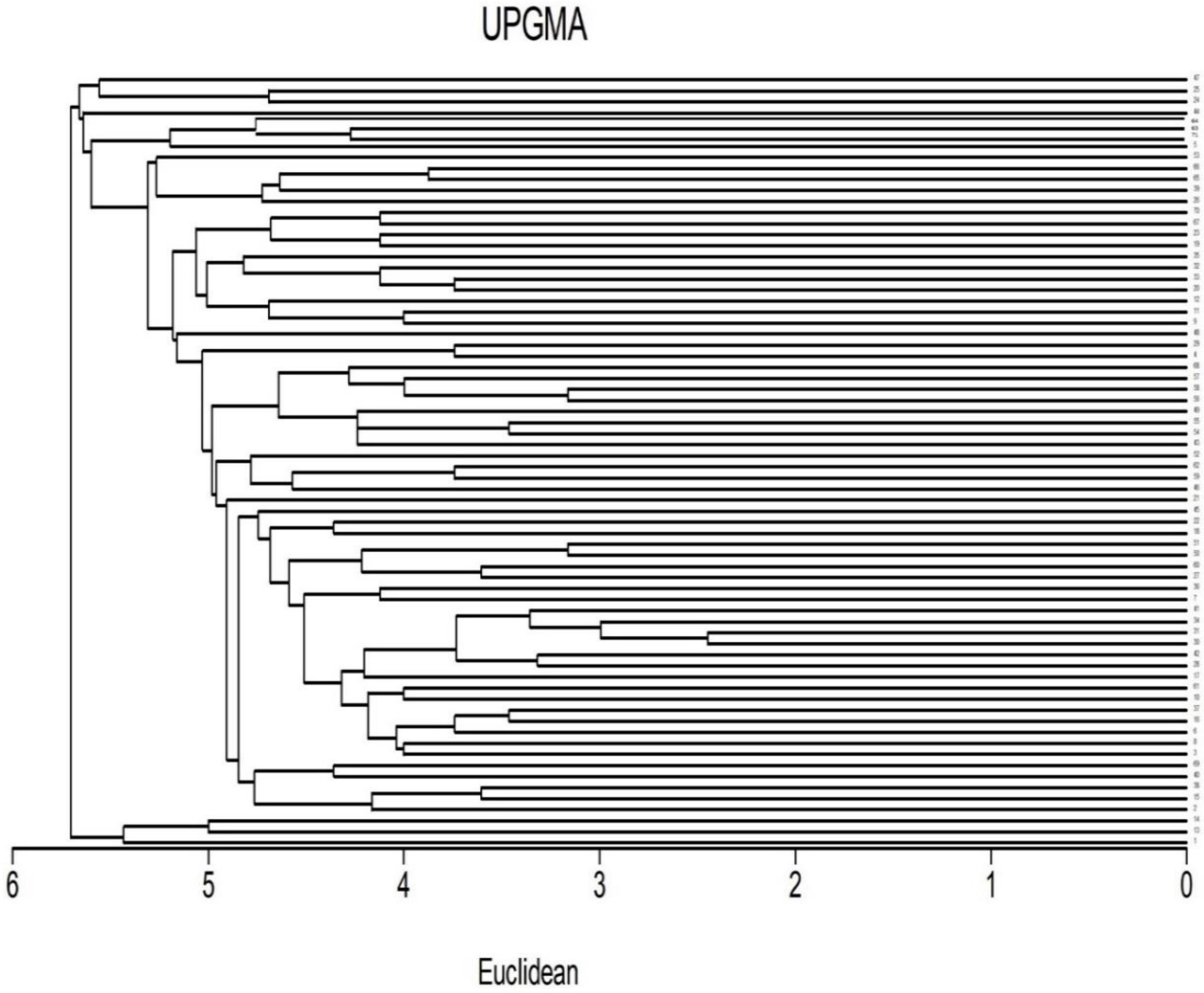

Figure 2. Dendrogram constructed by using arithmetic averages (UPGMA) method using the unweighted pair group method showing genetic diversity.

DISCUSSION

Genetic diversity of genotypes using SSR markers

The availability of genetic variability among and within closely associated genotypes is vital for convincing exploitation of germplasm (Ali *et al*., 2009). Utilization of genetic variation in hybridization programmes assist in the recognition of genotypes that can be most suited to variable environmental conditions (Zada *et al*., 2013; Farooq *et al*., 2014). To find out information on magnitude of genetic variation the researchers have to depend upon different procedures exploited for its measurement (Farooq *et al*., 2014). SSR markers have largely been used to study genetic diversity and character association mapping for biotic stresses (Ali *et al*., 2019; Jabran *et al*., 2023). SSR markers had a more allelic nature, were relatively abundant, were highly reproducible, and had genomic coverage with co-dominant inheritance (Liu *et al*., 2000; Semagan *et al*., 2006). Information on genetic variation in terms of susceptibility or resistance to CLCuD and the agronomic value of economically imperative traits are essential (Raza *et al*., 2016).

In the present study, a total of 71 cotton genotypes were screened for CLCuD resistance/tolerance, and out of these, only 9 cotton genotypes were found tolerant against the viral disease. The population of 71 genotypes was screened with 109 SSR markers linked with disease tolerance. Among 109 SSRs screened on 9 parents, only 14 were found polymorphic. These polymorphic markers were surveyed to find an association with CLCuD incidence. Four marker loci named NAU3254-1600, dPL0526-200, NAU5163-200, and CIR094-700 were found to be significantly associated with the CLCuD tolerance.

The identification of stable QTLs across variable environments and populations plays a vital role in marker-assisted selection (MAS) (Jamshed *et al*., 2016). In the present study, 23 stable QTLs were simultaneously identified under two variable environments i.e. two hotspots, of which five stable QTLs linked to Chromosome 11, two QTL on Chromosome 2, 7, 8, 20, and 26. In contrast, one QTLs was found located on chromosome numbers 3, 4, 5, 6, 16, 19, 21 and 25. The results of the present study suggested that it is reasonable and effective to map QTLs using the parental population together with its corresponding F_1 population across multiple environments and multiple years (Shang *et al*., 2016).

We contrasted the associated markers found in the present study with SSR markers previously found through linkage QTL and association mapping analyses (Zhang *et al*., 2005; Zhao *et al*., 2014; Wang *et al*., 2013; Wang *et al*., 2012). Association mapping studies revealed that five markers NAU3254-1600, dPL0526-260, NAU5163-200, CIR094-700, cgr6356-150 are associated with CLCuD resistance at CRS, Faisalabad. While At CRS, Vehari, the marker NAU3385-500 is different, while 4 others are common at both locations. Abbas et al. (2015) reported two SSR markers, PR-91 and CM-43, associated with CLCuD resistance in cotton.

Similarly, A total of 2400 SSRs were initially selected to explore the genomes of both species. Out of these, Rahman et al. (2014) reported 113 polymorphic SSRs and out of these found two QTLs, i.e., QCLCuD25 and QCLCuD26 associated with CLCuD resistance. Recently, Schoonmaker et al. (2023) have developed a QTL mapping App named “iCottonQTL”, which simplifies the process of QTL mapping from cotton genotyping arrays and the subsequent submission of data to the CottonGen database. They have identified multiple QTL in different locations on the genome, suggesting that there are multiple resistance genes present in geographically diverse cotton germplasm (Schoonmaker et al., 2023). Identification of these sources provides the basis for more durable resistance via the stacking of this resistance QTL. Up till now, very few microsatellite markers have been proficiently employed in marker-assisted selection programs of cotton crops as the bulk of existing marker information was derivative of populations resulting from biparental combinations possessing limited genetic backgrounds, encompassing very few meiotic events after experimental hybridization (Abdurakhmonov *et al.,* 2009). The latest association mapping of different cotton germplasms established the possibility of utilizing association study to explore intricate traits in Upland cotton and found valuable markers for MAS breeding programs (Zeng *et al*., 2009; Zhao *et al*., 2014; Joshi *et al*., 2023). Analogous to linkage mapping, association mapping by the means of diverse materials associating different genes and variable molecular markers can offer more useful information for MAS breeding programs along with the understanding of the genetic basis of valuable traits in cotton crops.

CONCLUSION

The findings of the current study delivered new and valuable molecular markers for MAS in cotton breeding programs and indications for the fine-tuned mapping of CLCuD resistance. These studies will also enrich our existing understanding of the genetic basis of genetic diversity and disease resistance in upland cotton at the genomic level.

AUTHOR CONTRIBUTIONS

AB mainly conducted the research developed the primary manuscript, JF and AF helped in data recording and analysis, MUS and GS helped in manuscript writing and reviewing.

CONFLICT OF INTREREST

The authors have no conflict of interest.

REFERENCES

Abbas, A., Iqbal, M.A., Rahman, M., & Paterson, A.H. (2015). Estimating genetic diversity among selected cotton genotypes and the identification of DNA markers associated with resistance to cotton leaf curl disease. Turkish Journal of Botany, 39, 1033–1041.

Akhtar, K.P., Haidar, S., Khan, M.K.R., Ahmad, M., Sarwar, N., Murtaza, M.A., & Aslam M. (2010). Evaluation of Gossypium species for resistance to cotton leaf curl burewala virus. Annals of Applied Biology, 157, 135–147.

Ali, M.A., & Khan I.A. (2007). Assessment of genetic variation and inheritance mode of some metric traits in cotton (*Gossypium hirsutum* L.). Journal of Agriculture & Social Sciences, 3(4), 112–116.

Ali, M.A., Farooq, J., Batool, A., Zahoor, A., Azeem, F., Mahmood, A., & Jabran, K. (2019). Cotton Production in Pakistan. In "Cotton Production Wordwide" (Eds. Jabran, K., & Chauhan, B.). John Wiley and Sons, Pages-249.

Ali, M.A., Nawab, N.N., Abbas, A., Zulkiffal, M., & Sajjad, M. (2009). Evaluation of selection criteria in *Cicer arietinum* L. using correlation coefficients and path analysis. Australian Journal of Crop Science, 3, 65-70.

Batool, A., Farooq, J., Jabran, M., Abbas, A., Ali, M.A. (2024). Exploring host resistance and character association in diverse cotton germplasm to manage Cotton Leaf Curl Disease (CLCuD) at two hotspots in Punjab, Pakistan. Journal of Crop Health, 76, 769–782.

Batool, A., Javed, N., Abbas, R.M., Iqbal, M.Z., & Ali, M.A. (2021). Assessment of cotton germplasm against CLCuD at different locations and determination of inheritance patterns for resistance. Pakistan Journal of Agricultural Sciences, 58, 1843-1855.

Dongre, A. & Parkhi, V (2005). Identification of cotton hybrid through the combination of PCR based RAPD, ISSR and microsatellite markers. Journal of Plant Biochemistry & Biotechnology, 14(1), 53-55.

Doyle, J.J. & Doyle, J.L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissue. Phytochemical Bulletin, 19, 11-15.

Farooq, A., Farooq, J., Mahmood, A., Shakeel, A., Rehman, A., Batool, A., Riaz, M., Shahid, M.T.H. & Mehboob, S. (2011). Overviews of cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity. Australian Journal of Crop Science, 5, 1823-1831.

Farooq, J., Anwar, M., Riaz, M., Farooq, A., Mahmood, A., Shahid, M.T.H., Shahid, M.R. & Ilahi, F. (2014). Correlation and path coefficient analysis of earliness, fiber quality and yield contributing traits in cotton (*Gossypium hirsutum* L.). Journal of Animal & Plant Sciences, 24(3), 781-790.

GOP, Economic Survey of Pakistan. (2023-24). Ministry of Finance.

Joshi, B., Singh, S., Tiwari, G.J., Kumar, H., Boopathi, N.M., Jaiswal, S., Adhikari, D., Kumar, Dinesh, Sawant, S.V., Iquebal, M.A., & Jena, S.N. (2023). Genome-wide association study of fiber yield-related traits uncovers the novel genomic regions and candidate genes in Indian upland cotton (*Gossypium hirsutum* L.). Frontiers in Plant Sciences, 14, 1252746

Hussain A., Azhar, F.M., Ali, M.A., Ahmad S., & Mahmood, K. (2010). Genetic studies of fiber quality characters in upland cotton. Journal of Animal & Plant Sciences, 20(4), 234-238.

Iqbal, M., Naeem, M., Aziz, U., Afzal, J., & Khan, M.A. (2014). An overview of cotton leaf curl virus disease, persistent challenge for cotton production. Bulgarian Journal of Agricultural Sciences, 20, 405-415.

Jabran, M., Ali, M.A., Zahoor, A., Muhae-Ud-Din, G., Liu, T., Chen W., & Gao, L, (2023). Intelligent reprogramming of wheat for enhancement of fungal and nematode disease resistance using advanced molecular techniques. Frontiers in Plant Sciences, 14,1132699.

Jamshed, M., Jia, F., Gong, J., Palanga, K.K., Shi Y., & Li, J. (2016). Identification of stable quantitative trait loci (QTLs) for fiber quality traits across multiple environments in *Gossypium hirsutum* recombinant inbred line population. BMC Genomics, 17,197.

Li, Z., Wang, X., Yan, Z., Guiyin, Z., Wu, L., Jina, C., & Ma, Z. (2008). Assessment of genetic diversity in glandless cotton germplasm resources by using agronomic traits and molecular markers. Frontiers of Agriculture in China, 2, 245-252.

Liu, S., Cantrell, R.G., McCarty, J.C. and Stewart, J.M. (2000). Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. Crop Science, 40, 1459-1469.

Mathews, R.E.F. (1987). The changing scene in plant virology. Annual Reviews of Phytopathology, 25, 10-23.

Rahman, M.A., Ahmad, A.Q., Khan, A., Abbas, Rahmat, Z., & Sarfraz, Z. (2014). Use of genetic and genomic approaches for controlling cotton leaf curl disease complex in Pakistan," *In Proceedings of the International Cotton Genome Initiative Conference* 26–28, Wuhan.

Raza, H., Khan, N.U., Khan, S.A., Gul, S., Latif, A., Hussain, I., Khan, J., Raza S., & Baloch, M. (2016). Genetic variability and correlation studies in F_4 populations of upland cotton. Journal of Animal & Plant Sciences, 26(4), 1048-1055.

Rehman, M.U., Khan, A.Q., Rahmat, Z., Iqbal, M.A., & Zafar, Y. (2017). Genetics and genomics of cotton leaf curl disease, its viral causal agents and whitefly vector: a way forward to sustain cotton fiber security. Frontiers in Plant Sciences, https://doi.org/10.3389/fpls.2017.01157.

Sattar, M.N., Kvarnheden, A. Saeed M., & Briddon, R.W (2013). Cotton leaf curl disease – an emerging threat to cotton production worldwide. Journal of General Virology, 94, 695–710.

Schoonmaker, A.N., Hulse-Kemp, A.M., Youngblood, R.C., Rahmat, Z., Iqbal M.A., Rahman, M.U., Kochan, K.J., Scheffler, B.E., Scheffler, J.A. (2023). Detecting Cotton Leaf Curl Virus Resistance Quantitative Trait Loci

in *Gossypium hirsutum* and iCottonQTL a New R/Shiny App to Streamline Genetic Mapping. Plants. 12(5):1153. https://doi.org/10.3390/plants12051153

Semagan, K., Bjornstad A., & Ndjiondjop, M.N. (2006). An overview of molecular marker methods for plants. African Journal of Biotechnology, 5 (25), 2540-2568.

Shang, L, Wang, Y. Wang, X. Liu, F., Abduweli, A., Cai, S., Li, Y., Ma, L., Wang, K., & Hua, J. (2016). Genetic analysis and QTL Detection on fiber traits using two recombinant inbred lines and their backcross populations in upland cotton. G3 Genes Genomics Genetics, 6(9), 2717–2724.

Tyagi, P., Michael, A.G., Bowman, D.T., Campbell, B.T., Udall, J.A. & Kuraparthy, V. (2014). Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.) Theoretical and Applied Genetics, 127, 283–295.

Van-Esbroeck, G.A. and Bowman, D.T. (1998). Cotton germplasm diversity and its importance to cultivar development. Journal of Cotton Science, 2, 121-129.

Wang, S., Basten, C.J., Zeng, Z.B. (2010). Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh. http://statgen.ncsu.edu/qtlcart/WQTL Cart.htm.

Wang, F.R., Xu, Z.Z. Sun, R., Gong, Y.C., &. Liu, G.D. (2013). Genetic dissection of the introgressive genomic components from *Gossypium barbadense* L. that contribute to improved fiber quality in *Gossypium hirsutum* L. Molecular Breeding, 32(3), 547–562.

Wang, P., Zhu, Y.J., Song, X.L., Cao, Z.B., and Ding, Y.Z. (2012). Inheritance of long staple fiber quality traits of *Gossypium barbadense* in *G.hirsutum* background using CSILs. Theoretical & Applied Genetics, 124(8),1415–1428.

Watkins, G.M. (1981). Compendium of cotton diseases. The American phytopathological society, 3340. Pilot knob Road, Minnesota, USA. pp. 87

Yu, J., Zhang, K., Li, S., Yu, S., Zhai, H., Wu, M., Li, X., Fan, S., Song, M., Yang, D. & Li, Y. (2013). Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. Theoretical & Applied Genetics, 126, pp.275-287.

Zada, M., Zakir, N., Rabbani, M.A., & Shinwari, Z.A. (2013). Assessment of genetic variation in Ethiopian mustard (*Brassica carinata* A. Braun) germplasm using multivariate techniques. Pakistan Journal of Botany 45(S1),583-593.

Zeng, L.H., Jr. Meredith, W.R., Gutiérrez O.A., & Boykin, D.L. (2009). Identification of associations between SSR markers and fiber traits in an exotic germplasm derived from multiple crosses among Gossypium tetraploid species. Theoretical & Applied Genetics, 119(1), 93–103.

Zhang, Z.S., Xiao, Y.H., Luo, M., Li, X.B., & Luo, X.Y. (2005). Construction of a genetic linkage map and QTL analysis of fiber-related traits in Upland cotton (*Gossypium hirsutum* L.). Euphytica, 144,91–99.

Zhao, Y., Wang, H. Chen, W. & Li, Y. (2014). Genetic structure, linkage disequilibrium and association mapping of Verticillium wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. PloS One, 9, e86308.