Asian Journal of Biotechnology and Genetic Engineering

Volume 6, Issue 2, Page 164-170, 2023; Article no.AJBGE.106008

Identification of Molecular Diversity in *Gossypium hirsutum* L. Against Whitefly Resistance Using SSR Markers

Syed Bilal Hussain [a], Muhammad Tayyab [a*], Muhammad Adnan Shah Bukhari [a] and Muhammad Zubair [b]

[a] Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Multan, Pakistan.
[b] Department of Forestry and Range Management, FAS and T, Bahauddin Zakariya University, Multan, Pakistan.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/106008

Original Research Article

Received: 08/07/2023
Accepted: 14/09/2023
Published: 21/09/2023

ABSTRACT

Whitefly is a major factor in reducing cotton yield worldwide more specifically in Asian Countries So to deal with this whitefly there is a need to separate genetically diverse varieties that can combat it due to its unique genetic source. In this study, genetic diversity analysis of cotton cultivars was done against whitefly resistance using simple sequence repeat (SSR) markers. For this, 75 cultivars of cotton were used to analyze genetic diversity by using 25 SSR markers out of which 9 markers were polymorphic. These markers amplified a total of 46 alleles with an average of 5.11 alleles per marker. These varieties showed a high value of gene diversity that ranged from 0.624 to 0.877 with a mean value of 0.735. Polymorphism information content (PIC) for selected genotypes ranged from 0.562 to 0.867 with a mean value of 0.704. Based on these 9 polymorphic markers 75 selected genotypes were divided into 3 major groups in a phylogenetic tree. The elected cultivars of cotton were highly polymorphic and could be used for further breeding to improve cotton accessions.

**Corresponding author: E-mail: muhammadtayyab4852@gmail.com;*

Keywords: Whitefly; CLCuD; gene diversity; CCRI; SSR; PIC.

1. INTRODUCTION

Humans have been using cotton for a long time. Historically, different samples of cotton (*Gossypium hirsutum* L) plants were found in different eras [1]. A complete plant of cotton is useable, but the primary reason for growing cotton is fiber. The importance of cotton is because its fiber and seed both have massive use in the industry [2]. Cotton is important for certain reasons. It is among the most crucial cash crops as far as the agriculture field is concerned. Industrially, its importance is because of its fiber which is most widely used in the medical, surgical, paper industry, food industries, accessories like clothing handbags [3], and home textiles. Lint of cotton is used for automobile products, fabrics, cords [4], etc. Cotton seed is important because of its usage in cottonseed oil for the food industry, animal feed, new plantings, and also for the rubber industry. Cotton is a source of employment in different countries mostly developing countries and many people in these countries are dependent on cotton. These people are dependent on cotton because their source of income is related to the production or processing of cotton. Cotton is not only grown for the sake of fiber it's also a food crop for animals as well as humans. Cotton seed oil constitutes 10% of the world's total oil for foodstuff and other uses of cottonseed oil are as biodiesel, cleanser, a creature of mash and compost. Lint also consists of cellulose which is industrially very important in the paper industry and also in the chemical industry. Linter which is the by-product of mills processing cotton seed for oil production is 99% cellulose so we can use it as an alternative to that cellulose. In Pakistan, cotton is also called "white gold" because of its ability to generate revenue for its cultivars. As cotton contribute 60% of the total export of Pakistan so it is a major source of rising gross domestic product (GDP). Cotton seed producers use the germplasm of cotton for breeding and improving cotton day by day with the emergence of latent techniques and technologies so it is a main business for cotton seed producers.

Many biotic factors destroy cotton pests are the leading problem of cotton. There are so many types of pests that attack cotton throughout its season. These pests are divided into two classes because of their mechanism of attack on plants these are chewing bugs and the second class is sap-sucking bugs. Bollworms, pink bollworms, spotted bollworms and armyworms are chewing bugs while whitefly, aphids, jassids, and mealy bug are sucking bugs. These are biotic factors that reduce cotton production. Currently, in Pakistan, there is massive usage of pesticides, fungicides, and insecticides to deal with these biotic factors but results are still disappointing rather it causes serious effects on labor using these pesticides. Previously two methods were common for dealing with these diseases, immunization, and removal of diseased plants.

There is a need for gene diversity in plants that will provide a broad genetic spectrum to deal with disease. Genetic diversity analysis is done in most studies using simple sequence repeat (SSR) markers. After the identification of a specific SSR marker associated with desired genes location of this gene throughout the whole genome of that species can be traced with the help of PCR. For this case, primers are designed consisting of a DNA sequence similar to the SSR marker identified and associated with the gene of interest and then PCR of that specific organism's DNA is done using these primers. In this study, we aimed to evaluate genetic diversity in cotton by employing SSRs.

2. MATERIALS AND METHODS

The current study was conducted at the Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University Multan from June 2022 to June 2023. In this study, genetic diversity analysis of whitefly-resistant cotton cultivars was performed using simple sequence repeat (SSR) markers. Seventy-five cotton cultivars of cotton were selected from the Central Cotton Research Institute of Multan (CCRI) for genetic diversity analysis using SSR markers (Table 2). DNA extraction of the obtained samples was done by using Cetyltrimethylammoniumbromide (CTAB) method [5].

After extraction of DNA, polymerase chain reaction (PCR) was performed using standard temperature and conditions. The annealing temperatures of the primers are mentioned in (Table 2). Agarose gel electrophoresis of PCR products was done to visualize the fragments produced in each variety. The Gel Doc XR+ (www.bio-rad.com) was used to visualize the gels after electrophoresis.

Table 1. List of 75 selected cultivars

Ser. No.	Cotton Variety	Ser. No.	Cotton Variety	Ser. No.	Cotton Variety
1	MNH- 552	26	1020	51	VH- 307
2	MNH- 554	27	VH- 327	52	1035
3	MNH- 147	28	BH- 178	53	FH- 490
4	NIAB- Noori	29	VH- 259	54	SS- 32
5	DNH- 105	30	SLH- 08	55	Sahara klean- 05
6	CA- 12	31	VH- 307	56	j- 05
7	1016	32	1035	57	BS- 18
8	MNH- 129	33	FH- 490	58	N- 878
9	RH- 662	34	SS- 32	59	Sahara Klean- 150
10	RH- 668	35	Sahara klean- 05	60	Sahara- 120
11	BH- 184	36	j- 05	61	SLH- 04
12	RH- 647	37	BS- 18	62	SADORI
13	AA- 703	38	N- 878	63	VH- 281
14	Cris- 578	39	Sahara Klean- 150	64	FH- 118
15	V- 14	40	Sahara- 120	65	FH- 114
16	1020	41	Sahara- 150	66	FH- 942
17	VH- 327	42	1045	67	FH- lallazar
18	BH- 178	43	Sahara Burak	68	Bt CIM- 599
19	VH- 259	44	IUB- 13	69	FH- 326
20	SLH- 08	45	BS- 15	70	FH- 941
21	GH- 102	46	MNH- 1086	71	CRSM- 38
22	NS- 141	47	BT- A1	72	Chandni- 95
23	CIM- 707	48	AGC- 555	73	CIM- 599
24	1026	49	NLA- UFAQ	74	CIM- 591
25	CIM- 534	50	NS- 161	75	CIM- 573

Table 2. List of SSR markers

Name	Repeat Motif	Annealing Temperature	Direction	Sequence
JESPR0013	CTT	54.7	F	GCT CTC AAA TTG GCC TGT GT
			R	GGT GGA GGC ATT CCT GCT AAC
BNL2835	GA	57.8	F	AAG ATA ATC GCC GGC TAG CT
			R	CCG CCA GTT TGC CAT AAT AT
JESPR0220	GA	54.7	F	CGA GGA AGA AAT GAG GTT GG
			R	CTA AGA ACC AAC ATG TGA GAC C
JESPR0274	CA	57.8	F	GCC CAC TCT TTC TTC AAC AC
			R	TGA TGT CAT GTG CCT TGC
BNL0409	GA	55.9	F	AGA AGT CGG ACG TGG AAA GA
			R	GCC GTT TTC CTC AGT GAT GT
BNL0387	AG	54.7	F	GAA GGG GAA TTT ATA GCG GG
			R	AGA GAC TCC CGC ACT TGA AA
BNL0341	GA	55.9	F	ACC TGG GGT ACT TGT CCA CA
			R	CCA TCC CAT TTG TGA TAC CC
BNL0243	AG	54.7	F	GGG TTT TCT GGG TAT TTA TAC AAC A
			R	TCA TCC ACT TCA GCA GCA TC
JESPR0042	CTT	57.8	F	CGT TGC CGT CTT CGA CTC CTT
			R	GTG GGT GGC TAA TAT GTA GTA GTC G

Data analysis was done by using two software, ImageLab (www.bio-rad.com/en-pk/ product/ image-lab-software)and Power Marker(http://statgen.ncsu.edu/powermarker/) [6]. ImageLab was used to detect the number of bands in each well of the gel. Data from these results was obtained in binary form 0 and 1 for the presence of the band at a specific location

was 1 and for the absence 0. This data in binary form was used in the PowerMarker to generate information on genetic diversity, PIC value, and the formation of a phylogenetic tree.

3. RESULTS AND DISCUSSION

For the analysis of genetic diversity within our 75 selected cultivars, we utilized 9 SSR markers. These markers amplified a total of 46 loci. Out of these 9 markers were polymorphic. The maximum number of loci amplified by a single marker was 8 by JESPR0013. In contrast, BNL0341 amplified 3 loci, the least number of loci amplified per marker.

As genetic diversity is concerned, our marker which showed the highest value of genetic diversity was BNL2835 with its value of 0.8779. BNL0409 and BNL0242 showed genetic diversity near to maximum with values 0.8661 and 0.8604 respectively. The lowest genetic diversity was shown by JESPR0274 which was 0.624. The mean genetic diversity of our research was 0.7358 which was calculated from the range between 0.879 and 0.624.

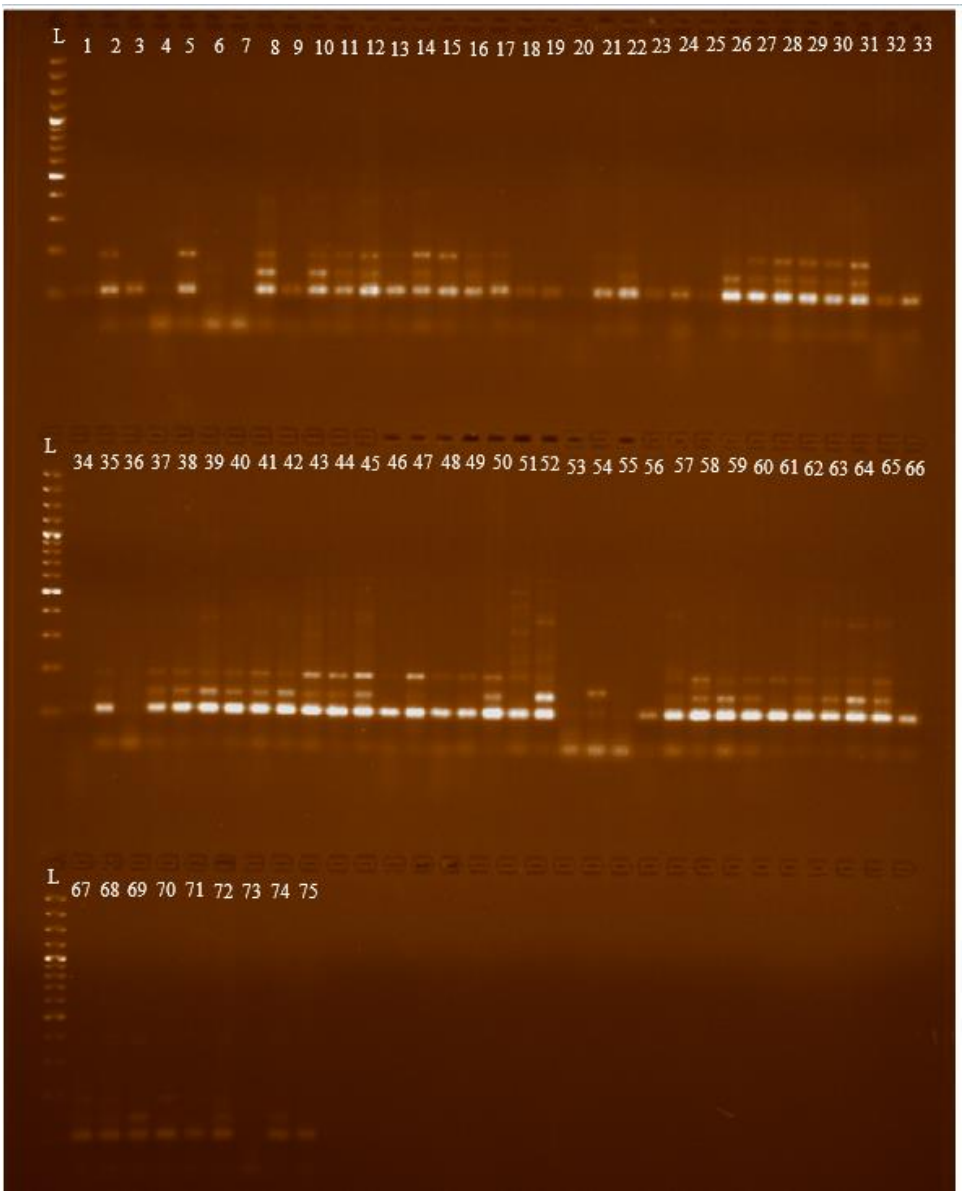

Fig. 1. Agarose gel visualization for BNL2835 on 75 samples

To study the value of polymorphism showed by our 9 markers we employed polymorphism information content. We obtained the highest value of PIC from marker BNL2835 with a value of 0.8672. BNL0409 and BNL0243 also showed maximum PIC value near BNL2835. The lowest PIC value was BNL0341 which was 0.562. The mean value obtained from a range of 0.8672 and 0.562 was 0.6340. All values we obtained from our markers related to PIC value and genetic diversity are mentioned in Table 3.

Based on the genetic distance table obtained, a phylogenetic tree was constructed (Fig. 2). From that phylogenetic tree, we analyzed that our selected varieties were grouped into 3 clusters. We named these clusters A, B, and C for our ease of data interpretation. Cluster A consisted of 4 varieties, Cluster B 61 varieties, and Cluster C consisted of 10 varieties in it. In cluster A, The variety with the highest distance was MNH-129 with its diversity value of 0.19 and the closest to the root of the tree was Cris-578 with its value of 0.11. In cluster B, the variety with the highest genetic distance was CIM-707 and other varieties that were close to this were SADORI, FH-114, FH-326, and NS-161 and most closest variety to the tree root was VH-307 with a value of 0.07. While in cluster C, BS-15 had the highest genetic distance value and the closest to tree root was line 1045.

With the implementation of the latest technologies and current software it becomes really easy to perform research to trace out desired characters within species [7,8]. We can utilize advanced technologies to find out the most diverse varieties within different genotypes for further breeding programs [9]. In our current study, we utilized 9 polymorphic SSR markers to find out genetic diversity within 75 cotton accessions. From that, we separated genetically different varieties based on data obtained from our results. Our results agreed with the earlier studies on the molecular diversity analysis of *G. hirsutum* [10,11].

The study of phylogenetic tree clusters that were formed by currently cultivated varieties helped to separate diverse ones from these as was done by Sagar [12]. The level of genetic diversity in our results was higher as the maximum genetic diversity showed by marker was 0.877 and polymorphism information content showed by selected accession was also higher with its maximum value of 0.867. These results indicated a higher level of polymorphism which is contrary to the results of Zhao [13].

In this research, a total of 46 alleles were amplified by 9 markers with a mean of 5.11 alleles per marker the average per marker was the same as obtained by Moiana et al. [14]. In their research. That mean value was much more than the results obtained by Fang et al. [15]. which were 2.6 alleles per marker. The reason for the low values of genetic diversity, PIC, and number of alleles amplified might be the selected genotypes for study. Genetic diversity among cultivated accessions will be higher as compared to wild types because currently cultivated types undergone a lot of modifications or changes as compared to wild ones.

Table 3. Statistics of genetic diversity, PIC, allelic frequency and allele number

Marker	Major Allele Frequency	Sample Size	Allele No.	Genetic Diversity	PIC	Allele number
BNL0409	0.2667	75	19	0.8661	0.8542	8
JESPR0220	0.3200	75	7	0.7783	0.7451	3
JESPR0274	0.5867	75	11	0.6240	0.6016	5
BNL2835	0.2267	75	20	0.8779	0.8672	8
BNL0341	0.4933	75	6	0.6276	0.5626	3
BNL0243	0.2667	75	17	0.8604	0.8478	7
JESPR0013	0.4667	75	5	0.6674	0.6132	3
JESPR0042	0.5467	75	7	0.6464	0.6120	4
BNL0387	0.5067	75	6	0.6745	0.6367	5
Mean	0.4089	75	11	0.7358	0.7045	5.11

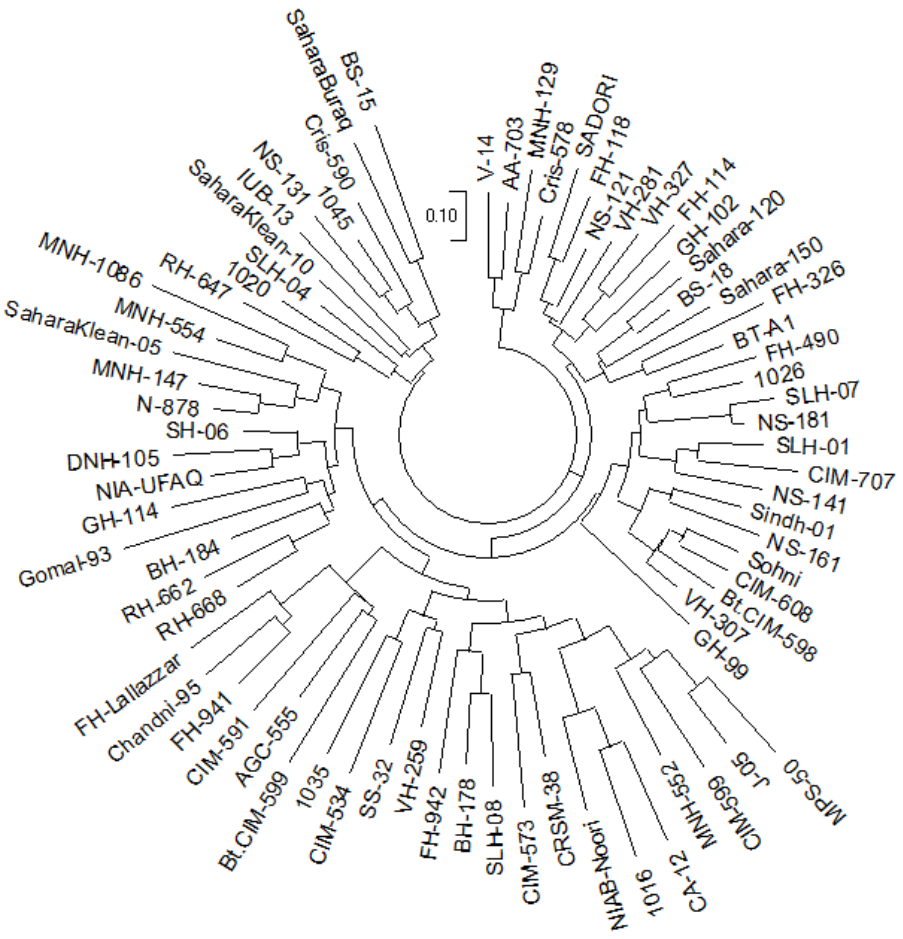

Fig. 2. Circular phylogenetic tree

Selected varieties showed medium to low resistance against whiteflies. It was necessary to find out the genetic diversity of these so that the reliability of these could be estimated. Results proposed that these are highly polymorphic according to our results as these showed high values of genetic diversity and PIC. According to these values, these varieties are best suited for further breeding programs. Our selected markers amplified a high number of alleles per marker which enhanced genetic diversity values as it was done by Zhu et al. [16]. in their research for the study of genetic diversity in upland cotton varieties.

4. CONCLUSION

From results generated from selected SSR markers, we concluded that our 75 *G. hirsutum* varieties are highly polymorphic. Our goal was to check the genetic diversity of cotton genotypes to prove their reliability and further use in breeding for improved varieties. These varieties showed a high mean value of genetic diversity of 0.877 proving that these are diverse as compared to their wild relatives.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Lee JA and Fang DD. Cotton as a world crop: origin, history, and current status. Cotton. 2015; 57:1-23.
DOI: 10.2134/agronmonogr57.2013.001
2. Munir H, Rasul F, Ahmad A, Sajid M, Ayub S, Arif M, et al. Diverse uses of cotton: From products to byproducts. Cotton Production and Uses: Agronomy, Crop Protection, and Postharvest Technologies. 2020: 629-641.
3. Imran MA, Khan MQ, Salam A, and Ahmad A. Cotton in nonwoven products. Cotton

Science and Processing Technology: Gene, Ginning, Garment and Green Recycling. 2020: 305-332.
DOI: 10.1007/978-981-15-9169-3_12

4. Esteve-Turrillas FA, de La Guardia M. Environmental impact of Recover cotton in textile industry. Resour Conserv Recycl. 2017;116:107-115.
DOI: 10.1016/j.resconrec.2016.09.034
5. Doyle JJ,Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem bull; 1987.
6. Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005. 21(9):2128-2129.
DOI: 10.1093/bioinformatics/bti282
7. Kilian B, Graner A. NGS technologies for analyzing germplasm diversity in genebanks. Brief Funct Genomics. 2012;11(1): 38-50.
DOI: 10.1093/bfgp/elr046
8. Van Treuren R, van Hintum TJ. Next-generation genebanking: Plant genetic resources management and utilization in the sequencing era. Plant genet resour. 2014;12(3):298-307.
DOI: 10.1017/S1479262114000082
9. Rafalski JA. Association genetics in crop improvement. Curr Opin Plant Biol. 2010;13(2):174-180.
DOI: 10.1016/j.pbi.2009.12.004
10. Sun F, Ma J, Shi W, and Yang Y. Genome-wide association analysis revealed genetic variation and candidate genes associated with the yield traits of upland cotton under drought conditions. Front in Plant Sci. 2023;14:1135302.
11. Venkatesulu S, Makula SP, Satihal MB, Puligundla SK, Srikanth K. Identification of SSR Molecular Markers for Jassid Resistance in Cotton.
DOI: 10.18805/ag.D-5740
12. Sagar T, Kapoor N, and Mahajan R. Microsatellites as potential molecular markers for genetic diversity analysis in plants, in molecular marker techniques: A Potential Approach of Crop Improvement. 2023; Springer 81-101
13. Zhao Y, Wang H, Chen W, Li Y, Gong H, Sang X, et al. Genetic diversity and population structure of elite cotton (*Gossypium hirsutum* L.) germplasm revealed by SSR markers. Plant Syst Evol. 2015;301:327-336.
14. Moiana LD, Vidigal Filho PS, Gonçalves-Vidigal MC, Lacanallo GF, Galván MZ, de Carvalho LP, et al. Genetic diversity and population structure of cotton *(Gossypium hirsutum L. race latifolium H.)* using microsatellite markers. Afr J Biotechnol. 2012;11(54):11640-11647.
DOI: 10.5897/AJB11.4084
15. Fang DD, Hinze LL, Percy RG, Li P, Deng D, Thyssen GA. microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton *(Gossypium hirsutum L.) cultivars from major cotton-growing countries.* Euphytica. 2013;191: 391-401.
DOI: 10.1007/s10681-013-0886-2
16. Zhu L, Tyagi P, Kaur B, and Kuraparthy V. Genetic diversity and population structure in elite US and race stock accessions of upland cotton *(Gossypium hirsutum).* J Cotton Sci. 2019;23(1):38-47.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/106008