

Original Research Article

MOLECULAR DISSECTION FOR BOLL SHEDDING IN UPLAND COTTON USING MICROSATELLITES

Abstract

DNA markers application in marker-assisted breeding of cotton is handicapped due to low genetic diversity in cotton germplasm. The present study was designed to identify DNA markers, predominately simple sequence repeats (SSRs), associated with tolerance/resistance to heat stress as a consequence of boll shedding. To find out the genetic diversity a total of 24 cotton genotypes and 50 SSR primers were used. Total 288 alleles were produced with an average of 5.7 alleles per primer. Bootstrap cluster analysis used to generate a dendrogram that cluster the 24 accessions into two main clusters. Eleven out of 24 genotypes fall in a single cluster. Phenotypically H-4074 gives more diversity, while genotypically H-4074 sheared the same genetic background as H-4070, H-4091 and H-4090. Low genetic diversity was observed among both genotypic and phenotypic as maximum varieties fall in single group. This study helps for selecting diverse accessions with multiple phenotypic traits, which were drought to boll shedding. It suggests further elaborating the molecular genetic diversity by using new SSR marker to improve the yield of cotton cultivars. These preliminary results set the stage for initiating in depth marker-trait association studies, which will be instrumental for initiating marker-assisted breeding in cotton.

Keywords: SSRs, Polymorphism, Boll shedding, Genetic diversity

INTRODUCTION

Cotton (*Gossypium* spp.) is considered as the foremost natural fiber and oil source worldwide. It is indigenous to subtropical and temperate areas of 80 countries around the world [1]. Cotton is classified in the genus *Gossypium* and family *Malvaceae* [2]. The genus *Gossypium* comprised about 50 species globally [3]. Economic impact of cotton around the world is about \$ 500 billion per year. Cotton has different essential tasks for example, give oil that is healthy for eaten, seed cake and good quality fabric. *Gossypium* genus contains 45 diploid species of cotton and 5 allotetraploid species [4]. There are four genotypes mostly used to cultivate, counting two diploids, *G. arboreum* L and *G. herbaceum* L. gives two percent of the world's cotton while two tetraploids species *Gossypium hirsutum* (*G. hirsutum*) that contribute about ninety percent and *G. barbadense* yield eight percent [5]. The most cultivated species is *G. hirsutum* cotton. It has highly extensive adaptability and increased product yield. These four varieties of cotton are reported round about greater than 95% of cotton producing all over the world [6].

Cotton farming is subsequently widespread, it comprised 32-34 million hectares worldwide in 2016 (USDANASS) [7]. Cotton production in 2016 around the world was 103.17 million bales. Cotton production plays an essential role in strong economy of Pakistan as it holds 1.0% share in GDP value of Pakistan. It was estimated that round about 1.5 million peoples of Pakistan income resource is farming cotton. The world's fourth large cotton producing country is Pakistan [8]. Despite its economic value, unfortunately the cotton yield in Pakistan among this year becomes low. Now it is the current requirement of Pakistan to sort out some recent cotton cultivars which are most resistant to any environmental state, gives better quality and high yield to meet the enough demand of textile fiber [9]. These can even better tolerate against biotic and abiotic stresses [10].

Researchers thought that fiber and yield traits are main goals for better development and improvement in cotton quality. For this purpose, Marker selection considered to be a recent technique over genetic engineering for improving crop plants. Markers are the specific landmarks or DNA tags to identify the certain other position of the DNA. These are special landmarks that show behavior similar to other landmarks of genome, concerned with expression or phenotype [11]. So, markers are used to identify such sequence of DNA that are not known. Hence the DNA markers are the enriched genomes assets which are used for getting the advanced worth and yield of cotton [12]. There are various numbers of markers used for various reasons, like they can differentiate among dominant and co-dominant loci, homozygosity and heterozygosity of particular trait under study. Among all of the markers co-dominant markers are more beneficial [13]. There is an undeniable need for highly polymorphic molecular markers if progress in plant breeding is to be made. Molecular marker technologies can be classified into hybridization based, PCR based and sequenced based markers on the basis of their working mechanism. Among these, PCR-based markers, simple sequence repeats microsatellites (SSRs) represent the major class of markers in cotton genomics due to their high utility and exploitation. These are repetitions in genome ranging from 2 to 5 nucleotides and randomly repeated 5-50 times. For a genomic characterizations and consistent estimation of gene diversity, SSR (simple sequence repeat) markers are one of the best option of markers as their high genome exposure, co-dominant inheritance, sufficient informative character, comparative abundance, multi-allelic character and their reproducibility have been generally used in cotton expansion [14]. Simple sequence repeat (SSR) markers are mainly categorized into expressed sequence tag (EST)-SSR and genomic SSRs (gSSRs) [15]. Understanding of the genetic relationships among plant

genotypes is significant to know the complexity of available germplasm, to discover the differences in available genotypes and to build up useful conservation plans.

The purpose of this study is to obtain the high yield and good quality cotton varieties for stable economic status of the Pakistan and identification of lines with different genotypes of cotton germplasm by decreased boll shedding and improved yield in cotton. This study will reveal the number of good quality SSR markers for reduction of boll shedding and improved the yield of cotton.

MATERIALS AND METHODS

Cotton Material

We obtained seeds of 24 genotypes of *G. hirsutum* two varieties 12 heat-tolerant and 12 heat-susceptible from Central Cotton Research Institute (CCRI) Multan, Pakistan. Heat tolerant genotypes were selected and grown in field of the Faculty of Agriculture Sciences and Technology Bahauddin Zakariya University, Multan.

Genomic DNA Extraction and Quantification

Fresh leaf specimens were taken from 10 particular plants of each variety. The extraction of DNA was carried out by using CTAB protocol [16]. DNA quality and quantity was analyzed by using 0.8% Agarose gel electrophoresis.

Molecular Markers

A total of 50 SSR primers were used to identify polymorphisms among the 24 cotton genotypes. The sequence of each primer was obtained from Cotton Marker Database [17]. These primers were design from Oligo Humanizing Genomics Macrogen.

RESULTS

Genotypic analysis

Allele number

Genetic analysis of boll shedding between the 24 genotypes of *G.hirsutum* was evaluated through utilizing 50 SSR primers which randomly cover the entire genome. With the help of these markers over-all 288 loci were amplified with an average of 5.7 loci of each SSR primer. The allele range of primers was found from one to eight. Seven SSR markers (JESPER 153, BNL-2448, NAU-2838, NAU-5465, NAU-4042, DPL-0323 and DPL-0519) were amplified the maximum number of bands (8) while the minimum number of loci were amplified (4 band) by 10 primers (BNL-2662, BNL-1066, BNL-3948, NAU-5269, NAU-2715, DPL-0079, NAU-2161, BNL-2449, BNL-2709 and NAU-4105). Average size of band was among 100 to 500 bp.

Genetic Frequency

To evaluate the gene diversity of 24 genotypes of cotton, power marker v 3.25 was operated. Having the mean 0.62 gene variability, was fluctuated from 0.00 (NAU2714, NAU1371) towards 0.93 (JESPER153). Although seven SSR primers (NAU4042, NAU5121, NAU1070, BNL1672, NAU2868, JESPER153, NAU3141) contributed the gene diversity at extreme point concerning boll shading as compared to other primers (SSR markers) having a gene diversity value 0.84, 0.85, 0.86, 0.88, 0.84, 0.93, 0.85 correspondingly. Least gene diversity was assessed among no three primers (NAU2714, NAU1371 and NAU1367) with value of 0.00, 0.00 and 0.15 respectively as given in Table 1.

Table 1: SSR polymorphism revealed by 50 SSR primer pairs.

Marker	Allele Frequency	Gene Diversity	PIC	Marker	Allele Frequency	Gene Diversity	PIC
NAU-5270	0.4167	0.7014	0.6548	NAU-5046	0.4583	0.7326	0.7064
NAU-2437	0.7500	0.4271	0.4153	BNL-1672	0.2083	0.8889	0.8791
NAU-2714	1.0000	0.0000	0.0000	BNL-3948	0.5417	0.6563	0.6251
BNL-2651	0.3750	0.7569	0.7221	NAU-5269	0.5417	0.5903	0.5208
JESPER-153	0.1250	0.9306	0.9263	NAU-5465	0.2917	0.8299	0.8097
NAU-2868	0.2917	0.8403	0.8233	NAU-3911	0.3333	0.7708	0.7382
NAU-980	0.5417	0.6701	0.6478	BNL-4096	0.4583	0.7083	0.6714
BNL-2448	0.5833	0.6250	0.6005	NAU-4042	0.2500	0.8472	0.8318
NAU-2838	0.6667	0.5069	0.4633	NAU-2715	0.5417	0.6563	0.6252
BNL-2662	0.8750	0.2188	0.1948	BNL-3502	0.4167	0.7500	0.7191
BNL-3442	0.2500	0.7986	0.7698	DPL-0079	0.8333	0.2951	0.2806
NAU-2954	0.2917	0.7882	0.7566	DPL-0323	0.5417	0.6806	0.6647
NAU-3414	0.2083	0.8472	0.8281	DPL-0519	0.7083	0.4826	0.4668
NAU-1070	0.2500	0.8681	0.8557	JESPER-36	0.6250	0.5903	0.5751
JESPER-134	0.3333	0.8021	0.7777	NAU-5121	0.2917	0.8472	0.8322
BNL-1066	0.3750	0.7396	0.6993	NAU-2095	0.5417	0.6667	0.6420
NAU-1366	0.5652	0.4915	0.3707	BNL-3171	0.3750	0.8021	0.7836
NAU-1367	0.9130	0.1588	0.1462	JESPER-292	0.3750	0.7569	0.7211

NAU-1368	0.7826	0.3403	0.2824	JESPER-101	0.6667	0.5243	0.4959
NAU-1369	0.7826	0.3403	0.2824	JESPER-110	0.4167	0.7083	0.6611
NAU-1370	0.7391	0.3856	0.3113	JESPER-274	0.3750	0.7639	0.7322
NAU-1371	1.0000	0.0000	0.0000	BNL-786	0.4167	0.6667	0.6057
NAU-1372	0.9130	0.1588	0.1462	NAU-998	0.8750	0.2292	0.2212
NAU-2836	0.2917	0.7986	0.7690	NAU-2161	0.4167	0.7743	0.7541
NAU-5172	0.7500	0.4167	0.3929	BNL-2449	0.7083	0.4514	0.4040
BNL-827	0.4167	0.6667	0.6057	BNL-2709	0.4583	0.6319	0.5606
NAU-4105	0.3750	0.7222	0.6775	BNL-3103	0.2917	0.8299	0.8100
Mean	0.5057	0.6213	0.5906	Mean	0.5057	0.6213	0.5906

PIC

In the assembly of genetic maps and molecular tagging, SSRs have been effectively utilized. Highest polymorphism was revealed in (JESPER153) having value (PIC) of 0.92, after that (BNL1672, NAU1070, NAU5121, NAU4042) having value of 0.87, 0.85, 0.84, 0.83 correspondingly. Least polymorphism stayed experiential in (NAU2714, NAU3172 and NAU3167) having a PIC value of 0.00, 0.14, and 0.14. The mean value of PIC is 0.59 of all traits that were observed as given in Table 1.

Similarity Index

A similarity matrix utilizing procedure was developed to set a level of similarity among the 24 genotypes. The average pair-wise similarity was found to be 0.76 fluctuated from 0.32 to 0.89. Least similarity was estimated amongst BH-160 and CYTO-124 (0.89), CRIS-134 and CYTO-124 (0.87) while maximum similarity was obtained between H-4090 and H-4091 (0.32). Genotypes like CIM-602, CIM707, CIM534, MNH886, HV-4088, HV-325, H-4091, H-4070 and FH-114 sheared common similarity index of (0.76) as given in Table 2.

Table 2 : Similarity index of 24 Cotton Genotype

	BH-160	BT-121	CEMB-66	CEMB-88	CIM-534	CIM-482	CIM-496	CIM-506	CIM-554	CIM-600	CIM-602	CIM-707	CRIS-134	CYTO 124	CYTO-608	FH-114	H-4070	H-4074	H-4090	H-4091	HV-325	HV-4088	MNH-786	MNH-886
BH-160	0.0000																							
BT-121	0.7679	0.0000																						
CEMB-66	0.5000	0.6786	0.0000																					
CEMB-88	0.6429	0.6607	0.6250	0.0000																				
CIM-534	0.7321	0.6786	0.7679	0.7321	0.0000																			
CIM-482	0.4286	0.6964	0.5536	0.5893	0.7500	0.0000																		
CIM-496	0.4107	0.6429	0.4643	0.6429	0.7679	0.3750	0.0000																	
CIM-506	0.5893	0.6964	0.6071	0.6607	0.7500	0.5179	0.5179	0.0000																
CIM-554	0.7321	0.5357	0.6786	0.6429	0.7321	0.6250	0.6071	0.5714	0.0000															
CIM-600	0.7321	0.6250	0.6964	0.6429	0.6607	0.6786	0.7143	0.7500	0.7679	0.0000														
CIM-602	0.7679	0.6607	0.8036	0.7500	0.6071	0.7857	0.7679	0.8393	0.6964	0.6429	0.0000													
CIM-707	0.7321	0.6786	0.6964	0.6071	0.8036	0.6607	0.5893	0.4821	0.5893	0.6607	0.7679	0.0000												
CRIS-134	0.4821	0.6786	0.4286	0.6786	0.7321	0.5536	0.4286	0.6429	0.7321	0.7321	0.7857	0.6964	0.0000											
CYTO 124	0.8980	0.7755	0.7959	0.7755	0.6122	0.7755	0.8163	0.7959	0.8367	0.6939	0.6531	0.7347	0.8776	0.0000										
CYTO-608	0.8214	0.7321	0.8036	0.8214	0.7143	0.8036	0.8393	0.7857	0.7857	0.7321	0.6964	0.8571	0.8214	0.7143	0.0000									
FH-114	0.6964	0.5536	0.5893	0.6071	0.7321	0.6964	0.6250	0.5714	0.5179	0.6607	0.7679	0.6429	0.6250	0.8163	0.8393	0.0000								
H-4070	0.5536	0.6250	0.5893	0.5893	0.7679	0.4821	0.3750	0.4821	0.5714	0.7143	0.7679	0.5357	0.5893	0.7959	0.8393	0.6607	0.0000							
H-4074	0.5893	0.6607	0.5357	0.6071	0.7857	0.6071	0.5357	0.5714	0.6607	0.6964	0.7857	0.5714	0.5179	0.8367	0.8393	0.6429	0.3571	0.0000						
H-4090	0.6250	0.6429	0.5000	0.5893	0.7857	0.5714	0.5179	0.5714	0.6607	0.6786	0.8036	0.6250	0.5357	0.8163	0.8214	0.6071	0.4821	0.4464	0.0000					
H-4091	0.5714	0.6429	0.5357	0.5893	0.8036	0.5179	0.4643	0.5536	0.6786	0.6786	0.7679	0.6607	0.5179	0.8367	0.7679	0.6250	0.4286	0.3750	0.3214	0.0000				
HV-325	0.7679	0.7143	0.7679	0.5536	0.6607	0.7679	0.7857	0.7321	0.7679	0.7679	0.7679	0.7500	0.7321	0.8163	0.7679	0.7321	0.6964	0.6607	0.7143	0.6964	0.0000			
HV-4088	0.6250	0.6786	0.6071	0.5357	0.7679	0.5714	0.5893	0.5893	0.6071	0.6964	0.8571	0.5714	0.6250	0.8367	0.7679	0.6429	0.5357	0.5714	0.5357	0.5893	0.6964	0.0000		
MNH-786	0.5714	0.5893	0.5536	0.5357	0.7321	0.5357	0.4821	0.4286	0.6607	0.6786	0.8036	0.5179	0.5893	0.7755	0.8214	0.5536	0.4464	0.5179	0.4464	0.4821	0.6071	0.5179	0.0000	
MNH-886	0.4643	0.7143	0.3750	0.6607	0.7321	0.4821	0.4286	0.5357	0.6786	0.7321	0.7679	0.7143	0.4464	0.8367	0.8214	0.6786	0.5357	0.5714	0.5714	0.5000	0.7321	0.6071	0.5893	0.0000

Phylogenetic Tree

In order to approximate the genetic variation present among all studied clusters, a Ward's dendrogram was constructed given in Figure 1 as described previously [18]. The dendrogram showed the presence of wide variation among the clusters suggesting high variability among genotypes. Based on the phylogenetic cluster given in Figure 1 and Ward's dendrogram analyses, all the accessions are clustered into two main clusters as A and B.

Cluster A contains 5 accessions and sub divided into two sub cluster named as A1 and A 2. Sub group A1 contained isolated CYTO-608 while A2 further separated into sub-sub group (A2-a) comprising variety CIM-600 and (A2-b) more divided into sub-sub group (A2-b-i) contained CYTO-124 and (A2-b-ii) contained two varieties CIM-534 and CIM-602. Cluster B contains 19 genotypes by dividing into sub groups B1 and B2. Sub groups B1 contains only one genotype HV-325. Sub group B2 more divided into (B2-a) and (B2-b). Sub-sub group (B2-a) is additionally distributed into two sub groups (B2-a-i) and (B2-a-ii). Sub-sub group (B2-a-i) contains only one genotype i.e. BT-121, while sub-sub group (B2-a-ii) is further sub-divided into two groups (B2-a-ii-a) contains one genotype i.e. CIM-554 and (B2-a-ii-b) contains the FH-114 genotype. Sub-sub group (B2-b) is further divided into (B2-b-i) and (B2-b-ii). Sub group (B2-b-i) is further divided into (B2-b-i-a) having only one variety i.e. CEMB-88 and (B2-b-i-b) contains HV-4088.

Sub-sub group (B2-b-ii) is further sub-divided into (B2-b-ii-a) and (B2-b-ii-b). Sub group (B2-b-ii-a) is sub-divided into (B2-b-ii-a-i) contains one genotype i.e. CIM-707 and (B2-b-ii-a-ii) further divided into (B2-b-ii-a-ii-á) having variety CIM-506 and (B2-b-ii-a-ii-â) having a variety MNH-786. (B2-b-ii-b) is further divided into two groups that are (B2-b-ii-b-i) and (B2-b-ii-b-ii). Sub group (B2-b-ii-b-i) is further two sub-divided into sub groups that are (B2-b-ii-b-i-á) contains two genotypes named as H-4090 and H-4091 and (B2-b-ii-b-i-â) contains two genotypes named as H-4070 and H-4074.

Sub group (B2-b-ii-b-ii) is further divided into two sub groups (B2-b-ii-b-ii-a) which is also further divided into two sub groups (B2-b-ii-b-ii-a-i) contains only one variety named as BH-160 and (B2-b-ii-b-ii-a-ii) contains two genotypes named as CIM-482 and CIM-496.

Sub-group (B2-b-ii-b-ii-b) is further divided into two groups that are (B2-b-ii-b-ii-b-i) contains only one genotype named as CRIS-134 and (B2-b-ii-b-ii-b-ii) contains two genotypes named as CEMB-66 and MNH-886.

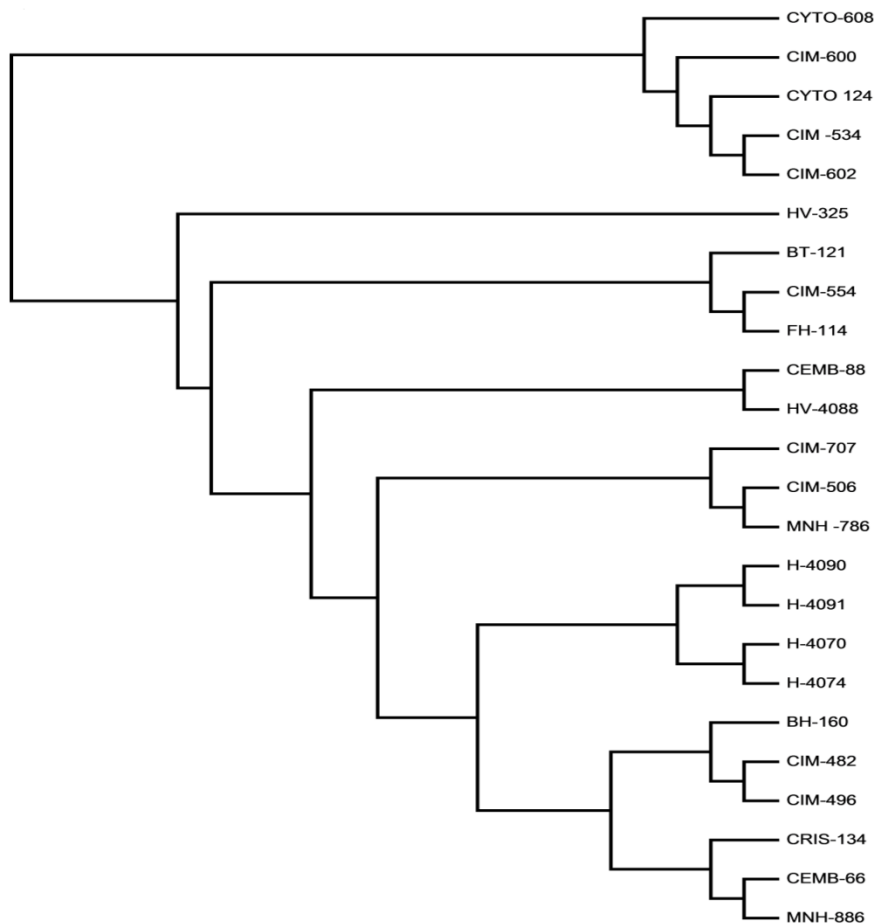


Figure 1: Bootstrap NJ tree showing genetic relationship among 24 *G. hirsutum* accessions by SSR markers

DISCUSSION AND CONCLUSION

The total 288 alleles were effectively amplified. The range was 4 to 8 alleles and mean was 5.7 alleles for each primer. It is combinely 76% instructive. The mainly informative primer was JESPER-153 that gives the polymorphism up to 92%. BNL-1672 formed 7 bands and accountable for 87% polymorphism. Approximately 7 primers NAU-1070, NAU-4042, NAU-5121, NAU-3414, NAU-2868, BNL-3103 and BNL-3410 are informational than 80%, however one was capable to distinguish the entire varieties. Followed by SSR makers be able to be given away major genetic assortment amongst closely connected genotypes still when a small number of loci are utilized. SSR markers represent large quantity of genetic differences amongst similar genotypes [19].

The similarity matrix from SSR markers, which were computed to construct a phylogenetic bootstrap neighbor joining tree based on Nei 1973 resemblance index was separated the 24

genotypes under research into diverse clusters according to phylogenetic resemblance. Two main clusters are created base on parentage i.e. tolerant and susceptible cotton varieties. Our results revealed that H-4090, H-4091, H-4070 and H-4074 have strong genetic association and fall in the same group. In our results CYTO-608, CIM-600, CYTO-124, CIM-534 and CIM-602 sheared the same cluster. These results describe the cultivars of cotton which are genotypically more stable regarding cotton boll shedding tolerant.

REFERENCES

1. Hinze LL, Hulse-Kemp AM Wilson, IW, Zhu Q H, Llewellyn, DJ, Taylor JM, Spriggs A, Fang DD, Ulloa M, Burke JJ, Giband M, Lacape JM, Van Deynze A, Udall JA, Scheffler JA, Hague S, Wendel JF, Pepper AE, Frelichowski J, Lawley CT, Jones DC, Percy RG, Stelly DM. Diversity analysis of cotton (*Gossypium hirsutum* L.) germplasm using the CottonSNP63K Array. *BMC Plant Biol.* 2017.17, 37. <https://doi.org/10.1186/s12870-017-0981-y>.
2. Khalid S, Zia-ur-Rehman M, Hameed U, Saeed F, Khan F, Haider MS. Transmission specificity and coinfection of mastrevirus with begomovirus. *Int. J. Agric. Biol.* 2017.19, 629-634.<https://doi.org/10.17957/IJAB/15.0249>.
3. Wendel JF, Grover CE. Taxonomy and evolution of the cotton genus, *Gossypium*. *Cotton* 2015.25-44.
4. Parekh MJ, Kumar S, Zala HN, Fougat R.S, Patel CB, Bosamia TC, Kulkarni KS, Parihar A. Development and validation of novel fiber relevant dbEST–SSR markers and their utility in revealing genetic diversity in diploid cotton (*Gossypium herbaceum* and *G. arboreum*). *Indus Crops Prod.* 2016.83, 620-629.
5. Tyagi P, Gore MA, Bowman DT, Campbell BT, Udall JA, Kuraparthi V. Genetic diversity and population structure in the US Upland cotton (*Gossypium hirsutum* L.). *Theoretical and Applied Genetics.* 2016.1;127(2):283-95.
6. Nie X, Huang C, You C, Li W, Zhao W, Shen C, Zhang B, Wang H, Yan Z, Dai B, Wang M. Genome-wide SSR-based association mapping for fiber quality in nation-wide upland cotton inbred cultivars in China. 2016 *BMC genomics.*17(1):1-6.
7. Kumar KS, Ashokkumar K, Ravikesavan R, Ashokkumar K. *Genetics of Yield Traits, Seed Cotton Yield and Fibre Quality Traits in Upland Cotton (Gossypium hirsutum L.)*. *Indian J. Sci.* 2016;9:66-75.
8. Ali I, Shakeel A, Ali A, Sadia B. Genetic basis of variation for within-boll yield components in cotton. *Turkish Journal of Agriculture and Forestry.* 2016 5;40(1):18-24.
9. Wang ML, Barkley NA, Jenkins, TM. Microsatellite markers in plants and insects. Part I: Applications of biotechnology. *Genes, Genom and Genomic.* 2007. 3, 54-67.
10. Zhao D, Reddy, KR, Kakani, VG, Reddy V. Nitrogen deficiency effects on plant growth, leaf photosynthesis, and hyperspectral reflectance properties of sorghum. *Eur J Agron.* 2005.22, 391-403.
11. Collard BCY and Mackill DJ, 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Phil. Trans. R. Soc.* 2008. B363 (1491): 557–572.
12. Mishra KK, Fougat RS, Ballani A. Validation of fiber quality linked SSR markers derived from allotetraploid (*Gossypium hirsutum*) in diploid (*Gossypium arboreum*). *Int J Scient Res Knowledge.* 2013.1, 349.

13. Qin H, Chen M, Yi X, Bie S, Zhang C, Zhang Y, Lan J, Meng Y, Yuan Y, Jiao C. Identification of associated SSR markers for yield component and fiber quality traits based on frame map and upland cotton collections. *PloS one*. 2015 30;10(1):e0118073.
14. Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, Iyer VN. Evolution of genes and genomes on the Drosophila phylogeny. *Nature*. 2007.450, 203-218.
15. Aamodt K, Abelev B, Quintana AA, Adamova D, Adare AM, Aggarwal MM, Rinella GA, Agocs AG, Agostinelli A, Salazar SA, Ahammed Z. Particle-Yield Modification in Jetlike Azimuthal Dihadron Correlations in Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Physical review letters*. 2012 Mar 1;108(9):092301.
16. S. Porebski, L. G. Bailey, and B. R. Baum, "Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components, *Plant Molecular Biology Reporter*, 1997.15(1), 8–15.
17. Lin Z, Wang Y, Zhang X, Zhang J. Functional markers for cellulose synthase and their comparison to SSRs in cotton. *Plant Molecular Biology Reporter*. 2012 1;30(5):1270-5.
18. Grenier C, Bramel-Cox PJ, Noirot M, Rao KP, Hamon P. Assessment of genetic diversity in three subsets constituted from the ICRISAT sorghum collection using random vs non-random sampling procedures A. Using morpho-agronomical and passport data. *Theoretical and Applied Genetics*. 2000,1;101(1-2):190-6.
19. Zhang X, Ye Z, Wang T, Xiong H, Yuan X, Zhang Z, Yuan Y, Liu Z. Characterization of the global transcriptome for cotton (*Gossypium hirsutum* L.) anther and development of SSR marker. *Gene*. 2014.10;551(2):206-13.