REVIEW ARTICLE

Tightly Linked Molecular DNA Markers with High Predictive Trait Value: A Current Increasing Demand in the Breeding Program of Upland Cotton (*Gossypium hirsutum* L.)

Ravi P. Shukla[1,3], Gopal J. Tiwari[1], Babita Joshi[1,4], Satya N. Jena[1], Sushma Tamta[2*]

DOI: 10.18811/ijpen.v7i02.1

Abstract

Cotton (*Gossypium* spp.) is a major cash crop of India and is the second-largest cotton producer in the world after China. *Gossypium hirsutum* and *Gossypium barbadense* are two tetraploid species that are majorly cultivated besides two diploid species (desi cotton). In cotton, fiber quality, drought tolerance, boll weight, boll number, and yield are essential quantitative traits with many components that are controlled by several genes present at different loci. Identifying such genes from different genomic resources of cotton using various molecular markers is necessary to accelerate the Quantitative Trait Loci (QTL) analysis. In the public domain of cotton, there is a vast number of molecular markers. However, not all are very useful for trait mapping, as most markers are away from the QTL region. Thereby, cotton improvement programs pay more attention to tightly linked markers with high predictive trait values.

The present review provides an overview and updates on the comparative studies and the application of various molecular markers, i.e., RFLP, AFLP, RAPD, SSR, EST-SSR, and SNP in the cotton-breeding program. Insights gained from the study may help in successful cotton breeding and improvement.

Keywords: EST-SSR, Mapping population, Molecular markers, QTLs, SNP, Tightly linked molecular markers.

International Journal of Plant and Environment (2021); **ISSN:** 2454-1117 (Print), 2455-202X (Online)

Introduction

Conventional plant breeding has been in practice for 9,000–11,000 years, resulting in domestication and genetic improvement of wild relatives of crop plants for better yield. However, improved management procedures were also only able to raise food output to a limited extent, which was insufficient to fulfill the need of the fast-rising human population. Therefore, plant breeding techniques were primarily used for the green revolution to develop high-yielding and disease-resistant varieties of cereal crops such as wheat, rice, maize, millets, and a few other kinds of cereal. Phenotypic evaluation of superior genotype is one of the fundamental aspects of traditional plant breeding practices. However, genetic variability created in the crosses and selection for a particular trait is not the only result of genotype, i.e., fixed effect, but also environmental influence on it, i.e., random effect. Thus, conventional breeding is a purposeful and effective process, where genetic loci are randomly reassembled in the recombination event of meiosis to make the resulted organism more useful.

Plant breeding for the development of new varieties is carried out methodically by government institutions and commercial companies. It can enhance the economic benefit to the industry by transferring the technologies to produce large-scale improved varieties. The advent of Deoxyribonucleic acid (DNA) markers in early 1980 have been a game-changer in the whole plant breeding process. With advancements in sequencing technologies, various types of molecular markers have been developed to aid crop development. As a result, genomics-driven plant breeding is a current agricultural demand that is crucial in developing, releasing and commercializing novel crop varieties. Thus, genomics-driven plant breeding is the present-day agricultural demand, which is crucial in developing, releasing, and commercializing novel crop varieties. Molecular breeding using DNA-based markers that are tightly linked to phenotypic traits will assist the selection for a particular breeding objective. In the absence of DNA markers, the targeted new changes/improvement are not quickly possible for the breeding program. Few researchers still use the phenotypic trait-based marker, but it is often inadequate as environmental factors influence it.

Depending on the study's selection and objective, each DNA marker may have its strengths and weaknesses. Nevertheless, DNA-based markers have much more advantages compared to morphology-based markers. They have been used in various aspects of research such as evolution and phylogeny,

[1]Plant Molecular Genetics Laboratory, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow, Uttar Pradesh, India.

[2]Department of Botany, D.S.B. Campus, Kumaun University, Nainital, Uttarakhand, India

[3]Aakash Institute, Bhopal, Madhya Pradesh, India

[4]Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, Uttar Pradesh, India

***Corresponding author:** Sushma Tamta, Department of Botany, D.S.B. Campus, Kumaun University, Nainital, Uttarakhand, India, Email: sushmatamta@gmail.com

How to cite this article: Shukla, R.P., Tiwari, G.J., Joshi, B., Jena, S.N., & Tamta, S. (2021). Tightly Linked Molecular Dna Markers with High Predictive Trait Value: A Current Increasing Demand in the Breeding Program of Upland Cotton (*Gossypium hirsutum* L.). International Journal of Plant and Environment, 7(2), 101-118.

Conflict of Interest: None

Submitted: 15/05/2021 **Accepted:** 30/06/2021 **Published:** 15/07/2021

investigation of heterosis, identification of haploid/diploid plants and cultivars genotyping, genetic diversity assessment, backcrossing for a gene of interest, genetic mapping, QTL mapping, and many more. The vast growth of plant genome sequences and many plant genes' physiological and molecular functions have revolutionized molecular genetics and its vital importance in breeding programs. After cotton, after China, India is the world's second-largest cotton grower (www.icac.org). But cotton yield per hectare is still low compared to other major cotton-producing countries such as China and the United States. Yields differ depending on the type of agro-climatic conditions and seed used. Mexico has the greatest diversity of wild cotton species, followed by Australia and Africa. Cotton was domesticated in both the Old and New Worlds separately. *Gossypium* comprises about 50 species of trees, shrubs, and herbs (Fryxell *et al.* 1971). Among them, 5 are allotetraploid (2n = 4x = 52) and 45 are diploid (2n = 2x = 26). Diploid species (2n = 26) are classified into eight genomic groups based on homologous chromosomal pairing during meiosis and karyotyping: A, B, C, D, E, F, G, and K, while allotetraploid species comprise of At and Dt sub-genomes. Out of all reported species, only four are widely cultivated, i.e., two allotetraploid (*G. hirsutum* L. and *G. barbadense* L.) and two diploids (G. *herbaceum* L., and *G. arboreum* L.) species. All these four cultivated species were domesticated independently at different times in ancient history (Brubaker *et al.* 1999). Cotton conventional breeders generally select advantageous traits by looking at their phenotype. This approach has sometimes failed to improve traits governed by many genes, their interactions among themselves and with environments. In such cases, quantitative genetic provides the solution for the collective effect of each locus involved in a polygenic trait. Finding these DNA markers may help plant breeders select desirable genotypes at the plantlet stage, while conventional breeding has drawbacks at this level for desirable genotype selection for a plant breeder. The identification of polygenic traits can be inundated with different molecular markers. Many researchers have published several inter and intra-specific genetic linkage maps in cotton by using several DNA markers. RFLPs, AFLPs, RAPDs, SSRs, EST-SSRs, and SNPs are among the marker technologies used to speed up the construction of a genetic linkage map in cotton. Tanksley (1983) cataloged five properties of markers that can distinguish molecular marker from morphological marker viz.,(1) Occurrence of natural alleles at many loci (2) Phenotypic neutrality (3) Determination of genotype at the cellular level (4) Co-dominant allele at many loci and (5) Low level of pleiotropism. The classification of DNA markers has been classified into five categories (Fig. 1).

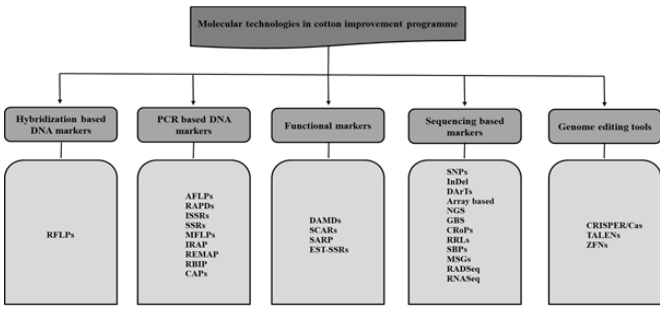

Fig. 1: Categorization of different classes of molecular DNA markers utilized for characterization, DNA finger printing, genetic linkage mapping, and genome editing

The present review emphasizes clearly understanding DNA markers and their applications for various purposes with particular attention to cotton. The review emphasizes tightly linked DNA markers with high predictive trait value for direct cotton improvement use. In addition, the review is also an emphasis on the available genetic map, various QTLs, which can be explored for cotton improvement through molecular breeding.

Molecular Markers in Cotton

Restriction Fragment Length Polymorphism (RFLP)

RFLP refers to the changes in two or more samples of homologous DNA molecules caused by different restriction site locations and a related laboratory technique in differentiating these segments. RFLP analysis was the first DNA profiling approach to become widely used due to its low cost. RFLP is an easy, absolute, and reproducible method and the first molecular marker technique developed for plant breeding. This technique has been widely used to monitor differentiation in DNA sequence-based based upon restriction sites among species. RFLP has been engaged in many crops to know the additional information of the relationship between parents, the genes for different traits, and the naming of measurable characteristic loci.

Meredith reported the first RFLP evaluation in upland cotton (1992), Reinisch *et al.,* (1994) constructed the first detailed RFLP map to explore chromosome organization and evolution in cotton. 705 RFLP loci comprising 4,675 cM of the cotton genome were classified into 41 linkage groups. The major strength of RFLP markers is co-dominance and high reproducibility. The main shortcoming of RFLP is the lack of polymorphism. Additionally, probe or sequence information for PCR analysis is required, making the development of markers in species with unknown molecular information challenging and time-consuming.

Random Amplified Polymorphic DNA (RAPD)

It's a type of PCR reaction in which the DNA segments amplified are random. Researchers using RAPD make multiple arbitrary short primers (8–12 nucleotides), then perform PCR on a large template of genomic DNA with an expectation that the fragment would amplify. A complete profile can be drawn from an RAPD reaction by resolving the generated patterns. The RAPD is fast, simply assessable, and regardless of sequence data for primer designing among different marker technology. A tri-specific F2 mapping population in cotton was developed in Arkansas (Khan *et al*., 1998). Ninety F2 plants were derived from a cross between *G. hirsutum* and synthetic tetraploid cotton made of two diploid species *G. arboreum* (A2 genome) and *G.trilobum* (D8 genome). Yu *et al.,* (1998) constructed a cotton framework map based mainly on RAPD and RFLP markers, with some SSRs based on 171 F2 TM-1 x 3–79 people with total coverage of 3,855 cM, the 219 loci were organized into 40 linkage groups. Using diploid and aneuploid cotton strains, linkage groups were allocated to chromosomes. They detected 10 fiber-related QTLs. The significant advantages of RAPD markers are no requirement of DNA probe and sequence information.

It does not require any blotting or hybridization, making it a rapid and efficient sample. The key disadvantages of RAPD are its lack of reproducibility, dominance, and lack of prior knowledge of the amplification product's identity.

Amplified Fragment Length Polymorphism (AFLP)

AFLP is a multi-locus PCR-based fingerprinting technology that Keygene developed in the early 1990s. It uses arbitrary primers based on restriction enzyme site sequences to amplify restriction fragments in genetic research and practice in genetic engineering (Vos *et al*., 1995). AFLP polymorphism arises as a result of insertions or deletions in the amplified fragments. Its key objective of genetic variation is similar to RFLP, but it allows many loci to be analyzed simultaneously instead of analyzing one locus at a time. The technique is divided into three phases;

- Selective amplification of certain fragments using two PCR primers with appropriate adaptor and restriction site-specific sequences, followed by band pattern visualization,
- Total cellular DNA is digested with one or more restriction enzymes, and site-specific adapters are ligated to all restriction fragments,
- The amplicon is separated electrophoretically on a gel matrix, then the band pattern is visualized. This method detects many locus amplicons in a single PCR reaction (Vos *et al.,* 1995) and uncovers a vast number of polymorphisms scattered throughout the genome (Lukonge *et al.,* 2007).

The first report of the application of AFLP in cotton appeared in 1997 when Reddy *et al.* successfully detected AFLPs in cotton, which led to applying this technique for genetic mapping. The AFLP technique has been used as a powerful tool to estimate genetic diversity in significant crops such as rice (Hashimoto *et al.,* 2004), wheat (Moghaddam *et al.,*2005), maize (Beyene *et al.,* 2006), pearl millet (Allinne *et al.,* 2007), sunflower (Dong *et al.,* 2007) and peanut (Jiang *et al.,* 2007). Moreover, this technique is also helpful for cultivar identification and evolutionary relationship studies. AFLP and RAPD markers were used to construct a cotton linkage map (Altaf *et al.* 1997). The main strengths of the AFLP method include a high frequency of genome-wide polymorphic markers. Because of the high PCR annealing temperatures, there is a high reproducibility level and is relatively economical compared to other arbitrary-primed PCR-based molecular marker systems. It does not require any prior sequence information, like RAPD marker, and consequently is appropriate for analyzing germplasm, biodiversity, and genetic relationship studies in a wide range of species. The major problem of AFLPs is their dominant mode of inheritance and the difficulty in finding allelic variations at a given locus, even though co-dominant AFLP markers have been detected at a frequency of 4–15 percent among all polymorphic AFLP markers. (Waugh *et al.,* 1997; Lu *et al.,* 1998; Boivin *et al.,* 1999).

Simple Sequence Repeat (SSR)

SSRs or microsatellites are PCR-based genetic marker systems. In 1989, Litt and Luty coined the term microsatellite for the first time. Eukaryotes have groups of repeated DNA sequences in their genomes, which are referred to as satellites. Satellite DNA is the result of centrifugation that removes repeated DNA from total genomic DNA. Satellites, minisatellites, and/or microsatellites have been given to the sequences depending on the length of the repeated sequence. Because defining micro and mini can be difficult, the term "simple sequence repeat" (SSR) will be more helpful. The technique for using SSRs as genetic markers is that the repeat region varies in length between genotypes, but the DNA flanking the repeat is sufficiently conserved that the same primers will work across many genotypes. The difference in the length of the repeats between the two conserved sequences creates an SSR polymorphism between the two types. Because of their locus identity, high PIC value, multiallelic nature, and PCR base, SSRs are excellent markers. SSRs can be tested on a variety of systems. Microsatellites, on the other hand, have significant research costs and practical challenges.

The first genetic map completely SSR-based using intra-specific F2 populations included 86 loci comprising 666.7 cM, or around 14.8 percent of the cotton genome's total recombination length (Shen *et al.,* 2005). Recently, comparatively dense SSR-based genetic maps utilizing recombinant inbred lines (RILs) derived from intra-specific populations have been reported with a genome coverage of 865 (Wang *et al.,* 2007a) and 1024 cM (Shen *et al.,* 2007). Comparatively dense genetic maps have been constructed in interspecific crosses of cotton using SSRs. Zhang *et al.* (2002) constructed a map for double haploids from an inter-specific cross between, *G. hirsutum* L. acc. TM-1, and *G. barbadense* L. cv. Hai-7124 consists of 510 SSRs and 114 RAPDs. The 489 loci were grouped into 43 linkage groups, covering 3314 cM of the cotton genome. They utilized the monosomic and telo-disomic lines of *G. hirsutum* in a TM-1 background for chromosome association. SSRs have been employed to develop a genetic map that exploited another BC1 population derived from the interspecific cross [{(Guazuncho2 (*G. hirsutum*) × VH8- 4602 (*G. barbadense*)} × Guazuncho2], which covered 4400 cM of the cotton genome (Lacape *et al.,* 2003). With the addition of the 233 new SSR loci (Nguyen *et al.,* 2004), the map comprised 1160 loci and 5519 cM and provided comprehensive coverage of the genome of tetraploid cotton. SSR-based linkage maps are also developed from Recombinant inbred line populations derived from inter-specific crosses. One hundred and eighty-three RILs were generated from a cross between an Upland cotton genotype (TM-1) and a Pima cultivar (3-79), along with complex sequence repeats (CRS) (Park *et al.,* 2005).

Constructions of SSR-based genetic maps have also been exploited by the F2 mapping population using an interspecific cross. A total of 205 SSRs, 107 RAPDs, and 437 SRAPs were used for genotyping the mapping population of F2 generation derived from a cross between *G. hirsutum* cv. Handan 208 and *G. barbadense* cv. Pima 90 (Lin *et al.,* 2005). The developed map divides 566 loci into 41 linkage groups, each including at least three loci. SSR markers with known chromosome locations were used to assign 28 linkage groups to associated chromosomes. With a mean inter-locus distance of 9.08 cM, the map covered 5141.8 cM. This backbone map was improved by integrating 463 new loci (He *et al*., 2007), resulting in the development of a map having 1029 loci (625 SSRs, 58 RAPDs, 341 SRAPs, and 5 REMAPs) assembled into 26 chromosomes/linkage groups, totally covering 5,472.3 cM of the allotetraploid cotton genome. Construction of the cotton genetic map subsequently led to the identification of QTLs for several traits, including fiber-related characteristics (Zhang *et al.,* 2003; Park *et al.,* 2005; Shen *et al.,* 2006b, 2007; Abdurakhmonov *et al.,* 2007), economic attributes

(He *et al.,* 2005; Hua *et al*., 2007; Shen *et al*., 2007; Wang *et al*., 2007a) leaf defoliation parameters (Abdurakhmonov *et al*., 2005) morphological traits (Lacape *et al*., 2005; Wang *et al*., 2006a) and chlorophyll contents (Song *et al*., 2005a). Other uses of cotton microsatellite markers include physical mapping of Rf1 fertility restorer gene (Yin *et al*., 2006) and tagging of genes related to root-knot nematode resistance (Shen *et al*., 2006a; Wang *et al*., 2006b; Ynturi *et al*., 2006), virticulum wilt resistance (Bolek *et al*., 2005). Moreover, markers associated with fiber traits have also been used in marker-assisted selection (Zhang *et al*., 2003).

Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR)

Since the dawn of the genomics era, massive volumes of publically available DNA sequence data have been generated, including big collections of expressed sequence tags (ESTs) from a range of species. According to various reports, ESTs can be a rich source of SSRs, revealing polymorphisms in the originating taxon and related taxa. Therefore, a different laboratory-created several EST-SSRs and other recently published SSRs were used to enrich the genetic linkage map.

Single Nucleotide Polymorphism (SNP)

A single nucleotide is a single strand of DNA. Polymorphisms are a form of genetic marker that enables more precise trait mapping. As a result, SNPs are the most often used genetic mapping tool. SNPs generally occur throughout the genome, both euchromatin and heterochromatin region but the frequency of SNP is found more in the heterochromatin region. According to the wobble hypothesis, when an SNP occurs within a gene, it may change the encoded protein, which may cause phenotypic variation, but it may not necessarily cause variation in protein structure. Another essential feature of SNPs is that they are co-dominant, meaning that both alleles can be genotyped individually, which aids in identifying whether a person is homozygous or heterozygous. SNPs offer the option of constructing dense genetic maps, map-based gene cloning, marker-assisted selection, marker-assisted breeding, and hap map-based association studies.

Cotton was the first crop to have a large-scale SNP finding. Re-sequencing of 24 Upland cotton genotypes yielded roughly 1,000 SNPs and 300 InDels, out of which 200 SNPs were mapped in the TM-1 3-79 genetic map (Fang *et al*., 2012; Yu *et al*., 2012). Cotton SNP discovery has recently accelerated, and several SNP markers have been published (Byers *et al*. 2012; Lacape *et al*. 2012; Rai *et al*. 2013; Zhu *et al*. 2014). Despite this, the number of cotton SNP markers is still low compared to other key crops such as maize, soybean, and rice. For example, only 39,862 SNPs have been presented in the public domain (dB SNP) in cotton, with 38,782 coming from G. hirsutum and the rest 1,080 is from G. barbadense.

Important Applications of DNA markers in a cotton breeding program

Construction of saturated genetic linkage map in cotton

Before 1980, all genetic maps were constructed using morphological and isozyme markers (Wendel *et al*., 1992). However, the number of these markers was limited, many were stage-specific, and they were susceptible to environmental influences. This problem was overcome with the discovery of DNA markers, followed by their utilization in map construction. DNA markers are pieces of DNA that disclose mutations or variations and can be used to detect variability between genotypes or alleles of a gene in a population or gene pool for a specific sequence of DNA. Those fragments are linked to a specific region in the genome and can be detected using molecular technologies. Simply defined, a DNA marker is a tiny section of DNA sequence that varies between individuals due to polymorphism (base deletion, insertion, and replacement).

The DNA-based marker has been utilized for molecular characterization, many more genetic linkage maps (Table 1), and DNA fingerprinting of cotton (Khan *et al*. 2010).

Table 1: Details of genetic map including number of loci, number of LGs, population size and type, number of QTL in cotton using different markers. (Source: Data derived from cottongen database (http://cottongen.org/find/featuremap/summary)

Sl. No.	*Year*	*Map Name*	*Genome*	*Female Parent*	*Male Parent*	*Mapping population Size*	*Population Type*	*Number of LG*	*Number of loci*	*Number of QTL*
				A-genome Group						
1	1999	A1-97 x A2-47, F2	A	A1-097	Gaorani	58	F2	18	161	0
2	2005	SMA-4 x A1-97, F2	A	SMA-4	A1-097	167	F2	13	276	27
3	2008	Jiang-Ling-Zhong-Mian x Zhe-Jiang-Xiao-Shan-Lu-Shu, F2	A	Jiangling Zhong Mian	Zhejiang Xiaoshan Green Fiber	189	F2	13	267	0
				D-genome Group						
4	2004	Gtr x Gra, F2	D	-	-	62	F2	13	763	0
5	2018	D3K-57 x D3D-1, F2:3	D	D3K-57	D3D-1	188	F2:3	13	728	0
6	2018	D3K-57 x D3D-1, F2:3-P	D	D3K-57	D3D-1	188	F2:3	13	750	0
G-genome Group										
7	2003	Gos-5024 x Hyb 601-2, F2-a	G	Gos-5024	Hyb 601-2	94	F2	13	213	0

Sl. No.	*Year*	*Map Name*	*Genome*	*Female Parent*	*Male Parent*	*Mapping population Size*	*Population Type*	*Number of LG*	*Number of loci*	*Number of QTL*
8	2003	Gos-5024 x Hyb 601-2, F2-n	G	Gos-5024	Hyb 601-2	94	F2	15	176	0
				AD-genome Group						
9	1998	CAMD-E x Seaberry, F2	AD	Tamcot CAMD-E (PI 529633)	Sea Island Seaberry	271	F2	28	254	12
10	1998	HS-46 x MARCABUCAG8US-1-88, F2:3	AD	HS-46	MARCABUC-AG8US 1-88	96	F2.3	31	120	100
11	2000	Deltapine-61 x Seaberry, F2	AD	Deltapine 61	Sea Island Seaberry	180	F2	29	226	50
12	2000	MD-5678ne x Prema, F2:3	AD	MD5678 ne	AcalaPrema	119	F2.3	17	81	26
13	2000	MD-5678ne x Prema, F2:3	AD	MD5678 ne	MD5678 ne	119	F2.3	17	81	26
14	2000	MD-5678ne x Prema, F2:3	AD	AcalaPrema	AcalaPrema	119	F2.3	17	81	26
15	2000	MD-5678ne x Prema, F2:3	AD	AcalaPrema	MD5678 ne	119	F2.3	17	81	26
16	2002	DES 119-5 x MD51ne, F2:3	AD	DES 119-5	MD51 ne	150	F2.3	16	56	0
17	2002	HQ95-6 x MD51ne, F2:3	AD	Tamcot HQ 95-6	MD51 ne	199	F2.3	24	82	0
18	2002	Vsg x	AD	Vsg	TM-1 x Hai-7124, F2	58	DH	43	487	0
19	2004	Acala-44 x Pima S-7, F2	AD	Acala 44	Pima S-7	94	F2	42	392	7
20	2004	Sic'on x F-177, F2	AD	Sic'on	F-177	406	F2	35	262	175
21	2005	Guazuncho-2 x VH8-4602, BC1	AD	Guazuncho-2	VH8-4602	75	BC1	27	1274	149
22	2005	Guazuncho-2 x VH8-4602, BC2	AD	Guazuncho-2	VH8-4602	-	BC2	26	513	0
23	2005	Tamcot-2111 x Pima S-6, BC3F2	AD	Tamcot 2111	Pima S-6	22-184	BC3F2	0	0	0
24	2005	TM-1 x WT-936, F2	AD	TM-1	WT-936	82	F2	51	589	11
25	2005	Vsg x	AD	Vsg	TM-1 x Hai-7124, F2	73	DH	40	444	0
26	2005	Yumian-1 x T586, F2:3	AD	Yumian-1	Texas 586	117	F2:3	20	70	0
27	2005	TM-1 x Hai-7124, BC1	AD	TM-1	Hai-7124	138	BC1	15	61	31
28	2005	Yumian-1 x T586, F2:3	AD	Yumian-1	Texas 586	117	F2:3	27	94	21
29	2005	H18 x 3-79, F2	AD	CS-B18	Mar-79	173	F2	1	25	0
30	2005	H18 x TM-1, F2	AD	TM-1	CS-B18	345	F2	1	21	0
31	2005	H18 x TM-1, F2	AD	TM-1	TM-1	345	F2	1	21	0
32	2005	H18 x TM-1, F2	AD	CS-B18	CS-B18	345	F2	1	21	0
33	2005	H18 x TM-1, F2	AD	CS-B18	TM-1	345	F2	1	21	0
34	2005	NemX x SJ-2, RIL	AD	AcalaNem-X	Acala SJ-2	69	RIL	1	11	0
35	2005	7235 x TM-1, F2:3	AD	7235	TM-1	-	F2:3	20	84	18
36	2005	HS427-10 x TM-1, F2:3	AD	HS 427-10	TM-1	-	F2:3	17	56	24

Sl. No.	*Year*	*Map Name*	*Genome*	*Female Parent*	*Male Parent*	*Mapping population Size*	*Population Type*	*Number of LG*	*Number of loci*	*Number of QTL*
37	2005	PD6992 x SM3, F2:3	AD	PD 6992	Simian-3	-	F2:3	21	71	11
38	2006	TM-1 x 3-79, RIL	AD	TM-1	Mar-79	186	RIL	57	433	25
39	2006	Yumian-1 x T586, RIL	AD	Yumian-1	Texas 586	270	RIL	59	672	0
40	2007	7235 x TM-1, RIL	AD	7235	TM-1	258	RIL	40	156	116
41	2007	CCRI-36 x Hai-7124, F2	AD	ZhongMian Suo 36	Hai-7124	186	F2	26	1077	0
42	2007	Deltapine-61 x Texas-701, F2	AD	Deltapine 61	PI 165329	251	F2	17	73	0
43	2007	Handan-208 x Pima-90, F2:3	AD	-	-	#VALUE!	F2:3	44	1028	43
44	2007	Palmeri x K-101, F2	AD	Palmeri (unknown ID)	AZK 101	57	F2	26	2636	329
45	2007	Pima S-7 x Empire, F2	AD		Pima S-7	108-150	F2	31	251	13
46	2007	Pima S-7 x im, F2	AD	im, Gh	Pima S-7	124	F2	39	364	51
47	2007	Pima S-7 x n2, F2	AD	n2, Gh	Pima S-7	124	F2	42	364	25
48	2007	TM-1 x Hai-7124, BC1	AD	TM-1	Hai-7124	138	BC1	26	1790	0
49	2007	Zhongmiansuo-12 x 8891, RIL	AD	ZhongMian Suo 12	8891	180	RIL	32	132	117
50	2007	TM-1 x Hai-1, F2	AD	TM-1	Hai-1	1599	F2	1	7	0
51	2007	TM-1 x Hai-1, F2	AD	TM-1	Hai-1	1599	F2	1	36	0
52	2007	ST-474 x Auburn 634 RNR, F2	AD	Stoneville 474	Auburn 634 RNR	200	F2	1	5	0
53	2007	TM-1 x Hai-7124, BC1S1	AD	TM-1	Hai-7124	-	BC1S1	10	24	15
54	2007	60182 x Jun Mian 1, F2:3	AD	60182	Jun Mian 1	229	F2:3	2	21	41
55	2008		AD	Simian-3 x Sumian-12	Zhong-4133 x 8891	280	4WC	25	183	0
56	2008	3-79 x NM24016, RIL	AD	Mar-79	NM24016	60	RIL	19	85	0
57	2008	Deltapine x Giza-83, F2	AD	Deltapine	Giza 83	71	F2	22	140	0
58	2008	Hai-7124 x Junmian-1, BC1	AD	Hai-7124	Jun Mian 1	180	BC1	7	129	0
59	2008	Hai-7124 x Junmian-1, F2	AD	Hai-7124	Jun Mian 1	96	F2	2	49	0
60	2008	TM-1 x Hai-7124, BC1	AD	TM-1	Hai-7124	138	BC1	26	2247	0
61	2008	Xinluzao-1 x Hai-7124, F2	AD	Xin Lu Zao 1	Hai-7124	76	F2:3	41	432	0
62	2008	Yumian-1 x T586, F2:7	AD	Yumian-1	Texas 586	270	RIL	22	113	0
63	2009	AD-genome wide Reference Map	AD	-	-	-	-	26	7413	0
64	2009	Guazuncho-2 x VH8-4602, consensus	AD	-	-	-	-	26	1745	0
65	2009	Guazuncho-2 x VH8-4602, RIL	AD	Guazun-cho-2	VH8-4602	-	RIL	26	800	47
66	2009	HS-46 x MARCABUCAG8US-1-88, RIL	AD	HS-46	MARCABUCA-G8US 1-88	188	RIL	26	125	0

Sl. No.	*Year*	*Map Name*	*Genome*	*Female Parent*	*Male Parent*	*Mapping population Size*	*Population Type*	*Number of LG*	*Number of loci*	*Number of QTL*
67	2009	Monsanto SSR Bin Map,	AD	Deltapine 33 B	GB-0679	-	F2	26	1553	0
68	2009	Tetraploid Species Polycross	AD	-	-	260	SP	0	0	0
69	2009	H16 x T586, F2	AD	CS-B16	Texas 586	1259	F2	1	88	0
70	2009	H16 x T586, F2sel	AD	CS-B16	Texas 586	237	F2sel	1	25	0
71	2009	Yumian-1 x T586, RIL	AD	Yumian-1	Texas 586	270	RIL	60	602	36
72	2009	EH083 x Hai-7124, F2	AD	EH083	Hai-7124	1214	F2	1	3	0
73	2009	DH962 x Jinmian-6, F2	AD	DH962	JinMian 6	137	F2	80	767	0
74	2010	T582 x Hai-7124, F2	AD	Texas 582	Hai-7124	290	F2	1	14	0
75	2011	Emian-22 x 3-79, BC1	AD	Emian-22	Mar-79	141	F2	26	2318	0
76	2011	Li2 x DP5690, F2	AD	Li2	Deltapine 5690	136	F2	1	7	0
77	2011	GB713 x Nem-X, F2	AD	Inca Cotton	AcalaNem-X	300	F2	2	70	9
78	2011	Pima-S7 x NemX, F2	AD	Pima S-7	AcalaNem-X	94	F2	1	10	2
79	2011	TM-1 x 3-79, RIL	AD	TM-1	Mar-79	186	RIL	5	64	6
80	2012	TM-1 x 3-79, RIL	AD	TM-1	Mar-79	186	RIL	29	2084	0
81	2012		AD	-	-	172	4WC	69	890	69
82	2012	TM-1 x 3-79, RIL	AD	TM-1	Mar-79	186	RIL	26	2490	0
83	2013	GX1135 x GX100-2, F2	AD	GX1135	GX100-2	173	F2	34	310	63
84	2013	SG 747 x Giza 75, BIL	AD	Sure-Grow 747	Giza 75	146	BIL	29	392	67
85	2014	Monsanto Jointed Map	AD	-	-	-	Joint	26	96	0
86	2014	Baimian 1 x TM-1, F2:3	AD	Bai Mian 1	TM-1		F2:3	6	30	7
87	2014	IL138-A11-3 x TM-1, F2:3	AD	IL138-A11-3	TM-1	570	F2:3	2	11	6
88	2015	MCU-5 x Siokra 1-4, RIL	AD	-	-	-		53	1244	4
89	2015	Yumian-1 x T586, RIL	AD	Yumian-1	Texas 586	270	RIL	26	1678	141
90	2015	Prema x 86-1, RIL	AD	AcalaPrema	China 86-1 (CN-ZM-01746)	179	RIL	103	4153	22
91	2015	DH962 X Ji Mian 5, F2	AD	DH962	Ji Mian 5	-	F2	55	1235	36
92	2015	W10 x TM-1, F2:3	AD	W10	TM-1	140	F2:3	42	411	6
93	2015	GX1135 x GX100-2, RIL	AD	GX1135	GX100-2	177	RIL	14	260	29
94	2015	Emian-22 x 3-79, BC1	AD	Emian-22	Mar-79	141	BC1	15	352	1
95	2015	0-153 x sGK9708, RIL	AD	0-153	sGK9708	196	RIL	1	209	47
96	2015	DH962 x Ji Mian 5, RIL	AD	DH962	Ji Mian 5	178	RIL	61	729	218
97	2015	CCRI 12-4 x	AD	ZhongMian Suo 12-4	(AD)5-007	188	F2	26	2763	0
98	2015	CCRI-35 x Yumian-1, RIL	AD	ZhongMian Suo 35	Yumian-1	180	RIL	26	1273	79
99	2015	Yumian-1 x 7235, RIL	AD	Yumian-1	7235	180	RIL	26	1538	114
100	2015	3-79 x TM-1, F2	AD	Mar-79	TM-1	118	F2	26	19191	160

Sl. No.	*Year*	*Map Name*	*Genome*	*Female Parent*	*Male Parent*	*Mapping population Size*	*Population Type*	*Number of LG*	*Number of loci*	*Number of QTL*
101	2015	Phytogen 72 x Stoneville 474, F2	AD	PHY 72 Acala	Stoneville 474	93	F2	26	7171	0
102	2015		ADB	86-1 x G. anomalum	Su 8289	384	BC2F1	13	230	0
103	2016	TM-1 x	AD	TM-1		115	CSIL	36	709	513
104	2016	GX1135 x GX100-2, RIL	AD	GX1135	GX100-2	177	RIL	57	622	266
105	2016	GX1135 x VGX100-2, RIL	AD	GX1135	VGX100-2	180	RIL	39	308	133
106	2016	0-153 x sGK9708, RIL	AD	0-153	sGK9708	196	RIL	70	767	355
107	2016	CCRI-36 x G2005, RIL	AD	Zhong Mian Suo 36	G-2005	137	RIL	26	6434	559
108	2016	HS-46 x MARCABUCAG8US-1-88, RIL	AD	HS-46	MARCABUCA-G8US 1-88	188	RIL	26	1729	87
109	2016	PD 94042 x	AD	PD 94042	(AD)4-08	92	F2	29	1055	132
110	2017	Phytogen 72 x Stoneville 474, consbin	AD	-	-	-	RIL	26	3824	0
111	2017	0-153 x sGK9708, RIL	AD	0-153	sGK9708	196	RIL	26	1714	326
112	2017	CCRI-35 x Yumian-1, RIL	AD	ZhongMian Suo 35	Yumian-1	180	RIL	26	2050	150
113	2018	CCRI-35 x NH, F2:3	AD	ZhongMian Suo 35	NandanBadi Big Flower	277	F2:3	26	5178	263
114	2018	Yumian-1 x AcalaMaxxa, RIL	AD	Yumian-1	AcalaMaxxa	180	RIL	26	12116	189
115	2018	MGG-293-793, MAGIC	AD	-	-	960	MAGIC	26	269	132
116		Hypoaneuploid Hybrids	AD	-	-	-	haplo	0	0	0
117	2014	TM-1 x 3-79, RIL	-	TM-1	Mar-79	186	RIL	26	2486	75
118	2018	LMY28 x XLZ24, RIL	-	Lu Mian 28	Xin Lu Zao 24	231	RIL	28	4851	411
119	2018	HS-46 x MARCABUCAG8US-1-88, RIL	-		-	-	-	25	380	263
120	2019	901-001 x sGK156, F2	-	sGK156	901-001	250	F2	23	312	155
121	2020	CwCRI-35 x TX-41, F2	-	ZhongMian Suo 35	TX-0041	392	F2	1	25	0

Various PCR-based DNA markers, such as RAPD, AFLP, RGA, SSR, SNP, SRAP, etc have been utilized for cotton genome research (Zhang *et al*.2008) and provided in the public domain database of cotton.

Marker-assisted Selection (MAS)

MAS is an indirect selection method in which a trait of interest (such as quality, abiotic stress tolerance, disease resistance, and productivity) is chosen based on a marker (DNA/RNA variation, morphological, or biochemical) linked to the trait of interest rather than the trait itself. MAS has developed a reputation for becoming an effective method to improve quantitatively inherited traits; however, MAS for polygenic traits such as drought and yield has also been successfully characterized. Even though the marker-assisted selection is very successful for introgression and pyramiding of significant genes, many concerns remain unresolved, and transgenic introgression programs and, to a smaller extent, backcross alteration programs for simple traits continue to dominate. However, it is expected that the most significant demand in public sector MAS will be for single and few oligogenic traits that are difficult or expensive to screen using conventional breeding methods, for implementing marker-assisted selection in breeding programs, there are some fundamentals: (i). Finding a tightly linked marker to the gene alarmed and (ii). A breeding population that is polymorphic for the concerned marker and the gene. Although MAS is meaningful, it is a time-consuming program.

Nevertheless, MAS dramatically enhances the productivity and effectiveness of plant breeding programs as compared to

conventional breeding methods. Furthermore, from first to last quantitative genetics research in cotton potentially accelerates new molecular tools like microarrays, ESTs and proteomics are being introduced, allowing the identification of many genes and QTLs in cotton. We conclude that MAS is an effective method to improve crop failure traits in cotton, but complementing these features with high yield potential, recombination, and selection is required.

Marker-based Genetic Relationship among Different Species of Cotton

The genetic diversity and relationships of crop species and their wild relatives were studied using various molecular-marker techniques. Random amplified polymorphic DNAs (RAPDs) have been the most commonly employed of these approaches in cotton (Iqbal *et al.* 1997; Tatineni *et al.* 1996; Multani and Lyon 1995). However, low levels of polymorphism at the intraspecific and interspecific levels have hampered studies employing RFLPs (Wendal and Brubaker 1993) and allozymes (Wendal *et al.* 1992). The amplified fragment length polymorphism (AFLP) technique has also been used to explore genetic diversity in several crop species and their wild counterparts (Zabeau and Vos 1993; Hill *et al.* 1995; Vos *et al.* 1995; Powell *et al.* 1996; Maughan *et al.* 1996). This marker can discover many potential polymorphic loci (known as a "high multiplex ratio"). AFLP markers were recently employed by Vroh Bi *et al.* (1999) to trace the introduction of genetic material from a synthetic hybrid (tri-species) into *G. hirsutum*. The study's main aim was to investigate if AFLP could be used to assess genetic diversity and phenetic connections among 29 diploid and tetraploid Gossypium accessions and identify genetically distinct cultivars that might be crucial new sources of alleles for cotton development. Many researchers have published articles in different years to determine inter and intra-specific genetic relationships within a broad range of cotton species (Table 2).

Genetic Diversity Studies in Cotton

The genetic relationship among genotypes or different species helps to address the diversity of available germplasm and uncover differences in available genotypes for useful conservation programs (Dahab *et al.,* 2013). The depiction of genetic similarity across genotypes is a helpful tool for selecting parental combinations to maintain genetic diversity in our live breeding program's fabrics. The genetic diversity among and within the available germplasm is critical to the breeding program's success, as it allows the breeder must identify parental sources capable of generating numerous populations for selection. A summary of some diversity analysis publications by different research groups has been summarized (Table 3).

QTL Mapping for Different Traits in Cotton

The acceleration of QTL mapping in cotton by finding molecular markers is a very appreciable task published by many researchers in different mapping populations in different years (Table 4). The principle of QTL mapping is to find a link between the phenotype and genotype of markers. Therefore, QTL mapping in cotton is not limited only to fiber-related trait studies, but boll weight and boll number QTLs have been mapped.

Basic Steps for Developing New Varieties

A. Collection of Germplasm:

A collection of plants/seeds includes multiple alleles for all genes in a crop of interest. As a result, any crop species' germplasm would include the following sorts of material: We (i) Cultivated enhanced varieties (ii) Improved varieties no longer in cultivation (iii) Desi or local varieties (iv) Plant breeding lines (v) Wild species related to the crop species in question.

All these materials contain priceless alleles of genes that are important for breeding. Genetic variability is the backbone or basis of any breeding program. Therefore, compilation and preservation of all genotypes, species, and varieties is a

Table 2: Summary of genetic relationship studies in cotton by using different markers

S.No.	*Species*	*Marker*	*Year*	*References*
1	*G. herbacium, G. arborium, G. hirsutum*	SSR	2013	Mishra *et al*.
2	*Gossypium* 20 diploid species	RAPD	2007	Yuxiang *et al*.
3	*G. herbacium, G. hirsutum, G. barbadense*	SSR	2015	Liu *et al*.
4	Among *Gossypium* species	AFLP	2001	Abdala *et al*.

Table 3 Summary of genetic diversity studies in cotton by using different molecular markers

S.N	*Country*	*Population type*	*Markers used*	*References*
1	Pakistan	31 *Gossypium* species, 3 subspecies and 1 inter-specific hybrid	45 RAPD Primer	Khan *et al*., 2000
2	USA	24 Cultivars of *G. hirsutum*	88 SSR Primers	Zhang *et al*., 2005
3	India	24 Lines of *G.hirsutum*	6 AFLP Primer	Rana *et al*., 2005
4	USA	24 Line of cotton	270 Snp Loci and 92 indel	Van *et al*., 2009
5	India	51 Cultivars of four cotton species	16 AFLP Primer	Jena *et al*., 2011
6	India	Intraspecific cotton F1 hybrids and its Parents	20RAPD and 19 ISSR	Dongre *et al*., 2012
7	India	150 *G. hirsutum* lines	50 SSR	Rajeev *et al*. 2014
8	USA	395 Diverse *Gossypium* genotype	SNP 63K	Hinze *et al*. 2017

Table 4: Summary of QTL mapping studies in cotton for qualitative and qualitative traits (Source: Data derived from cottongen database, https://www.cottongen.org/find/qtl)

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
1	Boll size	Morphological trait	qBS	3	7.08–25.51	7235 x TM-1, RIL (2007)
2	Seed cotton weight per boll	Yield trait	qBWT	13	-	-
3	Boll number	Yield trait	qBN	3	5.93–14.16	7235 x TM-1, RIL (2007), Palmeri x K-101, F2 (2007)
4	Fiber elongation	Quality trait	qFEL	37	-	
5	2.5 span length	Quality trait	qSL	162	-	MD-5678ne x Prema, F2:3 (2000), TM-1 x 3-79, RIL (2006)
6	Arealometer A	Quality trait	qAREOA	12	-	HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
7	Bacterial blight resistance	Biotic stress trait	qBBR	14	-	Pima S-7 x Empire, F2 (2007),Palmeri x K-101, F2 (2007)
8	Boll number	Yield trait	Qbn	32	3.586–14.88	Palmeri x K-101, F2 (2007), 7235 x TM-1, RIL (2007), DH962 X Ji Mian 5, F2 (2015),Sic'on x F-177, F2 (2004), 3-79 x TM-1, F2 (2015 SNP)
9	Boll size	Morphological trait	qBS	5	4.04–19.46	7235 x TM-1, RIL (2007)
10	Boron concentration	Biochemical trait	qBCONC	5	0.09–0.18	Guazuncho-2 x VH8-4602, RIL (2009)
11	Calcium concentration	Biochemical trait	qCACONC	3	0.11–0.13	Guazuncho-2 x VH8-4602, RIL (2009)
12	Canopy temperature	Physiological trait	qCT	8		Palmeri x K-101, F2 (2007), Sic'on x F-177, F2 (2004)
13	Carbon isotope ratio	Physiological trait	q13C	14		Sic'on x F-177, F2 (2004)
14	Coefficient variation of length by number	Quality trait	qLNCV	2		CAMD-E x Seaberry, F2 (1998)
15	Copper concentration	Biochemical trait	qCUCONC	4	0.1–0.19	Guazuncho-2 x VH8-4602, RIL (2009)
16	Days flowering to boll open	Phenological trait	qDOFB	54	1.7101–13.4664	CCRI-36 x G2005, RIL (2016),
17	Days to 50% boll opening	Phenological trait	qBOLL50	80	3.198–18.2694	CCRI-36 x G2005, RIL (2016), 3-79 x TM-1, F2 (2015 SNP)
18	Days to 50% flowering	Phenological trait	qFLW50	77	0.2831–29.3695	CCRI-36 x G2005, RIL (2016), 3-79 x TM-1, F2 (2015 SNP)
19	Dry matter yield	Yield trait	qDM	8		Palmeri x K-101, F2 (2007), Sic'on x F-177, F2 (2004)
20	Embryogenic callus formation	Stature or vigor trait	qECF	6		W10 x TM-1, F2:3 (2015)
21	Fiber cell wall thickness	Quality trait	qWALL	8		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
22	Fiber elongation	Quality trait	qFEL	614		TM-1 x 3-79, RIL (2006), 7235 x TM-1, RIL (2007), Palmeri x K-101, F2 (2007), MD-5678ne x Prema, F2:3 (2000), Yumian-1 x T586, RIL (2009), MD-5678ne x Prema, F2:3 (2000), Zhongmiansuo-12 x 8891, RIL (2007)
23	Fiber fineness	Quality trait	qFF	134		Palmeri x K-101, F2 (2007), Yumian-1 x T586, FTM-1 x 3-79, RIL (2014), Sic'on x F-177, F2 (2004),Acala-44 x Pima S-7, F2 (2004), CAMD-E x Seaberry, F2 (1998), Guazuncho-2 x VH8-4602, BC1 (2005),Palmeri x K-101, F2 (2007), Yumian-1 x T586, F2:3 (2005.v0), 2:3 (2005.v0),

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
24	Fiber immature content	Quality trait	qFIM	8		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998), TM-1 x 3-79, RIL (2014)
25	Fiber length	Quality trait	qFL	701	0.0533–98.2	Handan-208 x Pima-90, F2:3 (2007),7235 x TM-1, RIL (2007) Yumian-1 x T586, RIL (2009),(YM-1 x CCRI-35) x (YM-1 x 7235), 4WC (2012) 7235 x TM-1, F2:3 (2005),HS427-10 x TM-1, F2:3 (2005),PD6992 x SM3, F2:3 (2005),Zhongmiansuo-12 x 8891, RIL (2007),Palmeri x K-101, F2 (2007),DH962 X Ji Mian 5, F2 (2015),IL138-A11-3 x TM-1, F2:3 (2014),TM-1 x (TX-0256 + TX-1046), CSIL (2016),CCRI-35 x NH, F2:3 (2018),Emian-22 x 3-79, BC1 (2015),0-153 x sGK9708, RIL (2016),CCRI-36 x G2005, RIL (2016),CCRI-35 x Yumian-1, RIL (2015),Yumian-1 x 7235, RIL (2015),Handan-208 x Pima-90, F2:3 (2007), 7235 x TM-1, RIL (2007),Yumian-1 x T586, RIL (2009) (YM-1 x CCRI-35) x (YM-1 x 7235), 4WC (2012),7235 x TM-1, F2:3 (2005),HS427-10 x TM-1, F2:3 (2005),PD6992 x SM3, F2:3, (2005),Zhongmiansuo-12 x 8891, RIL (2007),Palmeri x K-101, F2 (2007),DH962 X Ji Mian 5, F2 (2015),IL138-A11-3 x TM-1, F2:3 (2014),TM-1 x (TX-0256 + TX-1046), CSIL (2016),CCRI-35 x NH, F2:3 (2018),Emian-22 x 3-79, BC1 (2015),0-153 x sGK9708, RIL (2016),CCRI-36 x G2005, RIL (2016),CCRI-35 x Yumian-1, RIL (2015),Yumian-1 x 7235, RIL (2015)
26	Fiber length by weight	Quality trait	qFLW	4		TM-1 x 3-79, RIL (2014)
27	Fiber mature content	Quality trait	qFM	200	0.003–7.6129	MD-5678ne x Prema, F2:3 (2000),Zhongmiansuo-12 x 8891, RIL (2007),HS-46 x MARCABUCAG8US-1-88, F2:3 (1998),TM-1 x 3-79, RIL (2014),CCRI-35 x NH, F2:3 (2018),Guazuncho-2 x VH8-4602, BC1 (2005)
28	Fiber mean length	Quality trait	qML	14		Guazuncho-2 x VH8-4602, BC1 (2005)
29	Fiber nep size	Quality trait	qFNS	3		TM-1 x 3-79, RIL (2014)
30	Fiber perimeter	Quality trait	qFP	4		MD-5678ne x Prema, F2:3 (2000)
31	Fiber spin consistency index	Quality trait	qFSCI	31	0.0019–6.2241	CCRI-35 x NH, F2:3 (2018)
32	Fiber strength	Quality trait	qFS	905	0.0101–32.6	Handan-208 x Pima-90, F2:3 (2007), TM-1 x 3-79, RIL (2006), Yumian-1 x T586, RIL (2009), (YM-1 x CCRI-35) x (YM-1 x 7235), 4WC (2012), MD-5678ne x Prema, F2:3 (2000),Yumian-1 x T586, F2:3 (2005.v0),7235 x TM-1, F2:3 (2005),HS427-10 x TM-1, F2:3 (2005),PD6992 x SM3, F2:3

Sl. No.	Triat name	Trait category	Keyword for QTL	No. of reported QTLs	R^2 range	References
						(2005),Prema x 86-1, RIL (2015),Zhongmiansuo-12 x 8891, RIL (2007),HS-46 x, MARCABUCAG8US-1-88, F2:3 (1998),Palmeri x K-101, F2 (2007),7235 x TM-1, RIL (2007)
33	Fiber uniformity	Quality trait	qFU	485	7.4–50.5	Yumian-1 x T586, RIL (2009), (YM-1 x CCRI-35) x (YM-1 x 7235), 4WC (2012), 901-001 x sGK156, F2 (2019),GX1135 x GX100-2, RIL (2016), HS-46 x MARCABUCAG8US-1-88, RIL (2018),Pima S-7 x im, F2 (2007),Guazuncho-2 x VH8-4602, BC1 (2005),MGG-293-793, MAGIC (2018), Sic'on x F-177, F2 (2004), Yumian-1 x Acala Maxxa, RIL (2018), PD 94042 x (AD)4-08, F2 (2016)
34	First fruit branch position	Stature or vigor trait	qFFBH	60	3.476–24.074	HS-46 x MARCABUCAG8US-1-88, F2:3 (1998), CCRI-36 x G2005, RIL (2016),3-79 x TM-1, F2 (2015 SNP)
35	First fruit branch position by node	Physiological trait	qFFBN	65	3.225–23.075	CCRI-36 x G2005, RIL (2016), 3-79 x TM-1, F2 (2015 SNP)
36	Floral bract length	Morphological trait	qFLBL	13		SMA-4 x A1-97, F2 (2005), Pima S-7 x im, F2 (2007),
37	Flower rate	Phenological trait	qFLWRT	2		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
38	Fruit branch node number	Yield component trait	qFBNDN	3	3.225–4.007	3-79 x TM-1, F2 (2015 SNP)
39	Fruit branch number	Yield trait	qFBN	27	0.0323–0.0849	LMY28 x XLZ24, RIL (2018)
40	Fusarium wilt resistance	Biotic stress trait	qFWR	8		Pima-S7 x NemX, F2 (2011), TM-1 x 3-79, RIL (2011)
41	Gin turnout	Quality trait	qGTO	2		TM-1 x 3-79, RIL (2014)
42	Gynoecium length	Morphological trait	qGYNOL	10		Pima S-7 x im, F2 (2007), SMA-4 x A1-97, F2 (2005)
43	Harvest index	Yield trait	qHI	24		Palmeri x K-101, F2 (2007), Sic'on x F-177, F2 (2004)
44	Iron concentration	Biochemical trait	qFECONC	3	0.1–0.18	Guazuncho-2 x VH8-4602, RIL (2009)
45	Leaf 2nd lobe length	Morphological trait	qL2W	31		Deltapine-61 x Seaberry, F2 (2000), TM-1 x WT-936, F2 (2005)
46	Leaf angle between 2nd and 3rd lobes	Morphological trait	qL2L3A	13		Deltapine-61 x Seaberry, F2 (2000), TM-1 x WT-936, F2 (2005)
47	Leaf chlorophyll content	Physiological trait	qCHL	15	6.19–6.56	Palmeri x K-101, F2 (2007), TM-1 x Hai-7124, BC1 (2005.v0), Sic'on x F-177, F2 (2004)
48	Leaf depth between 2nd and 3rd lobes	Morphological trait	qL2L3D	2		TM-1 x WT-936, F2 (2005)
49	Leaf lobe number	Morphological trait	qLBN	3		TM-1 x WT-936, F2 (2005)
50	Leaf main lobe length	Morphological trait	qL1L	11		Deltapine-61 x Seaberry, F2 (2000), TM-1 x WT-936, F2 (2005)
51	Leaf shape	Morphological trait	qLFSP	54	4.72–17.54	TM-1 x Hai-7124, BC1 (2005.v0), MCU-5 x Siokra 1-4, RIL (2015 SNP),Palmeri x K-101, F2 (2007)

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
52	Leaf sublobe number on main lobe	Morphological trait	qL1SLBN	1		Deltapine-61 x Seaberry, F2 (2000)
53	Lint color	Morphological trait	qLINTCL	11		Palmeri x K-101, F2 (2007)
54	Lint index	Yield trait	qLI	24	0.0353–16.62	Handan-208 x Pima-90, F2:3 (2007), Zhongmiansuo-12 x 8891, RIL (2007), DH962 X Ji Mian 5, F2 (2015), TM-1 x 3-79, RIL (2014)
55	Lint percent	Yield trait	qLP	414	0.0103–63.4	7235 x TM-1, RIL (2007), Yumian-1 x T586, RIL (2015), Zhongmiansuo-12 x 8891, RIL (2007),7235 x TM-1, RIL (2007,HS-46 x MARCABUCAG8US-1-88, F2:3 (1998), Baimian 1 x TM-1, F2:3 (2014), TM-1 x 3-79, RIL (2014),TM-1 x (TX-0256 + TX-1046), CSIL (2016), CCRI-35 x Yumian-1, RIL (2017), TM-1 x (TX-0256 + TX-1046), CSIL (2016), 3-79 x TM-1, F2 (2015 SNP) CCRI-35 x Yumian-1, RIL (2017)
56	Lint visible foreign matter	Quality trait	qLFVM	4		TM-1 x 3-79, RIL (2014)
57	Lint yield	Yield trait	qLY	91	0.0471–18.32	Handan-208 x Pima-90, F2:3 (2007), 7235 x TM-1, RIL (2007), MD-5678ne x Prema, F2:3 (2000), Zhongmiansuo-12 x 8891, RIL (2007), DH962 X Ji Mian 5, F2 (2015), TM-1 x 3-79, RIL (2014), DH962 x Ji Mian 5, RIL (2015),SG 747 x Giza 75, BIL (2013) 3-79 x TM-1, F2 (2015 SNP)
58	Log (locule number)/ Log (boll number)	Yield trait	qLB	2		CAMD-E x Seaberry, F2 (1998)
59	Magnesium concentration	Biochemical trait	qMGCONC	8	0.09–0.28	Guazuncho-2 x VH8-4602, RIL (2009)
60	Micronaire	Quality trait	qMIC	794	0.0156–35.9	Handan-208 x Pima-90, F2:3 (2007),TM-1 x 3-79, RIL (2006),7235 x TM-1, RIL (2007),MD-5678ne x Prema, F2:3 (2000),Yumian-1 x T586, RIL (2009),Yumian-1 x T586, RIL (2009) (YM-1 x CCRI-35) x (YM-1 x 7235), 4WC (2012) 7235 x TM-1, F2:3 (2005),HS427-10 x TM-1, F2:3 (2005),PD6992 x SM3, F2:3 (2005),Zhongmiansuo-12 x 8891, RIL (2007),7235 x TM-1, RIL (2007), HS-46 x MARCABUCAG8US-1-88, F2:3 (1998),Palmeri x K-101, F2, (2007),DH962 X Ji Mian 5, F2 (2015),IL138-A11-3 x TM-1, F2:3 (2014) TM-1 x (TX-0256 + TX-1046), CSIL (2016), CCRI-35 x Yumian-1, RIL (2017),GX1135 x VGX100-2, RIL (2016)
61	Osmotic potential	Physiological trait	qOP	24		Palmeri x K-101, F2 (2007), Sic'on x F-177, F2 (2004),

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
62	Petal length	Morphological trait	qPTLL	19		Pima S-7 x im, F2 (2007), SMA-4 x A1-97, F2 (2005)
63	Phosphorus concentration	Biochemical trait	qPCONC	5	0.09–0.17	Guazuncho-2 x VH8-4602, RIL (2009)
64	Plant height	Stature or vigor trait	qPH	86	0.0382–21.7701	HS-46 x MARCABUCAG8US-1-88, F2:3 (1998), CCRI-36 x G2005, RIL (2016), 3-79 x TM-1, F2 (2015 SNP),LMY28 x XLZ24, RIL (2018)
65	Pollen color	Morphological trait	qPOLLE-NCL	1		MCU-5 x Siokra 1-4, RIL (2015 SNP)
66	Potassium concentration	Biochemical trait	qKCONC	4	0.09–0.12	Guazuncho-2 x VH8-4602, RIL (2009)
67	Potassium to sodium concentration ratio	Biochemical trait	qKNACO-NCR	5	0.08–0.14	Guazuncho-2 x VH8-4602, RIL (2009)
68	Reflectance	Quality trait	qCRD	27	0.0016–7.3778	Zhongmiansuo-12 x 8891, RIL (2007), CCRI-35 x NH, F2:3 (2018), Guazuncho-2 x VH8-4602, BC1 (2005)
69	Reniform nematode resistance	Biotic stress trait	qNRFR	9		GB713 x Nem-X, F2 (2011)
70	Seed arginine content	Biochemical trait	qAAARG	2		TM-1 x Hai-7124, BC1S1 (2007)
71	Seed coat percent	Morphological trait	qSCP	5	39.55–69.27	
72	Seed cotton weight per boll	Yield trait	qBWT	334	0.0104–27.96	Palmeri x K-101, F2 (2007), Zhongmiansuo-12 x 8891, RIL (2007),7235 x TM-1, RIL (2007),DH962 X Ji Mian 5, F2 (2015),TM-1 x (TX-0256 + TX-1046), CSIL (2016),DH962 x Ji Mian 5, RIL (2015),CCRI-35 x Yumian-1, RIL (2017),HS-46 x MARCABUCAG8US-1-88, RIL (2016),3-79 x TM-1, F2 (2015 SNP), Sic'on x F-177, F2 (2004), LMY28 x XLZ24, RIL (2018), 901-001 x sGK156, F2 (2019)
73	Seed cotton yield	Yield trait	qSCY	61	0.0512–20.07	Handan-208 x Pima-90, F2:3 (2007),Zhongmiansuo-12 x 8891, RIL (2007),7235 x TM-1, RIL (2007),Palmeri x K-101, F2 (2007),TM-1 x 3-79, RIL (2014), SG 747 x Giza 75, BIL (2013),Sic'on x F-177, F2 (2004),Acala-44 x Pima S-7, F2 (2004)
74	Seed cysteine content	Biochemical trait	qAACYS	2		TM-1 x Hai-7124, BC1S1 (2007)
75	Seed fuzz color	Morphological trait	qSDFZCL	13		Sic'on x F-177, F2 (2004)
76	Seed glycine content	Biochemical trait	qAAGLY	2		TM-1 x Hai-7124, BC1S1 (2007)
77	Seed index	Yield trait	qSI	303	0.0434–28.2748	Handan-208 x Pima-90, F2:3 (2007), 7235 x TM-1, RIL (2007),MD-5678ne x Prema, F2:3 (2000),Zhongmiansuo-12 x 8891, RIL (2007),7235 x TM-1, RIL (2007),HS-46 x MARCABUCAG8US-1-88, F2:3 (1998), DH962 X Ji Mian 5, F2 (2015),CCRI-35 x Yumian-1, RIL (2017),TM-1 x (TX-0256 + TX-1046), CSIL (2016)
78	Seed isoleucine content	Biochemical trait	qAAILE	1		TM-1 x Hai-7124, BC1S1 (2007)
79	Seed kernel content	Morphological trait	qSKC	1		TM-1 x Hai-7124, BC1S1 (2007)

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
80	Seed leucine content	Biochemical trait	qAALEU	3		TM-1 x Hai-7124, BC1S1 (2007)
81	Seed linolenic acid content	Biochemical trait	qFALIN	8	5.4–15.1	Yumian-1 x T586, RIL (2015)
82	Seed number per boll	Yield trait	qSNPB	9	0.183–0.381	
83	Seed number per plant	Yield component trait	qSNPP	1		Acala-44 x Pima S-7, F2 (2004)
83	Seed oil content	Biochemical trait	qSOC	60	3.9–45.89	TM-1 x Hai-7124, BC1S1 (2007), Yumian-1 x T586, RIL (2015), GX1135 x GX100-2, RIL (2016),
84	Seed oleic acid content	Biochemical trait	qFAOLE	11	5.7–16.6	Yumian-1 x T586, RIL (2015)
85	Seed palmitic acid content	Biochemical trait	qFAPAL	13	4.2–14.8	Yumian-1 x T586, RIL (2015)
86	Seed phenylalanine content	Biochemical trait	qAAPHE	1		TM-1 x Hai-7124, BC1S1 (2007)
87	Seed protein content	Biochemical trait	qSPC	23	5.2–88.9	TM-1 x Hai-7124, BC1S1 (2007), Yumian-1 x T586, RIL (2015),
88	Seed serine content	Biochemical trait	qAASER	1		TM-1 x Hai-7124, BC1S1 (2007)
89	Seed stearic acid content	Biochemical trait	qFASTE	13	4.4–22.7	Yumian-1 x T586, RIL (2015)
90	Short fiber content	Quality trait	qSFC	93	0.0583–20.62	Zhongmiansuo-12 x 8891, RIL (2007), Palmeri x K-101, F2 (2007),TM-1 x 3-79, RIL (2014),3-79 x TM-1, F2 (2015 SNP),PD 94042 x (AD)4-08, F2 (2016),MGG-293-793, MAGIC (2018),Pima S-7 x n2, F2 (2007)
91	Short fiber content by weight	Quality trait	qSFCw	1		CAMD-E x Seaberry, F2 (1998)
92	Sodium concentration	Biochemical trait	qNACONC	4	0.11–0.17	Guazuncho-2 x VH8-4602, RIL (2009)
93	Standard fineness	Quality trait	qHS	13		Guazuncho-2 x VH8-4602, BC1 (2005)
94	Stem diameter	Stature or vigor trait	qSTEMD	3		TM-1 x 3-79, RIL (2014)
95	Stem height to node ratio	Stature or vigor trait	qRPHND	2		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
96	Stem node number	Stature or vigor trait	qSTEMNN	3		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
97	Stigma length	Morphological trait	qSTGL	12		Pima S-7 x im, F2 (2007), SMA-4 x A1-97, F2 (2005)
98	Stigma protrusion length	Morphological trait	qSTGPL	7		Pima S-7 x im, F2 (2007), SMA-4 x A1-97, F2 (2005)
99	Sulfur concentration	Biochemical trait	qSCONC	2	0.11–0.14	Guazuncho-2 x VH8-4602, RIL (2009)
100	Trichome density on leaf	Morphological trait	qDTL	22		Pima S-7 x Empire, F2 (2007), MCU-5 x Siokra 1-4, RIL (2015 SNP), MCU-5 x Siokra 1-4, RIL (2015 SNP),Palmeri x K-101, F2 (2007), Guazuncho-2 x VH8-4602, BC1 (2005)
101	Trichome density on leaf by number	Morphological trait	qDTLn	6		Guazuncho-2 x VH8-4602, BC1 (2005)
102	Trichome density on stem	Morphological trait	qDTS	3		Palmeri x K-101, F2 (2007), Guazuncho-2 x VH8-4602, BC1 (2005)
103	Upper half mean length	Quality trait	qUHML	18	0.052–0.09	CCRI-35 x Yumian-1, RIL (2017), Guazuncho-2 x VH8-4602, BC1 (2005),

Sl. No.	*Triat name*	*Trait category*	*Keyword for QTL*	*No. of reported QTLs*	*R^2 range*	*References*
104	Upper quartile length by weight	Quality trait	qUQLW	4		TM-1 x 3-79, RIL (2014)
105	Verticillium wilt disease incidence	Biotic stress trait	qVWDI	120	2.3–21.3	0-153 x sGK9708, RIL (2017),
106	Weight fineness	Quality trait	qWTFN	7		HS-46 x MARCABUCAG8US-1-88, F2:3 (1998)
107	Yellowness	Quality trait	qCB	46	0.0226–83.8311	Zhongmiansuo-12 x 8891, RIL (2007), CCRI-35 x NH, F2:3 (2018), Guazuncho-2 x VH8-4602, BC1 (2005)
108	Zinc concentration	Biochemical trait	qZNCONC	4	0.1–0.26	Guazuncho-2 x VH8-4602, RIL (2009)

prerequisite for the successful operation of natural genes in populations.

B. Evaluation and Selection of Parents:

To enhance the breeding program's chances of success, the complete germplasm collection is evaluated to find plants with a desirable combination of traits. Then, two or more types of plants are selected which possess all the needed traits amongst them. These selected plants are called the parent.

C. Cross Hybridization among the Selected Parents:

The process of making a cross between two genetically unlike parents to obtain offspring with desired superior traits and may be a prodigious and tedious task.

D. Selection and Testing of Superior recombinant:

It is an essential step in the breeding program's success; hence a thorough scientific examination of progeny is required. Plants that are superior to both parents emerge from the selection process.

E. New Cultivar Testing, Release, and Commercialization:

The newly selected genotypes must undergo evaluation for a different improved condition like ++yields, disease- resistance, drought–resistance, insect-resistance, high temperature –tolerant, insect/pest resistance, and other traits before releasing a good variety for cultivation in a given region. The Indian Council of Agricultural Research (ICAR), New Delhi, the evaluation is carried out in India. The material is compared to a superior locally available crop cultivar, a check or reference cultivar. Finally, the seeds of desirable plants are certified by the International/National seed corporation for marketing.

New Demands in Cotton Breeding

Global cotton production is presently facing many challenges with the rapid expansion in human population and the loss of arable land due to harsher climate conditions, urbanization, and many other issues. As a result, the demand for improving cotton yield is increasing. Thus, identifying the key limiting factors for increasing yield is critical to establishing the research initiatives for handling the challenges in cotton production. Biotic and abiotic stresses, global climate change, genotype by environmental interactions, limited germplasm resources, the negative association between yield and fiber quality, and most importantly, lacking tightly linked molecular DNA markers with high predictive trait value are the major limiting factors for increasing fiber yield and quality.

The primary purpose of cotton breeding is to develop genetically improved cultivars for yield and fiber quality by overcoming the limitations mentioned above. Thus, cotton researchers should aim to demonstrate the trends and the new approaches in cotton breeding towards the development of new cultivars, resolving/finding a new key on these above issues in understanding the genetic basis of attractive traits under differential environments, improving breeding efficiency, enhancing germplasm, and broadening genetic base in cotton germplasm.

GM Cotton *visa-vis* Marker-based Breeding Program

Cotton traits have been improved using gene transformation as well as marker-assisted breeding (MAS) techniques. Few cotton traits may be controlled by a single gene, whereas multiple genes influence complex fiber traits. Both traits, simple and complex, can be improved through MAS; however, multiple gene transformations of various traits in GM cotton are tedious and difficult. The horizontal gene transfer can be obtained in GM cotton for enhancing the level of trait choosing genes from other kingdoms. This horizontal transfer is not feasible in MAS but is limited to wild relatives. It is conceivable to reconstruct certain gene combinations over a relatively short number of generations using modern molecular breeding techniques. Additional genes close to the gene of interest are often transferred in MAS (linkage drag). This problem can be overcome if the gene can be genetically modified and delivered directly into a high-yield crop. There are significant advantages to combining MAS with GM. Combining multiple GM traits into superior cultivars would become a future challenge. This could be enhanced by the wise application of molecular markers in MAS to rise to a new era in cotton development.

Future Prospective of Cotton Improvement

A highly significant impact of cotton productivity resulted from meticulous efforts by conventional cotton breeders around the world. However, the advent of newer technology is yet to add its profuse advantages to cotton improvement. The prior knowledge of tightly linked molecular DNA markers with high predictive trait value bases molecular breeding assisted with novel genomic technologies. A simple mistake in the initial level of finding tightly linked molecular DNA markers with

high predictive traits may pay a heavy price in the end in the molecular breeding program. Thus, the data assets developed for finding tightly linked markers for traits can be cautiously used in cotton improvement programs. More research should focus on advanced technologies like gene editing, precise mapping, and genomic selection for sustainable cotton production.

Summary

Despite a well-targeted breeding program for enhancing cotton yield and fiber quality, there is a lack of data on the responses of available elite cotton cultivars from India. Molecular markers are a valuable tool used to develop genetic maps and as a steppingstone for map-based gene cloning. Several genetic maps of intra-specific and inter-specific cotton crosses for drought tolerance have been reported, and their utility in cotton genetic improvement has been discussed. (Reinisch *et al*., 1994; Smith, *et al*., 1999; Ulloa, *et al*., 2000). A molecular marker is an emerging technology in many crop breeding programs. It is an efficient molecular resource in the context of genetic and QTL mapping. These crucial genetic resources can demand a next-generation breeding program in many crops, including cotton. More molecular markers might meet the demand for next-generation breeding programs in various crops, notably cash crop cotton. These molecular marker resources are valuable in finding important traits as quantitative trait loci on specific chromosomes in various crops. Using a molecular marker to accelerate a breeding program is advantageous and time-consuming because no breeding program can be envisioned without one.

References

Abdurakhmonov, I. Y., Buriev, Z. T., Saha, S., Pepper, A. E., Musaev, J. A., Almatov, A.,... & Yu, J. Z. (2007). Microsatellite markers associated with lint percentage trait in cotton, *Gossypium hirsutum*. *Euphytica*, *156*(1-2), 141-156.

Abdalla AM, Reddy OUK, El-Zik KM and Pepper AE (2001) Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. *Theoretical and Applied Genetic, 102*:222-229.

Allinne, C., Mariac, C., Vigouroux, Y., Bezançon, G., Couturon, E., Moussa, D.,... & Robert, T. (2007). Role of seed flow on the pattern and dynamics of pearl millet (Pennisetum glaucum [L.] R. Br.) genetic diversity assessed by AFLP markers: a study in south-western Niger. *Genetica*, *133*(2), 167-178.

Altaf, M. K., Stewart, J. M., Wajahatullah, M. K., Zhang, J. F., & Cantrell, R. G. (1997). Molecular and morphological genetics of a trispecies F2 population of cotton. In Proceedings of the Beltwide *Cotton Conference, 1*, 448-452.

Beyene, Y., Botha, A. M., & Myburg, A. A. (2006). Genetic diversity in traditional Ethiopian highland maize accessions assessed by AFLP markers and morphological traits. *Biodiversity & Conservation*, *15*(8), 2655-2671.

Bolek, Y., El-Zik, K. M., Pepper, A. E., Bell, A. A., Magill, C. W., Thaxton, P. M., & Reddy, O. U. K. (2005). Mapping of verticillium wilt resistance genes in cotton. *Plant Science*, *168*(6), 1581-1590.

Boivin, K., Deu, M., Rami, J. F., Trouche, G., & Hamon, P. (1999). Towards a saturated sorghum map using RFLP and AFLP markers. *Theoretical and Applied Genetics*, *98*(2), 320-328.

Byers, R. L., Harker, D. B., Yourstone, S. M., Maughan, P. J., & Udall, J. A. (2012). Development and mapping of SNP assays in allotetraploid cotton. *Theoretical and Applied Genetics*, *124*(7), 1201-1214.

Dong, G. J., Liu, G. S., & Li, K. F. (2007). Studying genetic diversity in the core germplasm of confectionary sunflower (*Helianthus annuus* L.) in China based on AFLP and morphological analysis. *Russian Journal of Genetics*, *43*(6), 627-635.

Dongre, A. B., Raut, M. P., Paikrao, V. M. and S. S. Pande, (2012), Genetic purity testing of cotton F1 hybrid DHH-11 and parents revealed by molecular markers. *Int. Res. J. Biotechnol*., *3*(2), pp. 32–36, 2012

Fang, D. D., & Yu, J. Z. (2012). Addition of 455 microsatellite marker loci to the high-density *Gossypium hirsutum TM-1× G. barbadense 3–79 genetic map*. *J Cotton Sci*, *16*, 229-248.

Hashimoto, Z., Mori, N., Kawamura, M., Ishii, T., Yoshida, S., Ikegami, M.,... & Nakamura, C. (2004). Genetic diversity and phylogeny of Japanese sake-brewing rice as revealed by AFLP and nuclear and chloroplast SSR markers. *Theoretical and Applied Genetics*, *109*(8), 1586-1596.

He, D. H., Lin, Z. X., Zhang, X. L., Nie, Y. C., Guo, X. P., Zhang, Y. X., & Li, W. (2007). QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum× Gossypium barbadense*. *Euphytica*, *153*(1-2), 181-197.

Hinze LL, Hulse-Kemp AM, Wilson IW, Zhu QH, Llewellyn DJ, Taylor JM, Spriggs A, Fang DD, Ulloa M, Burke JJ, Giband M. (2017). Diversity analysis of cotton (Gossypium hirsutum L.) germplasm using the CottonSNP63K Array. *BMC Plant Biology, 17*(1):1-20.

Jena SN, Srivastava A, Singh UM, Roy S, Banerjee N, Rai KM, Singh SK, Kumar V, Chaudhary LB, Roy JK, Tuli R. (2011). Analysis of genetic diversity, population structure and linkage disequilibrium in elite cotton (*Gossypium* L.) germplasm in India. *Crop and Pasture Science, 62*(10):859-75.

Jiang, H., Liao, B., Ren, X., Lei, Y., Mace, E., Fu, T., & Crouch, J. H. (2007). Comparative assessment of genetic diversity of peanut (Arachis hypogaea L.) genotypes with various levels of resistance to bacterial wilt through SSR and AFLP analyses. *Journal of Genetics and Genomics*, *34*(6), 544-554.

Khan SA, Hussain D, Askari E, Stewart JM, Malik KA, Zafar Y (2000). Molecular phylogeny of Gossypium sp. by DNA fingerprinting. *Theor. Appl. Genet*., *101*: 931-938

Khan, M. A., Zhang, J., Stewart, J. M., & Cantrell, R. G. (1998). Integrated molecular map based on a trispecific F2 population of cotton. In *Beltwide Cotton Conferences (USA)*.

Lacape, J. M., Nguyen, T. B., Thibivilliers, S., Bojinov, B., Courtois, B., Cantrell, R. G.,... & Hau, B. (2003). A combined RFLP SSR AFLP map of tetraploid cotton based on a Gossypium hirsutum× Gossypium barbadense backcross population. *Genome*, *46*(4), 612-626.

Lin, Z. X., He, D., Zhang, X. L., Nie, Y., Guo, X., Feng, C., & Stewart, J. M. (2005). Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. *Plant breeding*, *124*(2), 180-187.

Liu J, Qu JT, Yang C, *et al*. (2015). Development of genome-wide insertion and deletion markers for maize, based on next-generation sequencing data. *BMC Genomics*; *16*: 601.

Lu, Z. X., Sosinski, B., Reighard, G. L., Baird, W. V., & Abbott, A. G. (1998). Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. *Genome*, *41*(2), 199-207.

Lukonge, E. P., Labuschagne, M. T., & Herselman, L. (2007). Combining ability for yield and fibre characteristics in Tanzanian cotton germplasm. *Euphytica*, *161*(3), 383-389.

Meredith, W. R. (1992, January). RFLP association with varietal origin and heterosis. In *Proc. Beltwide Cotton Prod. Res. Conf., Nashville, TN* (pp. 6-10).

Mishra KK, Fougat RS. (2013). Genetic relationship among different species of cotton as revealed by SSR markers for fiber quality traits. *International Journal of Pure & Applied Bioscience*. *1*:81-93.

Moghaddam, M. E., Trethowan, R. M., William, H. M., Rezai, A., Arzani, A., & Mirlohi, A. F. (2005). Assessment of genetic diversity in bread wheat genotypes for tolerance to drought using AFLPs and agronomic traits. *Euphytica*, *141*(1-2), 147-156.

Nguyen, T. B., Giband, M., Brottier, P., Risterucci, A. M., & Lacape, J. M. (2004). Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. *Theoretical and Applied Genetics*, *109*(1), 167-175.

Park, Y. H., Alabady, M. S., Ulloa, M., Sickler, B., Wilkins, T. A., Yu, J.,... & Cantrell, R. G. (2005). Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. *Molecular Genetics and Genomics*, *274*(4), 428-441.

Rai, K. M., Singh, S. K., Bhardwaj, A., Kumar, V., Lakhwani, D., Srivastava, A.,... & Sawant, S. V. (2013). Large-scale resource development in *Gossypium hirsutum* L. by 454 sequencing of genic-enriched libraries from six diverse genotypes. *Plant Biotechnology Journa*l, *11*(8), 953-963.

Rana, M. K., Singh, V. P. and Bhat, K. V., (2005). Assessment of genetic diversity in upland cotton (*Gossypium hirsutum* L.) breeding lines by using amplified fragment length polymorphism (AFLP) markers and morphological characteristics. *Genet.Resour. Crop Evol*., *52*: 989–997

Rajeev, S., Shreekanth S. P., Ranganath, H. M., Srivalli, P., Manjula, S. M. and Pranesh K. J., (2014), Genetic Diversity Analysis in Cotton (*Gossypium Hirsutum* L.) Based on Morphological Traits and Microsatellite Markers. *Int. J. Basic and Applied Biolog*., *1*(1), 19-22.

Reddy, A. S., Haisler, R. M., Yu, Z. H., & Kohel, R. J. (1997). AFLP mapping in cotton. In *Plant Genome Conf*.-V.

Reinisch, A. J., Dong, J. M., Brubaker, C. L., Stelly, D. M., Wendel, J. F., & Paterson, A. H. (1994). A detailed RFLP map of cotton, *Gossypium hirsutum x Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. *Genetics*, *138*(3), 829-847.

Smith, C. W., & Cothren, J. T. (Eds.). (1999). Cotton: origin, history, technology, and production. 4. John Wiley & Sons.

Shen, X., Guo, W., Zhu, X., Yuan, Y., John, Z. Y., Kohel, R. J., & Zhang, T. (2005). Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. *Molecular breeding*, *15*(2), 169-181.

Shen, X., Guo, W., Lu, Q., Zhu, X., Yuan, Y., & Zhang, T. (2007). Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. *Euphytica*, *155*(3), 371-380.

Ulloa, M., & Meredith Jr, W. R. (2000). Genetic linkage map and QTL analysis of agronomic and fibre quality traits in an intraspecific population. *Journal of Cotton Science*, *4*(3), 161-170.

Van Deynze, A., Stoffel, K., Lee, M., Wilkins, T. A., Kozik, A., Cantrell, R. G., & Stelly, D. M. (2009). Sampling nucleotide diversity in cotton. *BMC plant biology*, *9*(1), 125.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T. V. D., Hornes, M.,... & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, *23*(21), 4407-4414.

Wang, B., Guo, W., Zhu, X., Wu, Y., Huang, N., & Zhang, T. (2007). QTL mapping of yield and yield components for elite hybrid derived-RILs in upland cotton. *Journal of Genetics and Genomics*, *34*(1), 35-45.

Waugh, R., Bonar, N., Baird, E., Thomas, B., Graner, A., Hayes, P., & Powell, W. (1997). Homology of AFLP products in three mapping populations of barley. *Molecular and General Genetics* MGG, *255*(3), 311-321.

Yin, J., Guo, W., Yang, L., Liu, L., & Zhang, T. (2006). Physical mapping of the Rf 1 fertility-restoring gene to a 100 kb region in cotton. *Theoretical and Applied Genetics*, *112*(7), 1318-1325.

Ynturi, P., Jenkins, J. N., McCarty Jr, J. C., Gutierrez, O. A., & Saha, S. (2006). Association of root-knot nematode resistance genes with simple sequence repeat markers on two chromosomes in cotton. *Crop Science*, *46*(6), 2670-2674.

Yu, J., Park, Y. H., Lazo, G. R., & Kohel, R. J. (1998, January). Molecular mapping of the cotton genome: QTL analysis of fiber quality properties. In *Proc. Beltwide Cotton Conf* (p. 485).

Yuxiang, Wu, Sun Yuqiang, Chen Chongqian, and Zhu Shuijin. (2010). Analysis of genetic relationship among cotton species Gossypium spp. by RAPD marker. 909-913.

Zhang ZS, Xiao YH, Luo M, Li XB, Luo XY, Hou L, Li DM, Pei Y (2005). Construction of a genetic linkage map and QTL analysis of fibre related traits in upland cotton. Euphytica, *144*: 91-99.

Zhang, J., Guo, W. Z., & Zhang, T. (2002). Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L.× Gossypium barbadense L.) with a haploid population. *Theoretical and Applied Genetics*, *105*(8), 1166-1174.

Zhu, Q. H., Spriggs, A., Taylor, J. M., Llewellyn, D., & Wilson, I. (2014). Transcriptome and complexity-reduced, DNA-based identification of intraspecies single-nucleotide polymorphisms in the polyploid *Gossypium hirsutum* L. *G3:Genes, Genomes, Genetics, 4*(10), 1893-1905.