We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Association Mapping for Improving Fiber Quality in Upland Cottons

Khezir Hayat, Adem Bardak, Mehboob-ur-Rahman, Hafiz Muhammad Imran, Furqan Ahmad, Donay Parlak, Muhammad Azam, Muhammad Usmaan, Muhammad Adnan, Sidra Anjum and Rao Sohail Ahmad Khan

Abstract

Improved fiber yield is considered a constant goal of upland cotton (*Gossypium hirsutum*) breeding worldwide, but the understanding of the genetic basis controlling yield-related traits remains limited. Dissecting the genetic architecture of complex traits is an ongoing challenge for geneticists. Two complementary approaches for genetic mapping, linkage mapping and association mapping have led to successful dissection of complex traits in many crop species. Both of these methods detect quantitative trait loci (QTL) by identifying marker–trait associations, and the only fundamental difference between them is that between mapping populations, which directly determine mapping resolution and power. Nowadays, the availability of genomic tools and resources is leading to a new revolution of plant breeding, as they facilitate the study of the genotype and its relationship with the phenotype, in particular for complex traits. Next Generation Sequencing (NGS) technologies are allowing the mass sequencing of genomes and transcriptomes, which is producing a vast array of genomic information with the development of high-throughput genotyping, phenotyping will be a major challenge for genetic mapping studies. We believe that high-quality phenotyping and appropriate experimental design coupled with new statistical models will accelerate progress in dissecting the genetic architecture of complex traits.

Keywords: fiber quality, MAS, GBS, SNPs, association mapping

1. Introduction

Cotton is a crop of immense importance as being a dominant source of fiber and oil from cottonseed all over the world [1]. The improvement of cotton fiber quality has become more important because of changes in spinning technology and ever-increasing demands of fiber. Cotton is grown in more than 80 countries, and contributes to the world economy as a raw material for textile industry [2].

Gossypium" genus is made up of about 52 species of which 47 are diploid and 7 are as allotetraploids [3–7]. Of all the species of the genus, two most common

IntechOpen

diploids are *G. arboreum* L., *G. herbaceum* L., while *G. hirsutum* L., and *G. barbadense* L. are considered as the most commercially valuable tetraploids. *G. hirsutum*, is characterized by high yield, moderate fiber quality and wide adaptability contributes for 95% of overall cotton production [8]; while *G. barbadense* (Pima, and Egyptian) increases superior fiber quality [9, 10].

Efforts for broadening the genetic base of *Gossypium* genus have not generated successful outcomes due to the complex and large genetic architecture of its genome. Moreover, owing to its developmental barriers, genetic studies have not yet been able to produce the required traits in cotton [11]. Association among markers and characters can be used for fastening the breeding program. The hereditary variation present among the gene pool land races can be exploited by applying the mapping based on linkage disequilibrium. It will speed up the cotton breeding through identification of markers among trait of interest and ensure molecular breeding. Single reproducibility of genetic marker which govern a specific appearance on sequence of nucleotides can be analyzed with genome wide association [12, 13]. Association mapping relies upon the magnitude of different pair of genes for population analysis. Moreover, this mapping shows powerful connection between required character and a genetic marker while nonrandom combination between two quantitative trait loci or markers manifests linkage disequilibrium [8]. The valuable information about the origin of an individual is determined with the degree and the size of the population [13, 14]. Many loci relating to polygenic characters have been determined via genetic maps and linkage disequilibrium (LD) was measured in humans through diverse analysis methods [15, 16]. Population based polygenic characters mapping for desired traits became a widely used technique thanks to the innovations in omics and availability of advanced bioinformatic tools for analyzing genetic variations [17]. The ultimate benefits of this technique includes the ability to work with a large number of loci, producibilty of highly saturated maps, its speed and its low cost [18].

2. Fiber quality

Single cell elongation of ovule in cottonseed outer layer forms a natural fiber known as “trichome” which contains about 89–100% cellulose. [19–22]. As little as, 30% of lint primordia have the ability to be differentiated as mature fibers forming about 20,000 of it within a single ovule [23, 24]. The ideal cotton fiber should be white like frozen vapor, durable like iron, attractive like silk and stretched as a wool [25]. Nonetheless it is hard to include all these qualities within a breeding program for cotton production, but efforts have been made to obtain the most desired ones. Fiber quality is an array of quantitative traits (length, fineness, strength, uniformity and elongation) that enhance yarn value during spinning [26–28]. Fiber quality is a difficult association of physiology and genetic make-up of plant within a growing season of cotton [29, 30].

Fiber quality enhancement through genetics is the ultimate objective of breeding strategy in cotton. Cotton scientists have been involved in fiber quality improvement for a long time due to the increase in demand for multiple products from cotton. The critical goals of all cotton related techniques are fiber yield and quality, and the precise parameters which contribute its economic value on global level. Spinning automation renders fiber improvement according to interests of textile sector, as a result fiber quality measurements for breeders are considered. As an instance, prevailing spinning automation highly signify strength instead of fiber length and fineness [31]. Moreover, fiber quality improvement is a demanding task as it is determined after harvesting of crop.

The main goal of all genetic improvement is to increase yield. The intensity of improvement for lint production has deteriorated since the 1980s [32–34]. Nonetheless, genetic diversity has increased at the start of 21st century [35, 36].

3. Marker assisted selection

Due to the inverse relationship between seedcotton yield and fiber quality, and the complicated involvement of multiple genes in traits demand breeders to evolve varieties through more useful methods. In the past textile industry flourished principally via selection of new recombinants among germplasm entries with traditional breeding approaches [37, 38]. Elite grown cotton genotypes have narrow genetic base, therefore it has been thought that germplasm should be used for improvement of traits. Some of popular characters such as disease and insect resistance have been enhanced by introgression [39]. The advent of DNA markers paved the way for plant breeders to fasten breeding process through fast, authentic and substitutive techniques instead of the traditional methods for the selection to develop both agronomic and economic characters of plants [40].

Molecular marker is a specific DNA portion with a known position on the chromosome [41], or a gene whose phenotypic expression is frequently easily distinguished and used to detect an individual [42, 43]. DNA markers are having the property of polymorphism which can be used for the differentiation of homozygotes and heterozygotes [44]. Marker assisted selection has a great amount of advantages over conventional breeding, reviewed by many researchers [45–47]. Plant breeders utilize DNA markers for selection of desirable traits on molecular basis in spite of observing them phenotypically [48], furnishing the basis for using the molecular assisted selection [49–51]. Molecular markers are desired for improving traits in many essential crops; rice [52], wheat [53], maize [54, 55] and barley [56, 57]. Cotton is an important cash crop at global level and marker assisted selection has not got desired goals because of compatibility barriers through historic domestication and insufficient polymorphism [58–60].

Molecular characterization is the way to transfer required traits into modern genotypes [45, 61–64]. Quantitative trait loci (QTLs) allow gene pyramiding for yield and fiber quality through evolution of linkage maps. Association mapping using linkage disequilibrium on genome wide level is the most valuable strategy among scientists for searching QTLs in crop sciences. The association among trait of interest and germplasm entries is observed using population construction information and linkage disequilibrium (LD) with association mapping [65]. LD mapping is highly popular thanks to the sophistication of mathematical methods and accessibility of large number of DNA markers.

The traits controlled by multiple genes such as fiber quality can be studied more precisely with linkage maps after the availability of new genomic data of *Gossypium* spp. like *Gossypium raimondii* Ulbrich [66, 67], *Gossypium arboreum* L. [68] and *Gossypium hirsutum* L. [69, 70]. [71] revealed that tetraploid species derived from crossing of two diploid species *Gossypium arboreum* L. (A genome) and *Gossypium raimondii* Ulbrich (D genome) about 1–2 million years ago. Moreover, it may pave the way for fiber improvements as higher number of QTLs assigned to the Dt sub-genome compared to At sub-genome in hawian cotton [72–74].

Many researchers have observed QTLs for seedcotton yield and its components [9, 70, 75–79]. But, mostly filial generations were used for QTLs. Quantitative trait loci are highly effected by low heritability and more experimental error which are high in such plant materials, hence it is need of the day that a useful way should be

employed for the development of stable populations for overcoming these obstacles. The accuracy of QTL determination relies upon allelic frequency among QTL of the desired character and related marker [80]. Molecular breeding methods designed with the information obtained through quantitative trait loci analysis in association mapping creates valuable genetic variation from stable populations [81].

4. Association mapping of fiber traits using genotyping by sequencing (GBS)

Molecular markers are highly favored for linkage map development because they are polymorphic, easily transferred to next generation with Mendelian ratio and do not show epistasis. Molecular breeding with highly saturated maps having QTLs connected with economic traits through impactful genetic markers provides a good source for cotton improvement [64]. Genomic analysis in many crop species including cotton has been done using populations derived from hybridization of only two ancestors; which is major drawback for omics information. Therefore, there has been hindrance in applying QTL information gained from such populations to accomplishing breeding objectives, as, in these populations, the genetic aspects are the same owing to the share of genetically similar backgrounds.

The foundation of association mapping is on hypothesis about occurrence of markers as a panel in which the alleles are found almost adjacent to the required traits with co-segregation and thought to be in linkage disequilibrium. Germplasm entries are used for determining QTLs of interest using genome wide association mapping [82]. There are many agents including type of copulation, gene flow frequency and population structure can affect such mapping approach [18]. Association mapping allows to overcome drawbacks found in bi-parental mapping from traditional methods which include using populations which are found as well-established genotypes, detects only the required gene and identify high polymorphism [83–85]. This methodology also urges to use knowledge based on linkage disequilibrium instead of linkage mapping.

Marker assisted breeding involves recent approaches of genomics combined with traditional breeding procedures for improving traits in crop sciences. For this reproducibility is essential among genetic markers. Morphological characters grading and genotyping with molecular markers is accomplished [86]. Molecular markers are very effective for identifying and overcoming problems for transfer of traits from other species such as segregation distortion [87]. Genetic markers are effective for determining genetic variation in Gossypium gene pool. [88] classified DNA markers into groups: 1) non-hybridization based; which include Amplified Fragment Length Polymorphism (AFLP), Simple Sequence Repeats (SSR), Sequence Repeat Amplified polymorphism (SRAP), Inter-Simple Sequence Repeats (ISSR), Expressed Sequence Tag (EST-SSR), Single Nucleotide Polymorphism (SNPs) etc. Numerous linkage maps have been developed in allotetraploid cotton employing diverse mapping populations and different DNA markers techniques [76, 89–94]. Numerous SSRs and SNPs have been evolved in cotton [95–99]. Saturated genetic maps development through loci information of SSR and SNPs in cotton paves the way for ascertaining quantitative traits related to breeder objectives [100–104] Nonetheless, association analysis and very fine mapping is not possible owing to less information from these maps. It is need of the day that highly saturated mapping should be devised in cotton for overcoming the sequencing drawbacks and fastening the variety development.

Availability of microsatellites (SSR) and single-nucleotide polymorphisms (SNPs) have fastened genome mapping owing to their wider applicability in diverse

populations derived from discrete genetic backgrounds [93, 95, 99, 105–107]. Thanks to advances in genotyping and SNPs calling tools; broadening of genetic base is being explored excessively in plants owing to availability of valuable loci information [108–114].

Single nucleotide polymorphisms are distinct points of nucleotides on chromosomes between two genotypes differentiated by a single base [64]. [115] speculated that each SNP is found after 100-300 bp in any genome while revealed that such genetic markers are highest in occurrence than any other marker and manifest higher degree compared to microsatellites. SNPs can be formed rapidly with economical cost owing to availability of high-throughput tools for genotyping [116]. Assessment of gene expression [117, 118], genome wide association [68, 119] and SNPs detection has been carried among the individuals having different sizes of genomes and also polyploid species having limited genetic variation like cotton [10, 120] and wheat [121] through low-cost high-throughput genotyping tools. SNPs have been explored and genotyped among different species via diverse ways [10, 120–122].

Genotyping-by-sequencing (GBS) is powerful and easy approach which paves the way for the discovery of numerous SNPs concurrently among large number of genotypes [123]. Restriction enzymes with methyl sensitivity are used to mark the flanking restriction sites in the genome for the development of reduced representation of the genome via GBS [121, 122]. GBS method is much easier, requires lower amount of DNA and library preparation is achieved in just two steps on plates, circumvents DNA fragment analysis preceded by PCR amplification of pooled library in contrast to reduced representation libraries (RRL) and restriction site associated DNA (RAD) [122]. The discovery and verification of reproducibility is not required in this procedure and can be applied in any species having polymorphism or mapping population with diverse size [124]. A number of SNPs has been discovered in many species using GBS like maize [122], wheat, barley [121], sorghum [125], rice [126], soybean [127], oat [128] and cotton [10, 79, 129, 130].

Association mapping furnishes saturated map of desired trait in contrast to pair of genes harboring a required character [131]. Therefore, verification of QTLs is compulsory for mapping. Association mapping is the way to examine genetic variation of required characters; integrates the variation of the desired characters through reproducibility of the alleles and genetic markers are selected connected to economic traits using linkage disequilibrium extent [132]. Moreover, LD elaborates the ancestral pattern through information among populations and ecology [133, 134].

LD based association mapping has been applied by using different strategies for determining genetic diversity contributing source pattern and design of population [135, 136]. Grouping of population individuals with combined genetic distance among the entries established via LD [137–139]. LD extent among natural population is not contributed by linked loci but non-homologous chromosomes are also involved, accountable to selection, behavior of population and hybridization. Owing to which immense care should be considered for analyzing such relations. Reproducibility in a sequence controlling a specific character is the property of this mapping [140]. Moreover, considerable concern is prevailed among association studies and linkage mapping relating to depth and precision of QTLs, the magnitude of knowledge and evaluating procedures [132].

In spite of the fact, statistical analysis is not appropriate with LD derived tools. Natural population partitioned into distinct categories with model-based procedures [141]. Bayesian modeling is used widely for assessing the probability of a genotype related to a specific population category through allele repetition. With this technique the genotypes are allotted to particular population which can be

interspersed into statistical methods for association mapping with population organization. The population framework is analyzed by using STRUCTURE software [135] which has been used for association studies in many plants. Various studies have been conducted in cotton for different aspects in cotton through association mapping like seedcotton yield and components [142–144], salt tolerance [145], architecture of plant, earliness [146] and protein and oil contents [147] and fiber quality [8, 60, 132, 148–150].

In-contrast to genetic mapping in populations developed from hybridization of parents using conventional ways are not saturated, labor intensive, always in danger, high investment for development and more work after evaluating numerous genotypes of gene pool [84]. Nonetheless, association mapping use LD and overcomes the requirement of bi-parental populations by utilizing the extent of genetic variation present within the available stable populations like cultivars, accessions developed with the time and maintained as gene pool. Association mapping on whole genome has been studied in Arabidopsis [151], rice [152] for observing loci connected to economical characters. Association studies allow the development of highly saturated maps via determination of QTLs related to economic characters at whole genome level in permanent mapping populations.

Abdurakhmonov et al. [60] used association analysis for observing association among fiber traits in cotton among germplasm entries for utilizing the genetic variation in marker-based breeding. Linkage disequilibrium based association mapping determined in the germplasm having diverse genotypes from all over the world. 95 SSR were screened among all germplasm entries for ascertaining QTLs at whole genome level associated with fiber properties. They found about 11–12% LD among all SSRs. They also observed significant population orientation among all entries. They employed mixed linear model and general linear model using kinship and population structure and as a whole determined 6 & 13% pair of primers related to fiber quality. They concluded that the markers selected in this study can be used for refinement of fiber using hidden sources of genetic variability.

Genetic variation, population behavior and LD based association analysis for fiber conducted in germplasm under two different climatic zones [85]. The upland gene pool containing 335 elite entries screened with 202 SSRs. Mean of LD prolonged to 25 cM at whole genome level among all genotypes at 0.01 probability. They found that LD dropped to about 5 cM at ($r2 > 0.2$) showing potential for association among genotypes for yield contributing characters. They performed mixed linear model and population analysis for observing association contributing to permutation significance and population pattern. As a whole developed many common markers for fiber traits among genotypes in both locations. They revealed that mixed linear model associations ranged from 7 to 43% having strong to very strong relation to fiber properties as confirmed by Bayes factor which will be a very effective source for association analysis of yield improvement in marker based breeding techniques.

Wang et al. [153] found association among yield and fiber characters in using mixed linear model in pima cotton germplasm entries. They observed 72 loci, out of which 46 were connected to fiber while 26 related to cotton. They concluded that marker-associations among fiber characters are of vital value for enhancing quality.

Fang et al. [154] used multi-parents population for observing association among yield and fiber quality traits. They revealed that common and new QTLs deducted in this study can be used for overcoming problems in fiber quality enhancement. They screened 1582 polymorphic microsatellites among 275 RILs in first set developed from diverse parents for screening QTLs connected to fiber. 131 QTLs found for fiber quality sharing characters via association analysis with TASSEL while same

QTLs verified in second set of 275 RILs with 270 SSR. The distinction showed that 54 new QTLs and 77 QTLs are in accordance to previous studies.

Genetic map constructed using RIL developed from transference of superior fiber quality from *G. barbadense* (TM-1) to *G. hirsutum* cv. NM24016 and relationship determined among yield components and fiber. 429 SSR and 412 GBS-based single nucleotides were involved in the development of map which spanned to about half length of upland cotton genome [10]. They revealed that all makers are distributed randomly among all loci of the genome. The yield components and fiber characters showed extreme phenotypic expression under multiple locations. They found 28 QTLs which are useful from breeding perspectives for agronomic and fiber properties.

Cai et al. [8] used 99 upland cotton genotypes to ascertain the association for fiber traits. The relationship among fiber components determined with 97 polymorphic microsatellites. The genomic regions associated with fiber were 107 including 70 in 2 or more than 2 zones and 37 found in just one. It was revealed that most of the associations were reliable as verified from earlier findings for fiber quality. They also observed genomic regions related with 2 or more characters and assumed that such regions derived from the genotypes which are having minor allele frequency less than five, from local sources or acclimatized in china. They concluded that fiber traits can be renovated by using such loci from diverse resources.

Islam et al. [123] carried GBS for observing SNPs which can be used for improving economic traits in cotton gene pool. RILs and 11 contrasting parents were used in the study with two separate methods were applied for determining SNPs with variant allele frequency of >0.1. SNPs quality control performed and calling done with available G. raimondii Ulbrich genome. As a whole 1071 and 1223 SNPs observed among At and Dt genomes respective. Moreover these SNPs were found in coding region usually in higher frequency. GBS was conducted in germplasm consisting of 154 accessions for the verification of 111 of total SNPs and the SNPs verified in all parents and none of the genotype was found with same SNP. They revealed that SNPs can be determined in *G. hirsutum* with ease and genetic improvement can be done after getting true SNPs.

Association among fiber traits conducted in germplasm collection of Hawaiian cotton consisting of 503 genotypes [132]. They used 494 microsatellites at whole genome and as a whole 179 replicable SSRs were screened among genotypes under diverse climatic conditions. Population pattern and LD used for observing association among various fiber traits with mixed linear model via TASSEL program. The QTLs were selected among markers and phenological characters with association values. 426 alleles were evolved and germplasm was differentiated into seven subgroups upon the basis of hybridization, climate and topographical pattern. 216 polymorphic loci were associated with fiber contributing characters having mean of 2.7% and showed phenotypic variation from 0.58–5.12%. LD decreased significantly to 0-5 cM and observed 13 QTLs which are same to earlier findings and 3 connected to similar character while 7 QTLs were corresponded to fiber formation. They concluded that novel alleles identified based association mapping based LD for fiber quality can be applied in breeding cultivars for tagging genes of interest.

GBS carried in a population evolved using various parents for overcoming the inverse relation among yield and fiber traits [155]. They assumed that GBS will serve as a valuable source for the development of high saturated map with the development of large frequency of SNPs. Association analysis via mixed linear model in TASSEL observed among fiber traits in four separate climates with 5071 SNPs developed from GBS and 223 SSRs from 547 RILs. One QTL cluster related to fiber traits including length, short fiber content, strength and uniformity found and verified on locus A07. They also studied the ultimate genes connected to fiber

traits and revealed that SNP (CFBid0004) formed from deletion of 10 bp GhRBB1_A07 is directly associated with fiber traits among RIL and 104 approved American varieties. Moreover, GhRBB1_A07 can be used in MAS for the improvement of fiber traits among germplasm entries.

Sun et al. [150] studied the genetic architecture of major fiber traits in cotton germplasm using association mapping under different climatic zones. The mixed linear model association analysis showed that fiber length, strength and uniformity had 16, 10 and 7 SNPs respectively while G. raimondii 7th chromosome had two main genomic locations and fiber length contributing four genes were also observed. Moreover population structure showed that populations from low peaks were having less genetic variation among accessions compared to high peaks. The valuable allelic frequency was more in genotypes from less elevation in-contrast to high. They concluded that the desired allelic number among genotypes can be used for enhancement of fiber.

Association was observed for plant ideotype, heat tolerance, yield contributing traits and fiber quality among germplasm collection under different climatic conditions for consecutive three years at whole genome [156]. The genetic stock associations were observed using SNPs. Fiber characters were found to be low to highly heritable as value ranged from 0.26–0.89 for boradsense heritability as compared to yield components having 0.14–0.43. Phylogenetic analysis showed that the genotypes were developed from diverse parents having multiple characters from breeding perspectives. They pointed that less number of informative markers can be used for association mapping studies as LD value found upto 5Mbp which decreased to 2Mbp at $r2 \geq 0.2$. 17 significant SNPs connected fiber length while 50 SNPs for fineness were observed using mixed linear model. The results revealed that associations among most of the characters at whole genome were non-significant as numerous SNPs impact on phenotype was found lower than 5% and assumed this to be due to low reproducibility of markers among cotton or SNP Chip less coverage in the germplasm.

Sun et al. [150] used association analysis in germplasm containing wide variation among genotypes at multiple locations for fiber quality traits. Illumnia SNP array was used for genome-wide study for quality analysis. They found 10,511 SNPs which were distributed over all loci and 46 SNPs associated with fiber quality with significance. They observed two QTLs for strength and length on At07 and Dt11.

5. Conclusion

Fiber quality enhancement through genetics is the ultimate objective of breeding strategy in cotton. Cotton scientists have been involved in fiber quality improvement for a long time due to the increase in demand for multiple products from cotton. Furthermore, conventional ways would be tiresome and stagnant. Hence, the modern plant improvement methods should be integrated. Molecular characterization is the way to transfer required traits into modern genotypes. Genotyping-by-sequencing (GBS) is powerful and easy approach which paves the way for the discovery of numerous SNPs concurrently among large number of genotypes. Quantitative trait loci (QTLs) allow gene pyramiding for yield and fiber quality through evolution of linkage maps. Molecular breeding with highly saturated maps having QTLs connected with economic traits through impactful genetic markers provides a good source for cotton improvement. Association mapping using linkage disequilibrium on genome wide level is the most valuable strategy among scientists for searching QTLs in crop sciences. It is need of the day that highly saturated mapping should be devised in cotton for overcoming the sequencing drawbacks and fastening the variety development.

IntechOpen

Author details

Khezir Hayat[1*], Adem Bardak[2], Mehboob-ur-Rahman[3], Hafiz Muhammad Imran[1], Furqan Ahmad[4], Donay Parlak[2], Muhammad Azam[1], Muhammad Usmaan[1], Muhammad Adnan[5], Sidra Anjum[6] and Rao Sohail Ahmad Khan[7]

1 Central Cotton Research Institute, Multan, Pakistan

2 Department of Agricultural Biotechnology, Kahramanmaras Sutcu Imam University, Kahramanmaras, Turkey

3 Plant and Molecular Breeding Labs, NIBGE, Faisalabad, Pakistan

4 Institute of Plant Breeding and Biotechnology, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan

5 Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan

6 Bahauddin Zakariya University, Multan, Pakistan

7 Center of Agriculture and Biochemistry Faculty of Agriculture, University of Agriculture Faisalabad, Faisalabad, Pakistan

*Address all correspondence to: khezirso@gmail.com

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Bardak, A., Bolek, Y. 2012. Genetic diversity of diploid and tetraploid cottons determined by SSR and ISSR markers. Turkish J. of Field Crops. 17(2): 139-144.

[2] Tan, Z., Fang, X., Tang, S., Zhang, J., Liu, D., Teng, Z., et al. (2014). Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica* 203, 615-628. doi: 10.1007/s10681-014-1288-9.

[3] Fryxell, P.A. 1979. The Natural History of the Cotton Tribe. Texas A&M University Press, College Station, TX, USA

[4] Fryxell, P.A. 1992. A revised taxonomic interpretation of *Gossypium* L. (*Malvaceae*). *Rheeda*, 2(2): 108-165.

[5] Stewart J. McD. 1995. Potential for crop improvement with exotic germplasms and genetic engineering. In challenging the Future, World Cotton Research Conference, Brisbance Australia 14-17 February 1994. Constable CA and Forester NW (Eds.) CSIRO, Melbourne. pp.313-332.

[6] Grover, C.E., Gallagher, J.P., Jareczek, J.J., Page, J.T., Judall, J.A., Wendel, J.F. 2015. Re-evaluating the phylogeny of allopolyploid *Gossypium* L. *Mol. Phylognet. Evol*., 92: 45-52.

[7] Gallagher, J.P., Gover, C.E., Rex, K., Moran, M., Wendel, J.F. 2017. A new species of cotton from Wake Atoll. *Gossypium stephensii* (Malvaceae). *Systematic Botany*, 42:115-123.

[8] Cai, C., Ye, W., Zhang, T., Guo, W. 2014. Association analysis of fiber quality traits and exploration of elite alleles in upland cotton cultivars/ accessions (*G. hirsutum* L.). *J. of Integ. Plant Bio*., 56: 51-62.

[9] Ulloa, M., Saha, S., Jenkins, J.N., Meredith, W.R.Jr., McCarty, J.C. Jr., Stelly, D.M. 2005. Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (*Gossypium hirsutum* L.) joinmap. J. Here., 96(2): 132-144.

[10] Gore, M.A., Fang, D.D., Poland, J.A., Zhang, J., Percy, R.G., Cantrell, R.G., Thyseen, G., Lipka, A.E. 2014. Linkage map construction and quantitative trait analysis of agronomic and quality traits in cotton, *The Plant Genome*, 7(1): 1-10.

[11] Rahman, M., Ullah, I., Ashraf, M., Stewart JM, Zafar Y. 2008. Genotypic variation for drought tolerance in cotton. *Agron. Sustain Dev*., 28:439-447.

[12] Cerda, S., Cloutier, B.J.; S.; Quian, R.; Gajardo, H.A.; Olivos, M.; You, F.M. 2018. Genome-Wide Association Analysis of Mucilage and Hull Content in Flax (*Linum usitatissimum* L.) Seeds. Int. J. Mol. Sci. 19, 2870.

[13] Nordborg, M., Tavare, S. 2002. Linkage disequilibrium: what history had to tell us. *Trend Genet*., 18: 83-90.

[14] Slatkin, M. 2008. Linkage disequilibrium-understanding the evolution past and mapping the medical future. *Nat. Rev. Genet*. 9: 477-485.

[15] Weiss, K.M., Clark, A.G. 2002. Linkage disequilibrium and mapping of human traits. *Trend Genet*., 18: 19-24.

[16] Kruglyak, L. 2008. The road to genome-wide association studies. *Nat. Rev. Genet*., 9: 314-318.

[17] Zhu, C., Gore, M., Buckler, E., Yu, J. 2008. Status and prospects of association mapping in plants. *The Plant Genome*, 1(1): 5-20.

[18] Flint-Garcia, S.A., Thornsberry, J.M., Buckler, E.S. 2003. Structure of

linkage disequilibrium in plants. *Annual Review of Plant Bio*., 54: 357-374.

[19] Basara, A.S., Malik, C.P. 1984. Development of cotton fiber. *Inter Rev. Cyto*., 65-113.

[20] Ryser, U. 1985. Cell wall biosynthesis in differentiating cotton fibers. *Eur. J. Cell Bio*. 39: 236-256.

[21] Delmer, P.D., Amor, Y. 1995. Cellulose biosynthesis. *The Plant Cell*, 7: 987-1000.

[22] Haigler, C.H., Zhang, D., Wilkerson, C.G. 2005. Biotechnological improvement of cotton fiber. *Phsioyl. Plant*., 124: 285-294.

[23] Berlin, J.D. 1986. The outer epidermis of the cottonseed. In: Mauney, I.R., Srewart I.M. (Ed.), *Cotton physiology*. The cotton foundation reference book series. Memphis, TN: The Cotton Foundation, : 375-414.

[24] Tiwari, S.C., Wilkins, T.A. 1995. Cotton (*Gossypium hirsutum* L.) seed trichomes expand via diffuse growing mechanism. *Canadian Journal of Botany*, 73: 746-757.

[25] Bradow, J.W., Davidonis, H.D. 2000. Quantification of fiber quality and the cotton production-processing interface: A physiologist perspective. *The J. of Cotton sci*., 4: 34-64.

[26] Dutt, Y., Wang, X.D., Zhu, Y.G., Li, Y.Y. 2004. Breeding for high yield and fiber quality in coloured cotton. *Plant Breeding*, 123:145-151.

[27] Ali, M.A., Khan, I.A., Nawab, N.N. 2009. Estimation of genetic divergence and linkage for fiber quality traits in cotton. *J. Agri. Res*., 47(3): 229-236.

[28] Shen, X., Cow, Z., Singh, R., Lubbers, E.L., Smith, C.W., Paterson, A.H., Chew, P.W. 2011. Efficacy QFL-chr1 a quantitative trait locus for fiber length on cotton (Gossypium sp.). *Crop Sci*., 51:2005-2010.

[29] Rahman, H., Murtaza, N., Shah, M.K.N. 2007. Study of cotton fiber traits inheritance under different temperature regimes. *J. Agron. Crop Sci*. 193:45-54.

[30] Ali, M.A., Khan, I.A., Awan, S.I., Ali, S., Niaz, S. 2008. Genetics of fiber quality traits in cotton (*Gossypium hirsutum*). *Aus. J. of Crop Sci*., 2: 10-17.

[31] Shen, X.L., Guo, W.Z., Zhu, X.F., Yuan, Y.L., Yu, J.Z., Kohel, R.J., Zhang, T.Z. 2005. Molecular mapping of QTL for fiber qualities in three diverse lines in upland cotton using SSR markers. *Mol. Breed*. 15:169-181.

[32] Meredith, W.R.Jr., Heitholt, J.J., Pettigrew, W.T., Rayburn, S.T. 1997. Comparsion of obsolete and modern cultivars at two nitrogen levels. *Crop Sci*., 37: 1453-1457.

[33] Meredith, W.R.Jr. 2002. Factors that contribute to lack of genetic progress. Proc. Beltwide Cotton Conf. Atlanta, GA 8-13 Jan 2002. National Cotton Coun. Am, Memphis, TN. Pp: 528-532.

[34] Bayles, M.B., Verhalen, L.M., Johnson, W.M., Barnes, B.R. 2005. Trends over time among cotton cultivars released by the Oklahoma Agricultural Experiment Station. *Crop Sci*., 45: 966-980.

[35] Kerby, T.A., Hugie, W.V. 2006. Yield and quality improvement of significant D&PL varieties during the last 25 years. In Proc. Beltwide Cotton Conf. San Antonio, TX 3-6 Jan 2006 National Cotton Council Memphis, TN. pp: 858-865.

[36] Kuraparthy, V., Bowman, D.T. 2013. Gains in breeding upland cotton for fiber quality. *J. Cotton Sci.*,

[37] Flint-Garcia, S. A. (2013). Genetics and consequences of crop

domestication. Journal of Agricultural and Food Chemistry, 61(35), 8267-8276.

[38] Zhang, B.L., Yang, Y.W., Chen, T.Z., Yu, W.G., Liu, T.L., Li, H., Fan, X., Ren, Y., Shen, D., Liu, L., Dou, D., Chang, Y. 2012. Island cotton GbVe1 gene encoding a receptor-like protein confers resistance to defoliating and non-defoliating isolates of *Verticilium dahliae*. *PLoS One*, 7(12): 1-10.

[39] McCarty, J.C., Percy, R.G. 2001. Genes from exotic germplasm and their use in cultivar improvement in *Gossypium hirsutum* L. and *G. barbadense* L. pp. 65-80. In Jonkins J.J., Saha S. (ed.) Genetic improvement of cotton-emerging technologies. Sci. Publ. Enfield, NH.

[40] Tankley, S.D., Hewitt, J. 1988. Use of molecular markers in breeding for soluble solids in tomato- a re-examination. *Theor. And App. Genet*. 75: 811-823.

[41] Kumar, L.S. 1999. DNA markers in plant improvement: an overview, *Biotechnology Advances*, 17(2): 143-182.

[42] King R.C., Stansfield, W.D. 1990. A dictionary of genetics, Oxford University Press, New York, pp.188.

[43] Schulman, A. H. (2007). Molecular markers to assess genetic diversity. *Euphytica*, *158*(3), 313-321.

[44] Roychowhury, R., Taoutaou, A., Khalid, R.K., Mohamed R.A.G., Jagatpati, T. 2014. Crop improvement in the ear of climate change, In: Roychowhury (eds.), I.K International Publication Ltd.

[45] Collard, B.C., Mackill, D.J. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philo. Trans. Of the Royal Soc. Of Lon*., 363: 557-572.

[46] Kumpatla, S.P., Buyyarapu, R., Abdurakhmonov, I.Y., Mammadov, J.A. 2012. In: Abdurakhmonov I.Y., (ed.), Genomics-Assisted Plant Breeding in the 21st Century: Technological Advances and Progress, Plant Breeding, In Tech, Croatia, 131-184.

[47] Malik, W., Ashraf, J., Iqbal, M. Z., Ali Khan, A., Qayyum, A., Ali Abid, M., ... & Hasan Abbasi, G. (2014). Molecular markers and cotton genetic improvement: current status and future prospects. *The Scientific World Journal*, *2014*.

[48] Helentjaris, T., Slocum, M., Wright, S., Schaefer, A., Nienhuis, J. 1986. Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms, *Theoretical and Applied Genetics*, 72(6):761-769

[49] Welsh J., McClelland, M. 1990. Finger printing genomes using PCR with arbitrary primers, *Nucleic Acids Research*, 18(24): 7213-7218.

[50] Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van, D.L.T., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting, *Nucleic Acids Research*, 23(21): 4407-4414.

[51] Struss, D., & Plieske, J. (1998). The use of microsatellite markers for detection of genetic diversity in barley populations. *Theoretical and Applied genetics*, *97*(1-2), 308-315.

[52] Mackill, D.J., Nguyen, H.T., Zhang, J. 1999. Use of molecular markers in plant improvement programs for rainfed lowland rice. *Field Crops Res*., 64, 177-185.

[53] Koebner, R.M.D., Summers R.W. 2003, 21st century wheat breeding: plot selection or plate detection? *Trends in Biotechnology*, 21(2): 59-63

[54] Stuber, C.W., Polacco, M., Senior, M.L. 1999. Synergy of empirical

breeding, marker-assisted selection, and genomics to increase crop yield potential. *Crop Science*, 39(6): 1571-1583.

[55] Tuberosa, R., Salvi, S., Sanguineti, M.C., Maccaferri, M., Giuliani, S., Landi, P. 2003. Searching for quantitative trait loci controlling root traits in maize: a critical appraisal. *Developments in Plant and Soil Science*, 101: 35-54.

[56] Thomas, W. 2003, Prospects for molecular breeding of barley. *Annals of Applied Biology*, 142(1): 1-12.

[57] Williams, K. J. (2003). The molecular genetics of disease resistance in barley. *Australian Journal of Agricultural Research*, *54*(12), 1065-1079.

[58] Iqbal, M.J., Aziz, N., Saeed, N.A., Zafa, Y., Malik, K.A. 2001. Genetic diversity evaluation of some elite cotton cultivars by RAPD analysis. *Theo. Applied Gene*., 94:139-144. 13. Nordborg, M., Tavare, S. 2002. Linkage disequilibrium: what history had to tell us. *Trend Genet*., 18: 83-90.

[59] Rahman, M., Asif, M., Ullah, I., Malik, K.A., Zafar, Y. 2005. Overview of cotton genomic studies in Pakistan, Plant and Animal Genome Conference XIII, San Diego, USA

[60] Abdurakhmonov, I.Y., Kohel, R.J., Yu, J.Z., Pepper, A.E., Abdullaev, A.A., Kushanov, F.N., Salakhutdinov, I.B., Buriev, Z.T., Saha, S., Scheffler, B.E., Jenkins, J.N., Abdukarimov, A. 2008. Molecular diversity and association mapping of fiber quality in exotic *G. hirsutum* L. germplasm. *Genomics*, 92: 478-487.

[61] Paterson, A.H., Tanksely, S.D., Sorells, S.E. 1991. DNA markers in plant improvement. *Advances in Agronomy*, 46: 39-90.

[62] Mohan, M., Nair, S., Bhagwatt, A., Karisma, T., Yano, M., Bhatia, C.R., Sasaki, T. 1997. Genome mapping, molecular markers and marker assisted selection in crop plants. *Mol. Breed*., 3: 87-103.

[63] Zhu, Y., Chen, H., Fan, J., Wang, Y., Li, Y., Chen, J., et al. (2000). Genetic diversity and disease control in rice. Nature 406, 718-722. doi: 10.1038/35021046.

[64] Bolek, Y., Hayat K., Bardak, A., Azhar, M.T. 2016. Molecular breeding of cotton.: Adurakhmonov, I.Y. Cotton Research. Intech Publishers, in Cotton Research, ed I. Y. Abdurakhmonov (InTech): 123-166.

[65] Thornsberry, J.M., Goodman, M.M., Doebley, J., Kresovich, S., Nielson, D., Buckler, E.S. 2001. $Dwarf_8$ polymorphisms associate with variation in flowering time. *Nature Genet*., 28: 286-289.

[66] Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., Yue, G., Cong, L., Shang, H., Zhu, Z., Zou, C., Li, Q., Yuan, Y., Lu, C., Wei, H., Gou, C., Zheng, Z., Yin, Y., Zhang, X., Liu, K., Wang, B., Song, C., Shi, N., Kohel, R.J., Percy, R.J., Yu, Z.Y., Zhu, Y., Wang, J., Yu, S. 2012. The draft genome of diploid cotton *Gossypium raimondii* Ulbrich. *Nat. Genet*., 44: 1098-1103.

[67] Paterson, A.H., Wendel, J.F., Gundlach, H., Guo, H., Jenkins, J., et al. 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibers. *Nature*, 492: 423-427.

[68] Li, F., Fan. G., Wang, K., Sun, F., Yuan, Y., Guoli. Song., Qin, Li., Zhiying, Ma., Cairui, Lu., Zou, C., Chen, W., Liang, X., Shang, H., Liu, W., Shim C., Xiao, G., Goum C., Ye1, W., Xu. X., Zhang, X., Wei1, H., Li, Z., Zhang, G., Wang, J., Liu, K., Kohel, R.J., Percy, R.J., Yu, J.Z., Zhu, Y., Wang, J., Yu, S. 2014. Genome sequence of the cultivated *Gossypium arboreum*. *Nat. Genet*., 46:567-572.

[69] Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R. J., ... & Liang, X. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature biotechnology*, *33*(5), 524-530.

[70] Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., et al. 2015. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol*., 33: 531-537.

[71] Chen, Z., Scheffler, B., Dennis, E., Triplett B., Zhang, T., et al. 2007. Towards sequencing of (*Gossypium*) genomes. *Plant Physiol*., 145: 1303-1310.

[72] Jiang, C.X., Chee, P.W., Draye, X., Morrell, P.L., Smith, C.W., Paterson, A.H. 2000. Multi-locus interactions restrict gene introgression in interspecific populations of polyploid *Gossypium* (cotton). *Evolution.*, 54: 798-814.

[73] Paterson, A.H., sarangi, Y., Menz, M., Jiang, C.X., wright, R.J. 2003. QTL analysis of genotype x environment interactions affecting cotton fiber quality. *Theor. Appl. Genet*., 106: 384-396.

[74] Rong, J., Feltus, F.A., Waghmare, V.N., Pierce, G.J., Chee, P.W., Draye, X., et al. 2007. Meta-analysis of polyplois cotton QTL shows unequal distributions of sub-genomes to a complex network of genes and gene clustered implicated in lint fiber development. *Genetics*, 176: 2577-2588.

[75] Shappley, Z.W., Jenkins, J.N., Zhu, J., McCarty, J.C. 1998. Quantitative trait loci associated with agronomic and fiber traits of upland cotton. *J. Cotton Sci*., 4:153-163. 86. Lande, R., Thomson, R. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. *Genetics*, 124:743-756.

[76] Ulloa, M., Meredith Jr, W. R., Shappley, Z. W., & Kahler, A. L. (2002). RFLP genetic linkage maps from four F 2.3 populations and a joinmap of *Gossypium hirsutum* L. *Theoretical and Applied Genetics*, *104*(2-3), 200-208.

[77] He, D.H., Lin, Z.X., Zhang, X.L., Nie, Y.C., Guo, X.P., Feng, C.D., Stewart, Mc.D.J. 2005. Mapping QTLs contributing to yield and analysis of genetic effects in tetraploid cotton. *Euphytica*, 144(1-2):141-149

[78] Fang, D.D., Jenkins, J.J., Deng, D.D., McCarty, J.C., Li, P., Wu, J. 2014. Quantitative trait loci analysis of fiber quality using a random-mated recombinant inbred line population in upland cotton (*Gossypium hirsutum* L.). *BMC Genomics*, 15-397.

[79] Said, J.I., Lin, Z., Zhang, X., Song, M., Zhang, J.A. comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. *BMC Genomics,* 2013, 14:776.

[80] Mackay, I., Powell, W. 2007. Methods for linkage disequilibrium mapping in crops. Trends plant sci., *Trends Plant Sci*., 12: 57-63.

[81] Breseghello, F., Sorrells, M.E. 2006. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*, 172: 1165-1177.

[82] Nordborg, M., Borevitz, J.O., Bergelson, J., Berry, C.C., Chory J., Hagenblad, J., Kreitman, M., Maloof, J.N., Noyes, T., Oefner, P.J., Stahl, E.A., Weigel, D. 2002. The extent of linkage disequlibrium in *Arabdopsis thaliana*. *Nature Genetics*, 30: 190-193.

[83] Abdurakhmonov, I. Y., & Abdukarimov, A. (2008). Application of association mapping to understanding

the genetic diversity of plant germplasm resources. *International journal of plant genomics*, *2008*.

[84] Abdurakhmonov, I.Y., Buriev, Z.T., Saha, S., Pepper, A.E., Musaev, J.A., Almatov, A., Shermatov, S.E., Kushanov, F.N., Mavlonov, G.T., Reddy, U.K., Yu, J.Z., Jenkins., J.N., Kohel, R.J., Abdukarimov, A. 2007. Microsatellite Markers Associated with Lint Percentage Trait in Cotton, *Gossypium hirsutum*. *Euphytica,* 156(1-2): 141-156.

[85] Abdurakhmonov, I. Y. Saha, S., Jenkins, J.N., Buriev, Z.T., Shermatov, S.E., Scheffler, B.E., Pepper, A.E., Yu, J.Z., Kohel, R.J., Abdukarimov A. 2009. Linkage disequilibrium based association mapping of fiber quality traits in *G. hirsutum* L. variety germplasm. *Genetic*a, 136, 401-417.

[86] Lande, R., Thomson, R. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. *Genetics*, 124:743-756.

[87] Chee P., Draye, X., Jiang, C.X., Decaanini L., Delmonte, T.A. Bredjauer, R., Smith, C.W., Paterson, A.H. 2005. Molecular dissection of interspecific variation between *Gossypium hirsutum* and *Gossypium barbadense* (cotton) by a backcrossself approach: I. Fiber elongation. *Theor. Appl. Genet*. 111: 757-763.

[88] Kumar, P., Gupta, V.K., Misra, A.K., Modi, D.R., Pandey, B.K. 2009. Potential of molecular markers in plant biotechnology. *Plant Omics Journal*, 2: 141-162.

[89] 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: 829-847.

[90] Blenda, A., Fang, D.D., Rami, J.F., Garsmeur, O., Feng, L., Lacape, A. 2012, High density consensus genetic map of tetraploid cotton that integrates multiple component maps through molecular marker redundancy. *J. Genetics*, 91(3): 22-25.

[91] Rong, J.K., Abbey, C., Bowers J.E., Brubaker, C.L., et al. 2004. A 3347-locus genetic recombination map of sequence-tagged sites reveals types of genome organization, transmission and evolution of cotton *Gossypium*. *Genetics*, 166: 389-417.

[92] Mei, M., Syed, N.H., Gao, W., Thaxton, P.M., Smith, C.W. Stelly, D.M., Chen, Z.N. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium). *Theor. Appl. Genet*. 108:280-291.

[93] 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. *Theo. and App, Gene*. 109:167-175.

[94] Han, Z.G., Guo, W.Z., Song, X.L., Zhang, T.Z. 2004. Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. *Molecular Genetics and Genomics*, 272: 308-327.

[95] Guo, W., Cai, C., Wang, C., Han, Z., Song, X., Wang, K., Niu, X., Wang, C., Lu, K., Shi, B., Zhang, T. 2007. A microsatellite-based gene-riched linkage map of reveals genome structure, function, and evolution in *Gossypium*. *Genetics*, 176: 527-541.

[96] Lacape, J. M., Jacobs, J., Arioli, T., Derijcker, R., Forestier-Chiron, N., Llewellyn, D., ... & Viot, C. (2009). A new interspecific, *Gossypium hirsutum* × *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. *Theoretical and applied genetics*, *119*(2), 281-292.

[97] Blenda, A., Fang, D.D., Rami, J.F., Garsmeur, O., Luo, F., Lacape, J.M. 2012. A high-density consensus genetic map of tetraploid cotton that integrates multiple component maps through marker redundancy check. *PLoS ONE*, 7: e45739.

[98] 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.

[99] Yu, J.Z., Fang, D.D., Kohel, R.J., Ulloa, M., Hinze, L.L., Percy, R.G., Zhang, J., Chee, P., Scheffler, B.E, Jones, D.C. 2012. Development of a core set of SSR markers for the characterization of *Gossypium* germplasm. *Euphytica*, 187: 203-213.

[100] Zhu, H., Han, X., Lv, J., Zhao, L., Xu, X., Zhang, T., & Guo, W. (2011). Structure, expression differentiation and evolution of duplicated fiber developmental genes in *Gossypium barbadense* and *G. hirsutum*. *BMC plant biology*, *11*(1), 40.

[101] Marathi, B., Guleria, S., Mohapatra, T., Parsad, R., Mariappan, N., Kurungara, V. K., ... & Singh, A. K. (2012). QTL analysis of novel genomic regions associated with yield and yield related traits in new plant type based recombinant inbred lines of rice (Oryza sativaL.). *BMC plant biology*, *12*(1), 137.

[102] Lacape, J. M., Gawrysiak, G., Cao, T. V., Viot, C., Llewellyn, D., Liu, S., ... & Palaï, O. (2013). Mapping QTLs for traits related to phenology, morphology and yield components in an inter-specific *Gossypium hirsutum*× *G. barbadense* cotton RIL population. *Field Crops Research*, *144*, 256-267.

[103] Li, X., Yuan, D., Zhang, J., Lin, Z., & Zhang, X. (2013). Genetic mapping and characteristics of genes specifically or preferentially expressed during fiber development in cotton. *PLoS One*, *8*(1), e54444.

[104] Lacape, J. M., Jacobs, J., Arioli, T., Derijcker, R., Forestier-Chiron, N., Llewellyn, D., ... & Viot, C. (2009). A new interspecific, *Gossypium hirsutum*× *G. barbadense*, RIL population: towards a unified consensus linkage map of tetraploid cotton. *Theoretical and applied genetics*, *119*(2), 281-292.

[105] Reddy, O. U. K., Pepper, A. E., Abdurakhmonov, I., Saha, S., Jenkins, J. N., Brooks, T., ... & El-Zik, K. M. (2001). New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research.

[106] 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.

[107] Xiao, J., Wu, K., David, D.F., Stelly, D.M., Yu, J., Roy, G.C. 2009. New SSR markers for use in cotton (*Gossypium* spp.) improvement. *J. Cotton Sci*., 13: 75-157.

[108] McCouch, S.R., Zhao, K., Wright, M., Tung, C.W., Ebana, K., Thomson, M., Reynolds, A., Wang, D., DeClerck, G., Ali, M.L., McClung, A., Eizenga, G., Bustamanteet, C. 2010. Development of genome-wide SNP assays for rice. *Breed Sci*., 60, 524-535.

[109] Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet*. 12: 499-510.

[110] Feuillet, C., Leach, J.E., Rogers, J., Schnable, P.S., Eversole, K. 2011. Crop genome sequencing: lessons and rationales. *Trends Plant Sci*., 16: 77-88.

[111] Morrell, P. L., Buckler, E. S., & Ross-Ibarra, J. (2012). Crop genomics:

advances and applications. *Nature Reviews Genetics*, *13*(2), 85-96.

[112] Poland, J.A., Rife, T.W. 2012. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome*, 5: 92-102.

[113] Cai, C., Zhu, G., Zhang, T., and Guo, W. (2017). High-density 80 K SNP array is a

[114] Huang, B. E., Clifford, D., & Cavanagh, C. (2013). Selecting subsets of genotyped experimental populations for phenotyping to maximize genetic diversity. *Theoretical and applied genetics*, *126*(2), 379-388.

[115] Gupta, P.K., Roy, J.K., Prasad, M., 2001. Single nucleotide polymorphisms: a new paradigm molecular marker technology and DNA polymorphism detection with empahais on their use in plants. *Curr. Sci*., 80:524:535.

[116] Maughan, P.J., Yourstone, S.M., Jellen, E.N., Udall, J.A. 2009. SNP discovery via genome reduction, barcoding and 454-pyrosequencing in amaranth. *Plant Genome*, 2: 260-270.

[117] Harper, A.L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P., Bancroft I. 2012. Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nature Biorechnol*. 30: 798-802.

[118] Naoumkina, M.G., Thyssen, G., Fang, D.D., Hinchliffe, D.J., Florane, C., et al. 2014. The Li2 mutation results in reduced subgenome expression bias in elongation fibers of allotetraploid cotton (*Gossypium hirsutum* L). *PLoS One*, e90830.

[119] Xu, Y., Xu, C. & Xu, S. Prediction and association mapping of agronomic traits in maize using multiple omic data. Heredity 119, 174-184 (2017). https://doi.org/10.1038/hdy.2017.27.

[120] 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.

[121] Poland, J.A., Brown, P.J., Sorrells, M.E., Jannink, J.L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*, 7:32253 Genomics 18:654.

[122] Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E. 2011. A robust, simple genotyping-by- sequencing (GBS) approach for high diversity species. *PLoS ONE*, 6: e19379.

[123] Islam, M.S., Thyssen, G.N., Jenkins, J.N., Fang, D.D. 2015. Detection, validation, and application of genotyping-by-Sequencing based single nucleotide polymorphisms in upland cotton. *The Plant Genome*, 8:1-10.

[124] Schnable, P.S., Liu, S., Wu, W. 2013. Genotyping by next-generation sequencing. U.S. Patent Appl. No. 13/739,874.

[125] Ma, X. F., Jensen, E., Ale

[126] Spindel, J., Wright, M., Chen, C., Cobb, J., Gage, J., Harrington, S., Lorieux, M., Ahmadi, N., McCouchet, S. 2013. Bridging the genotyping gap: Using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet*., 126:2699-2716.

[127] Sonah, H., Bastien, M., Iquira, E., Tardivel, A., et al. 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE*, 8(1): 1-9.

[128] Huang, Y.F., Poland. J.A., Wight, C.P., Jackson, E.W., Tinker, N.A. 2014. Using genotyping-bysequencing (GBS) for genomic discovery in cultivated oat. *PLoS One*, 9: e102448

[129] Islam, M., Zeng, L., Thyseen, G., Delhom, C., Kim, H. 2016b. Mapping by sequencing in cotton (*Gossypium hirsutum*) line MD52ne identified genes for fiber strength and its related quality attributes. *Theor. Appl. Genet*., 129: 1071-1086.

[130] Hayat, K.B. 2018. Association analysis and Mapping of Fiber quality in upland cotton. Ph. D Thesis submitted to Kahramanmaras Sutcu Imam University/Turkey.

[131] Yang, X.H., Yan, J.B., Zheng, Y.P., Yu, J.M., Li, J.S. 2007. Reviews of association analysis for quantitative traits in plants. *Acta Agron Sin*., 33(4):523-30.

[132] Nie X., Huang, C., You, C., Li, W., Zhao, W et al. 2016. Genome-wide SSR-based association mapping for fiber quality in nation-wide upland cotton inbreed cultivars in China. *BMC Genomics,* 17:1-16. 131. Yang, X.H., Yan, J.B., Zheng, Y.P., Yu, J.M., Li, J.S. 2007. Reviews of association analysis for quantitative traits in plants. *Acta Agron Sin*., 33(4):523-30.

[133] Gould, S. J., Johnston, R.F. 1972. Geographic variation. *Annu. Rev. Ecol. Syst*. 3: 457-498

[134] Roesti, M., Moser, D., Berner, D. 2013 Recombination in the three-spine stickleback genome-patterns and consequences. *Mol. Ecol*., 22: 3014-3027.

[135] Pritchard, J.K., Stephens, M., Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959.

[136] Peleg, Z., Fahima, T., Abbo, S., et al. 2008. Genetic structure of wild emmer wheat populations as reflected by transcribed versus anonymous SSR markers. *Genome*, 51: 187-195.

[137] Nei, M. 1972. Genetic distances between populations. *American Naturalist*, 106, 283-292.

[138] Rogers, J.S. 1972. Measures of genetic similarity and genetic distance. Studies in genetics. VII. Univ. Texas Publ. 7213,145-153.

[139] Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89, 583-590.

[140] Yan, J. B., Kandianis, C.B., Harjes, C.E., Bai, L., et al. 2010. Rare genetic variation at Zea mayscrtRB1 increases beta-carotene in maize grain. *Nat. Genet*., 42:322-327.

[141] Badigannavar, A., and Myers, G. O. (2015). Genetic diversity, population structure.

[142] Mei, H., Zhu, X., Zhang, T. 2013. Favorable QTL alleles for yield and its components identified by association mapping in Chinese upland cotton cultivars. *PLoS One*, 8: e82193.

[143] Zhang, T., Qian, N., Zhu, X., Chen, H., Wang, S., Mei, H., Zhang, Y. 2013. Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in china. *PLoS One*, 8(2): 1-12.

[144] Qin, H., Chen, M., Yi, X., Bie, S., Zhang, C., Zhang, Y., Lan, J., Meng, Y., Yuan. Y., Jiao, C. 2015. Identification of associated SSR markers for yield component and fiber quality traits based on frame map and upland cotton collections. *PLoS One*, 10(1): 1-16.

[145] Saeed, F., Farooq, J., Mahmood, A., Hussain, T., Riaz, M., Ahmad, S. 2014. Genetic diversity in upland cotton for cotton leaf curl virus disease, earliness

and fiber quality. *Pak. J. Agric. Res*. 27, 226-236.

[146] Li, C., Zhang, J., Hu, G., Fu, Y., Wang, Q. 2016. Association mapping and favorable allele mining for node of first fruiting/sympodial branch and its height in Upland cotton (*Gossypium hirsutum* L.). *Euphytica*, 210:57-88

[147] Liu, G.Z., Mei, H.X., Wang, S., Li, X.H., Zhu, X.F., Zhang, T.Z. 2015. Association mapping of seed oil and protein contents in upland cotton. *Euphytica*, 205:637-645.

[148] Zeng, L., Meredith, W. R., Gutiérrez, O. A., & Boykin, D. L. (2009). Identification of associations between SSR markers and fiber traits in an exotic germplasm derived from multiple crosses among Gossypium tetraploid species. *Theoretical and applied genetics*, *119*(1), 93-103.

[149] Iqbal, M.A., Rahman, M. 2017. Identification of marker-trait associations for lint traits in cotton. *Front Plant Sci*., 8: 86.

[150] Sun, Z., Wang, X., Liu, Z., Gu, Q., Zhang, Y.,. 2017. Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. *Plant Biotechnology Journal*, 1-15.

[151] Atwell, S., Huang, Y.S., Vilhjalmsson, B.J., Willems, G., Horton, M. et al. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature, 465 (7298): 627-631.

[152] Huang, B. E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K., & Cavanagh, C. R. (2012). A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant biotechnology journal*, *10*(7), 826-839.

[153] Wang, X., Yu, Y., Sang, J., Wu, Q., Zhang, X., Lin, Z. 2013. Intraspecific linkage map construction and QTL mapping of yield and fiber quality of *Gossypium barbadense*. *Aust. J. Cop Sci*., 7:1252-1261

[154] Fang, H., Zhou, H., Sanogo, S., Lipka, A. E., Fang, D. D., et al. (2014). Quantitative trait locus analysis of Verticillium wilt resistance in an introgressed recombinant inbred population of Upland cotton. *Molecular breeding*,*33*(3), 709-720.

[155] Islam, A., Gregory, N., Thyssen, Johnie, N., Jenkins, J.J., 2016a. A MAGIC population-based genome-wide association study reveals functional association of GhRBB1_A07 gene with superior fiber quality in cotton. *BMC Genomics*, 17:903.

[156] Gapare, W., Conaty, W., Zhu, Q. H., Liu, S., Stiller, W., Llewellyn, D., & Wilson, I. (2017). Genome-wide association study of yield components and fibre quality traits in a cotton germplasm diversity panel. *Euphytica*,*213*(3), 66.