Journal of Life and Social Sciences
eISSN: 3006-2675
www.bbasrjlifess.com
*J. Life Soc. Sci, Volume, **1**: 1*

Review Article ***Open Acces***

EXPLORING THE GENETIC RESOURCES OF COTTON

ABBAS A[1*], REHMAN AU[1], BUKHARI MS[2], ABBAS MZ[1]

[1]*Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences, University of the Punjab, Lahore, Pakistan*
[2]*Agricultural Research Station, Bahawalpur, Pakistan*
**Correspondence author email address: ali.bukhari91112@gmail.com*

(Received, 10th January 2022, Revised 8th July 2022, Published 12th July 2022)

Abstract *Since its first use in 6000 BC, cotton (genus Gossypium) has become a major natural textile in the global market. Modern tools such as web databases, microsatellite databases, and comparative QTL resources have been developed to evaluate the consequences of human dispersal and selection on different strains of Gossypium. Out of the fifty species of Gossypium, only four have been domesticated, leading to significant changes in lint percentage (40%), fibre length (22%), and boll size. The biggest challenge with domestication is the lack of genetic variety. This is particularly evident in Gossypium hirsutum L. upland cotton cultivars compared to Pima and Egyptian cotton cultivars of Gossypium barbadense. The latter possess a higher degree of genetic diversity due to the introduction of G. hirsutum genes into G. barbadense cultivars. Including genes from all types of Gossypium in cultivated cotton species is very important. The use of genome-wide markers such as Simple Sequence Repeat (SSR), Restriction Fragment Length Polymorphism (RFLP), Amplified Fragment Length Polymorphism (AFLP), and Random Amplified Polymorphic DNA (RAPD) enabled the discovery of 16,162 public SSRs and 312 mapped RFLP sequences. These markers were further employed to study various plant traits reported in 26 mapping populations. These included qualitative traits such as fiber quality, yield, leaf and flower shape, trichome density and placement, disease protection, and quantitative traits such as quantitative trait loci (QTLs). After a suitable comparison of the mapped populations, 432 QTLs were associated with 3,475 loci within 11 mapping groups. Furthermore, a meta-analysis of over 1,000 QTLs derived from backcross and hybrid inbred line populations with the same parents revealed the most accurate meta-clusters for fiber color, fineness, and length. The cotton genome has undergone enrichment by incorporating genes obtained from distantly related organisms via diverse transformation techniques.*

[*Citation: Abbas, A., Rehman, A.U., Bukhari, M.S., Abbas, M.Z. (2022). Exploring the genetic resources of cotton. Biol. Agri. Sci. Res. J., **2022**: 1*]

Keywords: *Gossypium hirsutum, Diversity, Cultivated Species, QTL Map, Hybrid cotton, Structural genomic resources, DNA marker, RFLP, RAPD, SSR*

Introduction

Cotton is widely regarded as the most important natural fiber crop globally, as it sustains one of the largest textile industries in the world. It has been estimated that the global economic impact of this industry is around $500 billion. An opportunity arises to understand the divergence of domesticated cotton from its wild origins by examining the genetics of approximately 50 cotton species. This could assist us in comprehending how the quality of lint fiber altered over time and how the phenomenon of polyploidy influences the creation of lint. Moreover, research is being conducted to determine the amount of genetic variation in cotton and how this can be used to make cotton production and lint quality more sustainable (Batool et al., 2023; Hafeez et al., 2021; Iqbal et al., 2023ab; Abbas et al.,2015). Additionally, scientists are exploring potential bio-based alternatives to petrochemicals. Scientists have employed cutting-edge genetic techniques to analyze the genome of the Gossypium (cotton) genus to explain why cotton production has decreased over the last few decades. This research has enabled people to keep producing raw cotton. However, since cotton has been domesticated for a considerable time, it has a limited genetic range (Abbas et al., 2016; Zafar et al., 2022; Puspito et al., 2015; Rehman et al., 2017). Although some classical DNA resources have been used, we must seek new solutions in the face of contemporary problems. Genomic tools have enabled the alteration of fiber aesthetics and the development of varieties resistant to pests, diseases, and environmental stressors (Abelson, 1998). This study analyzes cotton's genetic reserves' past, present, and potential future applications. It is possible to address challenges

related to cotton yield by using the available genetic reserves more effectively.

Range of Genetic Variation

Cotton (Gossypium spp., family Malvaceae) is a vital crop and commodity in the global economy. Fryxell et al. (1992) identified 45 diploid species, with 2n = 2=26, and 5 tetraploid species, with 2n = 4=52. Wendel (1989) proposed that interbreeding between two ancestral species ("A" and "D") resulted in the emergence of tetraploid species around 1-2 million years ago (MYA). Before this, it was believed that the common ancestor of these two species had lived between 4-11 MYA. Tetraploid cotton has two sets of genetic material, or sub-genomes, with most genes duplicated, creating multiple versions of each gene. Brubaker et al. (1999) showed that the gene structure of this variety of cotton is identical to that of the diploid version. *Gossypium hirsutum* L. and *Gossypium barbadense* L., GH and GB, are the two most commonly cultivated cotton species. Today's cotton has been carefully selected to produce the most efficient fiber for harvesting and processing. However, this selection has reduced genetic variation for certain traits, such as drought resilience (Rosenow et al. 1983). To address this, wild alleles can be introduced into cultivation, which may help to solve various agricultural problems (Gur and Zamir 2004). Wild tetraploid species such as *Gossypium darwinii* (GD), *Gossypium tomentosum* (GT), and *Gossypium mustelinum* (GM) can be crossed with cultivated cotton species to create normal hybrids, which can then produce several fertile offspring (Waghmare et al. 2005). These wild tetraploid species are comparatively more drought and heat-tolerant, making them a valuable source of genetic variation.

Exploring Intraspecific Genetic Variation

The low genetic diversity of cotton has contributed to the current stagnation in cotton yield improvement. This issue has been addressed by forming the "Blue Ribbon Committee", comprising cotton scientists from the governmental and commercial sectors (Helms 2000). The committee examined the data from the National Cotton Variety Tests (Rayburn et al. 1999) and the National Agricultural Statistics Service (USDA) to investigate the reasons for yield deflation. A linear model was used, which estimated that the growth in cotton production over the previous 39 years was 6.7 kg ha/year (1.3% annual growth rate). Additionally, it was established that between 1970 and 1985, there was an increase in cotton yield. Beginning in 2000, the average global cotton output increased from 566 kg/ha in 1998/99 to 793 kg/ha in 2007/2008, a record high. This growth can be attributed to technological advancements, better resource utilization, and changes in cotton production regions. Unfortunately, since 2007/2008, cotton production has decreased, with the world's cotton crop predicted to yield 725 kg/ha in 2009/2010, a 5% decrease from the previous season. This decline can be attributed to the reduced yields in the USA and China over the past three seasons. May et al. (1995) and Rahman et al. (2002, 2005, 2008) have demonstrated that there is limited genetic diversity within the cotton gene pool used to create new cultivars due to the consistent use of the same few genetic sources as well as the planting of similar seeds over large areas. This has led to a heightened level of genetic homogeneity in the field, with Van Esbroeck et al. (1998) reporting a field uniformity of 38% in the mid-south US production area by 1995. The potential of incorporating foreign accessions to increase the genetic diversity of cultivated cotton has been widely understudied, however, several accessions with advantageous traits have been genetically modified to become more "breeder-friendly". McCarty and Jenkins (1992) noted that exotic cottons are often sensitive to light and flower too late for temperate production. Calhoun et al. (1997) further discovered that the pedigrees of current cotton cultivars lack any exotic germplasm. Adkisson et al. (1982) suggested that early growth of cotton can protect plants from various stresses such as insects, diseases, drought, and cold. Quisenberry (1977) reported a high heritability (0.62) for maturity date, and Kohel et al. (1974) found that day-neutrality is partially dominant for cotton flowering. In light of the difficulty of breeding photoperiodic cottons, McCarty and Jenkins (1992) developed the "Cotton Conversion Program" which has since "converted" numerous exotic cottons through backcrossing.

Previous studies that attempted to classify exotic germplasm into botanical races based on shape and location (Fryxell, 1979) have been disproved by limited neutral marker data (Brubaker and Wendel, 1993). Moreover, Jiang et al. and Liu et al.'s (2000b) discovery of introgressed chromatin being lost from the progeny of advanced backcrosses faster than anticipated suggests that traditional studies on cotton introgression may have overestimated its impact. Due to the lack of DNA markers, measuring the effects of particular chromosome segments was previously hard. This hypothesized artifact further complicates distinguishing the pros and cons of introgression. However, by pinpointing individuals who still possess foreign chromatin, a comprehensive molecular map can help solve these issues and analyze the benefits and drawbacks of each introgressed chromosome segment.

Genomic Databases

Exploring Structures of Genomes Using Traditional Methods

Cytogenetic Analysis of Stock Populations

Through monosomics and telosomics, new stocks of the tetraploid TM-1 (GH) 3–79 (GB) (Endrizzi and Ramsay, 1979) and TM-1 GT (Saha et al., 2006) were developed. This enabled the successful placement of 23 of the 26 chromosomes in their respective locations; the remaining three were not identified. An abundance of SSRs and RFLPs were assigned to the correct chromosomes with the help of this technique.

Nested Association Mapping (NAM), a relatively new method in the realm of QTL mapping, is used to identify genes that control complicated trait expression (Fig. 1). This method combines the high-resolution resolution of association mapping in measuring QTLs by combining data from the parental genotypes and the genotypes of segregates. The 'parental' SNPs identified are those observed in only one parent, allowing for the determination of whether a chromosomal region in each segregate is inherited from the common parent or the alternate parent. Yu et al. (2008) proposed a nested NAM approach that employs a RIL (Recombinant Inbred Line) population whose parents have been crossed and backcrossed multiple times. This specific QTL mapping method should be able to identify QTLs through a genome-wide approach with higher accuracy when both linkage and LD (Linkage Disequilibrium) are simultaneously used (Yu et al., 2008).

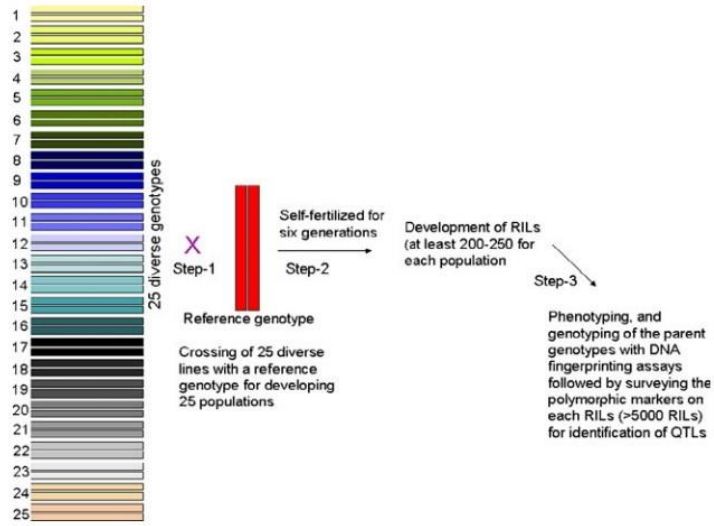

Fig. 01 Schematic illustration of development of nested association mapping (NAM) populations. *RIL* recombinant inbred line, *QTL* quantitative trait loci

Exploring the Role of Genetics in Mapping Genomes

Approximately 5,000 markers, including 3,300 RFLP, 700 AFLP, 1,000 SSR, and 100 SNP markers, are available in the public database for genetic mapping. This has enabled the construction of a reference map for the tetraploid cotton Gossypium with a length of 2,324.7 cM, containing 3,016 loci discovered by 2,337 probes. This map is believed to have originated from the At, Dt, and D genes in the past. Subsequent interspecific crosses have resulted in maps that are abundant with markers, which are often utilized as reference maps for locating QTLs (Lacape et al., 2003; Guo et al., 2007; Lacape et al., 2007; Yu et al., 2007). The molecular maps of cotton have also been applied to identify DNA markers linked to fibre traits (Rahman et al., 2009; Cai et al., 2010) and to investigate the genomics of drought tolerance (Saranga et al., 2001; Saranga et al., 2004) and resistance to diseases such as Xanthomonas (Wright et al., 1998).In recent years, there has been a surge of research and development on nuclear restorers of cytoplasmic male sterility to cultivate hybrid cotton. Leaf morphology, leaf color, and pubescence, a specific characteristic of GH, have been extensively studied (Song et al. 2005; Hao et al. 2008; Ali et al. 2009a). The physical features and nutritional value of cotton seeds have also been examined in great detail (Song and Zhang, 2007). The genetic map presented by Jiang et al. (2005) provides an example of how an in-depth analysis of the correlation between genes and traits and their organization in the genome can be achieved. This was made possible by combining and assessing numerous QTL mapping studies with a shared reference map. The genetic map comprises 3,475 loci and 432 QTLs in 11 populations (Rong et al. 2005b). The development of consensus maps has enabled researchers to explore the correlation between cotton QTLs and Arabidopsis genes relevant to fibre or trichome formation. Rong et al. (2005b) used such a map to map QTLs onto the map, while Lacape et al. (2010) conducted a meta-analysis of around 1000 QTLs from BC and RIL populations that had descended from the same parents. This analysis revealed the most consistent meta-clusters for fibre colour, fineness, and length. To facilitate various genomic projects, the CMap resource can be accessed from http://chibba.agtec.uga.edu/cgi-bin/cmap/viewer.

3.1.3 Exploring the Possibilities of Cultivating New Populations

Mapping groups use molecular markers to generate genetic maps. Upland cotton intraspecific crosses have produced numerous linkage maps (Wang et al., 2007a; Ma et al., 2008). When GH and GB were crossed, many polymorphisms were identified, which is encouraging. Most of the larger cotton genetic maps available originate from these interspecific populations (Rahman et al., 2009). Nested association mapping and linkage disequilibrium (LD) are newer techniques for gene mapping. Linkage disequilibrium is particularly strong in populations generated through crossing two inbred parents. Association mapping is a technique used to identify QTLs with greater

resolution in natural populations, typically resulting from multiple rounds of recombination. Nested Association Mapping (NAM) is a more recent method that combines the statistical power of QTL mapping with the high (potentially gene-level) resolution of association mapping. This technique is 'nested' as the offspring share the same parent, but each has a different one. With NAM, several unique SNPs (e.g. between 1,000 and 2,000) from the common parent are genotyped alongside super-resolution genotyping of the parents (e.g. whole-genome sequencing). This data is then used to determine which chromosomal sections in the offspring come from the common parent and which come from the alternative parent. This allows for high-resolution genetic data to be inferred from parents to offspring. By analyzing associations in the entire population, it is sometimes possible to map traits down to the level of a single gene (Yu et al., 2008). Additionally, Yu et al. (2008) proposed a nested NAM approach utilizing data from RIL populations created when inbred parents are crossed multiple times. This method integrates linkage and LD information to map QTLs in a genome-wide approach, while providing improved mapping resolution when both linkage and LD are combined (Yu et al., 2008).

Based on LD, association mapping is an effective tool for studying and exploiting natural genetic variation and has been employed in marker-assisted selection (MAS). For lint quality factors in G. hirsutum, genome-wide LD and association mapping with SSRs revealed that the exotic germplasm of this species contains valuable genetic variation (Abdurakhmonov et al., 2008). These findings suggest that association mapping has great potential for MAS programs. Linkage disequilibrium is particularly strong in populations generated through crossing two inbred parents. Association mapping is a technique used to identify QTLs with greater resolution in natural populations, typically resulting from multiple rounds of recombination. Nested Association Mapping (NAM) is a more recent method that combines the statistical power of QTL mapping with the high (potentially gene-level) resolution of association mapping. This technique is 'nested' as the offspring share the same parent, but each has a different one. With NAM, many unique SNPs (e.g. between 1,000 and 2,000) from the common parent are genotyped alongside super-resolution genotyping of the parents (e.g. whole-genome sequencing). This data is then used to determine which chromosomal sections in the offspring come from the common parent and which come from the alternative parent. This allows for high-resolution genetic data to be inferred from parents to offspring. By analyzing associations in the entire population, it is sometimes possible to map traits down to the level of a single gene (Yu et al., 2008). Additionally, Yu et al. (2008) proposed a nested NAM approach utilizing data from RIL populations created when inbred parents are crossed multiple times. This method integrates linkage and LD information to map QTLs in a genome-wide approach while providing improved mapping resolution when both linkage and LD are combined (Yu et al., 2008). Based on LD, association mapping is an effective tool for studying and exploiting natural genetic variation and has been employed in marker-assisted selection (MAS). For lint quality factors in G. hirsutum, genome-wide LD and association mapping with SSRs revealed that the exotic germplasm of this species contains valuable genetic variation (Abdurakhmonov et al., 2008). These findings suggest that association mapping has great potential for MAS programs.

Comparison of Dry Lab Resources in Genomics

Transferring Genome Data from Model Species to Cotton

Bowers et al. (2003) proposed that the similarities between Malvales (including cultivated cotton species) and Brassicales (including Arabidopsis) might be attributed to a shared ancestor from around 80 million years ago. To further explore this, Rong et al. (2005b) analyzed the homology of ESTs from cDNA libraries of *Gossypium Raimondi*, *G. arboreum*, and *G. hirsutum*. Their findings revealed that 55-70% of the cotton ESTs had homologs in the Arabidopsis genome while 30-45% of cotton genes were not found in Arabidopsis. Moreover, using two types of analysis, they discovered that approximately half of the cotton genome is the same as that of Arabidopsis. This implies that functional changes, gene loss, and genomic alterations must have occurred to enable these two plants' evolution. Given the size of the cotton genome (2,246 mega bases), it is challenging to pinpoint or alter genes that influence the quality and growth of cotton fiber. However, research on Arabidopsis trichomes has been beneficial in this regard. Several scientists, such Jacoby et al., (2008), and Plett et al. (2009), have studied the action and development of trichomes. The MYB gene is a potential area where these findings can be applied to cotton. Walker et al. (1999) discovered that the R2/R3 repeat-containing MYB genes GL1, GL3, and WER serve as positive regulators, while TRY and CPC act as negative regulators Wada et al. (2002). MYB genes have three main functional domains: a basic DNA-binding domain (DBD) with two to three imperfect repeats (R1, R2, and R3), a negative regulatory domain, and an acidic activating domain. Leveraging the DBD as a guide, Hsu et al. (2005) identified 65 MYB genes in cotton, some of which were associated with the production of cotton thread. According to its length and origin, cotton fibre is classified as "fuzz" or "lint" (Fryxell, 1963). The research conducted by Rong et al. (2007) has been a major success for cotton and Arabidopsis comparative genomics. They used cotton-Arabidopsis synteny to examine the positions of fiber-related QTLs and mutants in correlation with cloned fiber prospect genes. Such correlations have been documented in the CMap database

(www.plantgenome.uga.edu/cmap). Moreover, Rong et al. (2005b) found a strong connection between the number of QTLs associated with cotton fiber growth and the number of Arabidopsis candidate genes. This may be useful in the process of cloning genes from cotton.

Exploring Further Possibilities

Identification of Genes for Cotton Variety Improvement

Model plants such as Arabidopsis have been tremendously useful in providing insight into some metabolic pathways, including salt tolerance (Tester and Davenport 2003), flower development and disease resistance. By comparing genetic information between Arabidopsis and cotton, agriculturally important genes can be identified. Additionally, the full genome sequence of Arabidopsis can be used to fill in a desired region in the cotton map by finding similar cotton ESTs to Arabidopsis genes. A system was developed to facilitate further gene transfer to produce fertile and stable GH cv. Coker-312 transformants produce much better results than particle bombardment (Asad et al., 2008).

Bioengineered cotton types, known as Bollgard cotton, produce a protein from *Bacillus thuringiensis*, effectively eliminating certain lepidopterous cotton pests. The first-generation Bollgard cotton contains the Cry 1Ac gene, while the second-generation contain both Cry 1Ac and Cry 2Ab genes. Since their introduction in 1996, these cotton have been planted on 62% of the world's cotton-growing land, reducing the cost of making cotton by acting as an alternative to chemical herbicides for controlling the cotton bollworm, Helicoverpa zea, Heliothis virescens, and Pectinophora gossypiella (Perlak et al., 2001).

The Consequences of Implications

Optimizing Breeding Decisions Using Knowledge

In the 1930s, hybrid corn was grown on a large scale for the first time, resulting in its eventual spread worldwide due to its capacity to generate more varieties of maize than open-pollinated maize. As per Dong et al. (2006), heterosis breeding has since been identified as a significant genetic tool that can increase food production in multiple species. Zhu et al. (2008) have studied the heterosis of cotton, with the primary aim of breeders being to utilize heterosis to enhance lint yield. While hybrid energy for fibre output and quality and physiological traits such as photosynthetic rate in okra leaf cotton have been claimed, the yield benefits of cultivating cotton hybrids are not as high as those reported for maize.

Marker-Assisted Selection for Traits of Interest

Several scientific studies have demonstrated the use of MAS to improve cotton characteristics. Wright et al. (1997) identified QTLs related to leaf pubescence while Ullah (2009) found QTLs related to dryness. Rahman et al. (2002) and Ali et al. (2009a,b) discovered DNA markers linked to hairiness, nectar absence, and red leaf colour. Lan et al. (1999) identified a gene that could change male cotton plants into fertile females. Zhang and Stewart (2004) and Feng et al. (2005) identified markers for the restorer genes Rf1 and Rf2 in two cotton lines from the D2 genome, which were then used to create restorer maternal lines. Rahman et al. (2008) tested later generations of cotton plants for markers of gene resistance to CLCuD, creating two resistant cotton lines, NIBGE-2 and NIBGE-115. Wang et al. (2009) highlighted the usefulness of Fusarium wilt QTLs in MAS, while Asif (2010) showed that MAS can be used to improve fibre quality. Guo et al. (2003) reported that nine DNA markers (three SSRs and six RAPDs) mapped to one linkage group were linked to two QTLs for fibre strength. On the Chr-24 (D8) chromosome, Chen et al. (2009) improved the mapping of an important QTL for fibre strength. The use of markers associated with G. barbadense QTLs discovered in MAS has been beneficial in generating high-quality lint. Mumtaz (2007) reported that SSRs have been used to monitor the entry of genomic or loci from GB into GH, resulting in a 2- to 3-mm increase in fibre length. In addition to being used in cotton breeding, AFLPs have been linked to various agronomic and fibre properties. Wu et al. (2009) identified 56 QTLs with a LOD greater than 3.0, including one QTL related to fiber lengthening and involving 14 different agronomic and fiber variables. Cotton fiber is a great single-celled model system for studying processes such as the construction of cellulose and cell proliferation, given its distinction from the protoderm of developing seeds, which consists of one cell and does not branch out. About half a million quasi-synchronous processes must be produced to produce this fiber.

Translating Genes to Genomes

Plant Translational Genomics has been instrumental in understanding how different food plants react to various stresses, their growth patterns, and the flavor they produce. This is an especially difficult when dealing with polyploidy plants (Salentijn et al., 2007). Plant genomes such as Arabidopsis (The Arabidopsis Genome Initiative, 2000), poplar (Tuskan et al. 2006), sorghum, and soybean have been utilized to discover new solutions to complex issues. One system employed is the candidate gene strategy (Pflieger et al. 2001).

Exploring Genetic Resources of Non-Human Species

A sequence of 2,062 bases, known as the GaMYB2 promoter, was taken from G. arboreum and fused with a -glucuronides (GUS) gene, allowing scientists to track its activity in developing fiber cells and trichome in various parts of the plant, such as stems, leaves, and bracts. An activator was used in Arabidopsis to activate the gene only in its trichome. However, tobacco plants contain several trichomes, such as multicellular simple trichomes and glandular secreting trichomes, which differ from the unicellular, non-glandular simple trichomes found on cotton and Arabidopsis. By transferring genes from cotton plants

to other species, scientists can study the various functions of these trichomes.

Conclusion

A deep concern exists regarding the potential stagnation or even decrease in cotton production. The lack of genetic diversity among cotton cultivars and germplasm makes them more vulnerable to threats from living and nonliving sources. To address this issue, hybridization through G. arboreum, Gossypium gossypioides, G. herbaceum, G. barbadense, and Gossypium laxum with G. hirsutum and/or G. barbadense may help to overcome cytogenetic boundaries and provide advantageous traits to cultivated cotton varieties, thus improving the genetic diversity and introducing genes for traits not present in the cultivated species. Research such as nested association mapping studies and TILLING populations should be conducted to maximize the potential outcomes of genome sequencing efforts. Additionally, comparative mapping is an important tool, as it allows the information from well-studied genomes to be applied to those that have not been as thoroughly investigated. Once the cotton genome has been sequenced, molecular genomic tools can be further improved. Comparing the genomes of cotton and Arabidopsis has the potential to aid in understanding how biological processes operate. Sequencing the cotton genome can provide insight into the minor alterations in the cotton genome at the nucleotide level, which can modify the functioning of genes and, thus, the essential economic features of cotton.

REFERENCES

Abbas, H. G., Mahmood, A., & Ali, Q. (2016). Zero tillage: a potential technology to improve cotton yield. *Genetika*, *48*(2), 761-776.

Abbas, H. G., Mahmood, A., & Ali, Q. (2015). Genetic variability and correlation analysis for various yield traits of cotton (Gossypium hirsutum L.). *Journal of Agricultural Research*, *53*(4), 481-491.

Abdurakhmonov, I. Y., Kohel, R. J., Yu, J. Z., Pepper, A. E., Abdullaev, A. A., Kushanov, F. N., ... & Abdukarimov, A. (2008). Molecular diversity and association mapping of fiber quality traits in exotic G. hirsutum L. germplasm. *Genomics*, *92*(6), 478-487.

Abelson, P. H. (1998). A third technological revolution. *Science*, *279*(5359), 2019-2109.

Adkisson, P. L., Niles, G. A., Walker, J. K., Bird, L. S., & Scott, H. B. (1982). Controlling cotton's insect pests: a new system. *Science*, *216*(4541), 19-22.

Asad, S., Mukhtar, Z., Nazir, F., Hashmi, J. A., Mansoor, S., Zafar, Y., & Arshad, M. (2008). Silicon carbide whisker-mediated embryogenic callus transformation of cotton (Gossypium hirsutum L.) and regeneration of salt tolerant plants. *Molecular biotechnology*, *40*, 161-169.

Asif, M. (2010). *Genomic analysis for quality traits in cotton (Gossypium hirsutum L.) by DNA fingerprinting technology* (Doctoral dissertation, Bahauddin Zakariya University (BZU), Multan).

Batool, F., Hassan, S., Azam, S., Sher, Z., Ali, Q., & Rashid, B. (2023). Transformation and expressional studies of GaZnF gene to improve drought tolerance in Gossypium hirsutum. *Scientific Reports*, *13*(1), 5064.

Bowers, J. E., Chapman, B. A., Rong, J., & Paterson, A. H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature*, *422*(6930), 433-438.

Brubaker, C. L., Paterson, A. H., & Wendel, J. F. (1999). Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. *Genome*, *42*(2), 184-203.

Brubaker, C. L., Paterson, A. H., & Wendel, J. F. (1999). Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. *Genome*, *42*(2), 184-203.

Buckler, E. S., Yu, J. I. A. N. M. I. N. G., Holland, J. B., & McMullen, M. D. (2008). Genome-wide complex trait dissection through nested association mapping. *Genetics*, *178*, 539-551.

Cai, Y., Xie, Y., & Liu, J. (2010). Glandless seed and glanded plant research in cotton. A review. *Agronomy for sustainable development*, *30*, 181-190.

Dong, J., Wu, F., Jin, Z., & Huang, Y. (2006). Heterosis for yield and some physiological traits in hybrid cotton Cikangza 1. *Euphytica*, *151*, 71-77.

Endrizzi, J. E., & Ramsay, G. (1979). Monosomes and telosomes for 18 of the 26 chromosomes of Gossypium hirsutum. *Canadian Journal of Genetics and Cytology*, *21*(4), 531-536.

Feng, C. D., Stewart, J. M. D., & Zhang, J. F. (2005). STS markers linked to the Rf<? A3B2 show $132#?> 1 fertility restorer gene of cotton. *Theoretical and applied genetics*, *110*, 237-243.

Fryxell, P. A. (1979). *The natural history of the cotton tribe (Malvaceae, tribe Gossypieae)*. Texas A & M University Press..

Fryxell, P. A., Craven, L. A., & McD, J. (1992). A revision of Gossypium sect. Grandicalyx (Malvaceae), including the description of six new species. *Systematic Botany*, 91-114.

Guo, W., Cai, C., Wang, C., Han, Z., Song, X., Wang, K., ... & Zhang, T. (2007). A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in Gossypium. *Genetics*, *176*(1), 527-541.

Gur, A., & Zamir, D. (2004). Unused natural variation can lift yield barriers in plant breeding. *PLoS biology*, *2*(10), e245.

Hafeez, M. N., Khan, M. A., Sarwar, B., Hassan, S., Ali, Q., Husnain, T., & Rashid, B. (2021). Mutant Gossypium universal stress protein-2 (GUSP-2) gene confers resistance to various abiotic stresses in E. coli BL-21 and CIM-496-Gossypium hirsutum. *Scientific reports*, *11*(1), 20466.

Hao, J. J., Yu, S. X., Dong, Z. D., Fan, S. L., Ma, Q. X., Song, M. Z., & Yu, J. W. (2008). Quantitative inheritance of leaf morphological traits in upland cotton. *The Journal of Agricultural Science*, *146*(5), 561-569.

Hsu, C. Y., Jenkins, J. N., Saha, S., & Ma, D. P. (2005). Transcriptional regulation of the lipid transfer protein gene LTP3 in cotton fibers by a novel MYB protein. *Plant Science*, *168*(1), 167-181.

Iqbal, A., Aslam, S., Ahmed, M., Khan, F., Ali, Q., & Han, S. (2023a). Role of actin dynamics and GhACTIN1 gene in cotton fiber development: A prototypical cell for study. *Genes*, *14*(8), 1642.

Iqbal, A., Aslam, S., Akhtar, S., Ali, Q., Rao, A. Q., & Husnain, T. (2023b). Over-expression of GhACTIN1 under the control of GhSCFP promoter improves cotton fiber and yield. *Scientific Reports*, *13*(1), 18377.

Jakoby, M. J., Falkenhan, D., Mader, M. T., Brininstool, G., Wischnitzki, E., Platz, N., ... & Schnittger, A. (2008). Transcriptional profiling of mature Arabidopsis trichomes reveals that NOECK encodes the MIXTA-like transcriptional regulator MYB106. *Plant Physiology*, *148*(3), 1583-1602.

Jiang, C. X., Chee, P. W., Draye, X., Morrell, P. L., Smith, C. W., & Paterson, A. H. (2000). Multilocus interactions restrict gene introgression in interspecific populations of polyploid Gossypium (cotton). *Evolution*, *54*(3), 798-814.

Kohel, R. J., Richmond, T. R., & Lewis, C. F. (1974). Genetics of Flowering Response in Cotton. VI. Flowering Behavior of Gossypium hirsutum L. and G. barbadense L. Hybrids 1. *Crop science*, *14*(5), 696-699.

Lacape, J. M., Dessauw, D., Rajab, M., Noyer, J. L., & Hau, B. (2007). Microsatellite diversity in tetraploid Gossypium germplasm: assembling a highly informative genotyping set of cotton SSRs. *Molecular Breeding*, *19*, 45-58.

Lacape, J. M., Gawrysiak, G., Cao, T. V., Viot, C., Llewellyn, D., Liu, S., ... & Giband, M. (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.

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.

Lan, T. H., Cook, C. G., & Paterson, A. H. (1999). Identification of a RAPD marker linked to a male fertility restoration gene in cotton (Gossypium hirsutum L.). *J Agric genomics*, *4*, 1-5.

Liu, S., Cantrell, R. G., McCarty Jr, J. C., & Stewart, J. M. (2000). Simple sequence repeat–based assessment of genetic diversity in cotton race stock accessions. *Crop Science*, *40*(5), 1459-1469.

Ma, X. X., Zhou, B. L., Lü, Y. H., Guo, W. Z., & Zhang, T. Z. (2008). Simple sequence repeat genetic linkage maps of a-genome diploid cotton (Gossypium arboreum). *Journal of Integrative Plant Biology*, *50*(4), 491-502.

May, O. L. (2001). Registration of PD 94045 germplasm line of upland cotton. *Crop science*, *41*(1), 279-279.

May, O. L., Bowman, D. T., & Calhoun, D. S. (1995). Genetic diversity of US upland cotton cultivars released between 1980 and 1990. *Crop Science*, *35*(6), 1570-1574.

May, O. L., Chee, P. W., & Sakhanokho, H. (2004). Registration of GA98033 upland cotton germplasm line. *Crop science*, *44*(6), 2278-2280.

McCarty Jr, J. C., & Jenkins, J. N. (1992). Cotton germplasm: characteristics of 79 day-neutral primitive race accessions. *Technical bulletin-Mississippi Agricultural and Forestry Experiment Station (USA)*.

McCarty Jr, J. C., & Jenkins, J. N. (1992). Cotton germplasm: characteristics of 79 day-neutral primitive race accessions. *Technical bulletin-Mississippi Agricultural and Forestry Experiment Station (USA)*.

Mumtaz, H. (2007). *Identification of structural and functional genomic markers for fiber quality traits in cotton using interspecific population (G. hirsutum x G. barbadense)* (Doctoral dissertation, MPhil Thesis, QA Univ Islamabad Pakistan).

Perlak, F. J., Oppenhuizen, M., Gustafson, K., Voth, R., Sivasupramaniam, S., Heering, D., ... & Roberts, J. K. (2001). Development and commercial use of Bollgard® cotton in the USA–early promises versus today's reality. *The Plant Journal*, *27*(6), 489-501.

Pflieger, S., Lefebvre, V., & Causse, M. (2001). The candidate gene approach in plant genetics: a review. *Molecular breeding*, *7*(4), 275-291.

Plett, J. M., Mathur, J., & Regan, S. (2009). Ethylene receptor ETR2 controls trichome branching by regulating microtubule assembly in Arabidopsis thaliana. *Journal of experimental botany*, *60*(13), 3923-3933.

Puspito, A. N., Rao, A. Q., Hafeez, M. N., Iqbal, M. S., Bajwa, K. S., Ali, Q., ... & Husnain, T.

(2015). Transformation and evaluation of Cry1Ac+ Cry2A and GTGene in Gossypium hirsutum L. *Frontiers in plant science*, *6*, 943.

Quisenberry, J. E. (1975). Inheritance of Fiber Properties Among Crosses of Acala and High Plains Cultivars of Upland Cotton 1. *Crop Science*, *15*(2), 202-204.

Rahman, M., Hussain, D., & Zafar, Y. (2002). Estimation of genetic divergence among elite cotton cultivars–genotypes by DNA fingerprinting technology. *Crop Science*, *42*(6), 2137-2144.

Rahman, M., Hussain, D., Malik, T. A., & Zafar, Y. (2005). Genetics of resistance to cotton leaf curl disease in Gossypium hirsutum. *Plant pathology*, *54*(6), 764-772.

Rahman, M., Yasmin, T., Tabbasam, N., Ullah, I., Asif, M., & Zafar, Y. (2008). Studying the extent of genetic diversity among Gossypium arboreum L. genotypes/cultivars using DNA fingerprinting. *Genetic Resources and Crop Evolution*, *55*, 331-339.

Rehman, I., Aftab, B., Bilal, S. M., Rashid, B., Ali, Q., Umair, M. M., ... & Husnain, T. (2017). Gene expression in response to Cotton Leaf Curl Virus Infection In Gossypium hirsutum under variable environmental conditions. *Genetika*, *49*(3), 1115-1126.

Rong, J., Pierce, G. J., Waghmare, V. N., Rogers, C. J., Desai, A., Chee, P. W., ... & Paterson, A. H. (2005). Genetic mapping and comparative analysis of seven mutants related to seed fiber development in cotton. *Theoretical and applied genetics*, *111*, 1137-1146.

Rosenow, D. T., Quisenberry, J. E., Wendt, C. W., & Clark, L. E. (1983). Drought tolerant sorghum and cotton germplasm. *Agricultural Water Management*, *7*(1-3), 207-222.

Saha, S., Raska, D. A., & Stelly, D. M. (2006). Upland Cotton (Gossypium hirsutum L.) x Hawaiian Cotton (G. tomentosum Nutt. Ex. Seem.) F1 hybrid hypoaneuploid chromosome substitution series.

Salentijn, E. M., Pereira, A., Angenent, G. C., van der Linden, C. G., Krens, F., Smulders, M. J., & Vosman, B. (2007). Plant translational genomics: from model species to crops. *Molecular Breeding*, *20*, 1-13.

SARANGA, Y. E., Jiang, C. X., Wright, R. J., Yakir, D., & Paterson, A. H. (2004). Genetic dissection of cotton physiological responses to arid conditions and their inter-relationships with productivity. *Plant, Cell & Environment*, *27*(3), 263-277.

Saranga, Y., Menz, M., Jiang, C. X., Wright, R. J., Yakir, D., & Paterson, A. H. (2001). Genomic dissection of genotype× environment interactions conferring adaptation of cotton to arid conditions. *Genome research*, *11*(12), 1988-1995.

Shaheen, T., Tabbasam, N., Iqbal, M. A., Ashraf, M., Zafar, Y., & Paterson, A. H. (2012). Cotton genetic resources. A review. *Agronomy for sustainable development*, *32*, 419-432.

Song, X. L., & Zhang, T. Z. (2007). Identification of quantitative trait loci controlling seed physical and nutrient traits in cotton. *Seed Science Research*, *17*(4), 243-251.

SONG, X. L., GUO, W. Z., HAN, Z. G., & ZHANG, T. Z. (2005). Quantitative trait loci mapping of leaf morphological traits and chlorophyll content in cultivated tetraploid cotton. *Journal of Integrative Plant Biology*, *47*(11), 1382-1390.

Tester, M., & Davenport, R. (2003). Na+ tolerance and Na+ transport in higher plants. *Annals of botany*, *91*(5), 503-527.

Ullah, I. (2009). Molecular genetic studies for drought tolerance in cotton. *Ph. D thesis Quaid-i-Azam University*.

Van Esbroeck, G. A., Bowman, D. T., Calhoun, D. S., & May, O. L. (1998). Changes in the genetic diversity of cotton in the USA from 1970 to 1995. *Crop Science*, *38*(1), 33-37.

Wada, T., Kurata, T., Tominaga, R., Koshino-Kimura, Y., Tachibana, T., Goto, K., ... & Okada, K. (2002). Role of a positive regulator of root hair development, CAPRICE, in Arabidopsis root epidermal cell differentiation.

Waghmare, V. N., Rong, J., Rogers, C. J., Pierce, G. J., Wendel, J. F., & Paterson, A. H. (2005). Genetic mapping of a cross between Gossypium hirsutum (cotton) and the Hawaiian endemic, Gossypium tomentosum. *Theoretical and Applied Genetics*, *111*, 665-676.

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., ... & Gray, J. C. (1999). The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *The Plant Cell*, *11*(7), 1337-1349.

Wang, B., Wu, Y., Guo, W., Zhu, X., Huang, N., & Zhang, T. (2007). QTL analysis and epistasis effects dissection of fiber qualities in an elite cotton hybrid grown in second generation. *Crop science*, *47*(4), 1384-1392.

Wendel, J. F. (1989). New World tetraploid cottons contain Old World cytoplasm. *Proceedings of the National Academy of Sciences*, *86*(11), 4132-4136.

Wright, R. J., Thaxton, P. M., El-Zik, K. M., & Paterson, A. H. (1998). D-subgenome bias of Xcm resistance genes in tetraploid Gossypium (cotton) suggests that polyploid formation has created novel avenues for evolution. *Genetics*, *149*(4), 1987-1996.

Yu, J., Yu, S., Lu, C., Wang, W., Fan, S., Song, M., ... & Zhang, J. (2007). High-density linkage map of cultivated allotetraploid cotton based on SSR,

TRAP, SRAP and AFLP markers. *Journal of Integrative Plant Biology*, *49*(5), 716-724.

Zafar, M. M., Mustafa, G., Shoukat, F., Idrees, A., Ali, A., Sharif, F., ... & Li, F. (2022). Heterologous expression of cry3Bb1 and cry3 genes for enhanced resistance against insect pests in cotton. *Scientific Reports*, *12*(1), 10878.

Zafar, Y., & Paterson, A. H. (2009). Gossypium DNA markers: types, numbers, and uses. *Genetics and genomics of cotton*, 101-139.

ZAFAR, Y., ASIF, M., KAUSAR, A., RIAZ, S., NIAZ, M., WAHID, A., & ABBAS, S. Q. (2009). Development of genetic linkage map of leaf red colour in cotton (Gossypium hirsutum) using DNA markers. *Pak. J. Bot*, *41*(3), 1127-1136.

Zhu, W., Liu, K., & Wang, X. D. (2008). Heterosis in yield, fiber quality, and photosynthesis of okra leaf oriented hybrid cotton (Gossypium hirsutum L.). *Euphytica*, *164*, 283-291.

Declarations

Data Availability statement

All data generated or analyzed during the study are included in the manuscript.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Funding

Not applicable

Conflict of Interest

Regarding conflicts of interest, the authors state that their research was carried out independently without any affiliations or financial ties that could raise concerns about biases.