Mehboob-ur-Rahman
Yusuf Zafar
Tianzhen Zhang *Editors*

Cotton Precision Breeding

Springer

Cotton Precision Breeding

Mehboob-ur-Rahman • Yusuf Zafar
Tianzhen Zhang
Editors

Cotton Precision Breeding

Springer

Editors
Mehboob-ur-Rahman
Agricultural Biotechnology Division
National Institute for Biotechnology &
Genetic Engineering (NIBGE) College
Faisalabad, Pakistan

Pakistan Institute for Engineering and
Applied Science (PIEAS)
Islamabad, Pakistan

Tianzhen Zhang
College of Agriculture and Biotechnology
Zhejiang University
Hanzhou, China

Yusuf Zafar
Pakistan Agriculture Research Council
Islamabad, Pakistan

ISBN 978-3-030-64503-8 ISBN 978-3-030-64504-5 (eBook)
https://doi.org/10.1007/978-3-030-64504-5

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Cotton is cultivated for harvesting natural lint fiber that sustains textile industry worldwide, generating revenue of around US$ 899 billion. Economies of few countries are heavily dependent on cotton production. In the present scenario, harvesting sustained cotton yield has become a major issue owing to changing environments, high pressure of expanding human population, and high cost of inputs including pesticides, fertilizers, water, etc. The availability of manmade fibers at low price also competes out the natural fiber in the international market. Resultantly, it depresses the cotton farm profitability that may cause significant reduction in cotton growing area. This phenomenon has been witnessed in many regions, including Pakistan, Argentina, Brazil, and Paraguay.

The sustainable enhancement in cotton production can be achieved largely by developing cotton varieties that can mitigate the fast evolving resistance in insect pests and pathogens and instable prevailing environments. Cotton cultivars with brilliant genetics can be bred in the shortest possible time by the application of genomic knowledge as well as genetic transformation assay. The revolution in genomic science together with parallel amazing innovations in bioinformatics and transformation technologies particularly witnessed after 2010s has made us to initiate "Precision Breeding in Cotton."

In this particular book, efforts are made to summarize the key issues depressing cotton production as well as means to combat these challenges for awareness and better utilization of natural fiber cotton. Five different research areas including cotton germplasm, genomic resources and their evolution, genetically modified (GE) cotton developed through transformation assay, genome editing, skeptical views about GE cotton and cotton production beyond 2030 have been discussed.

Breeding for improved cotton cultivars through avoiding the repetitive use of few highly adaptive germplasm—stagnating cotton production worldwide, has got the top priority for bringing new genes under cultivation. Diverse untapped genetic resource (under-exploited cotton species) is not only an asset for breeders to enrich the cultivated gene pool with new genes and/or alleles but also for molecular biologist for unraveling various genetic mechanisms.

In this book, extraordinary natural diversity exists in *Gossypium* for various traits, including fiber morphology and stress tolerance, and also to agronomic traits have been described. It has been explained that high resilience and yield potential in *G. hirsutum* than that of *G. barbadense* is the outcome of divergent evolution of several genes that was the major driving force for imparting this species a wider adaptability. Similarly, within a species, several genetically distinct landraces, accessions, etc. exist. However, within cultivated cotton varieties, the genetic diversity is very narrow as indicated in several studies. It makes the cotton improvement very challenging.

At the moment, cumulatively 70,000 accessions, cultivars, landraces, etc. are available in nine gene banks of different cotton-growing countries, particularly at Fort Collins, USDA, and CIRAD, Montpellier, France. These germplasm resources can be utilized for tapping genes for adding resilience to abiotic and biotic stresses. However, it is expected that several accessions are duplicated, which can be avoided by characterizing the available germplasm resources phenotypically as well as genotypically. For bringing diversity in cotton varieties, exchange of cotton genetic resource among the cotton breeders of different countries is the key to success. Such activities would add synergy to cotton improvement programs. Possible means to explore such kind of untapped genetic resources have been discussed.

The potential of genetic diversity occurs in germplasm and has not been fully utilized owing to several inherent issues, including linkage drags of undesirable traits, requirement of several rounds of backcrosses for recovering the genetic background of recurrent parent genotype, and incompatibility barriers. Generation of enormous amount of genomic data has made it possible to use such genetic resources in breeding programs through designing DNA markers around the genes followed by making selection using these markers in succeeding generation, which can reduce sufficient years for releasing a cotton variety. For example, availability of gap-less genome assemblies of all cultivated cotton species and their progenitor species and high-density genetic maps developed using interspecific and intraspecific populations are available, which can be used in warranting marker-assisted selection. Several articles on cotton genome sequencing appeared from 2012 till date. This information was also helpful in knowing the evolutionary course of the genus *Gossypium*, landscape of cotton genome, and also the function of various genes involved in conferring various traits. Also the use of mutagens for inducing mutations by exposing cottonseed with mutagens like EMS and gamma radiations would be instrumental in expanding the allelic diversity at various loci that can buffer the potential epidemics of insect pests and diseases. Utilization of genetic diversity and means to enhance the genetic diversity for improving resistance to diseases as well as to water-limited conditions have been explained in this book

Improvement in fiber traits has remained the top priority area of several cotton-breeding programs. Multiple genes have been identified which confer high fiber traits. These are found in abundance in *G. barbadense* as elucidated by undertaking genome-wide association analysis (GWAS) and re-sequencing of cotton germplasm. Genetic mechanism for fiber traits development has also been explained. Similarly, genetics of colored lint fiber has been discussed in one of the chapters and also the scope of breeding for developing cotton varieties producing colored natural fiber.

Another major section of this book is related to the development of GE cotton and its cultivation in the field. First generation of GE cotton has offered protection from bollworms and herbicides, which together depreciate the cost of production across all the cotton-growing countries. Limited number of genes of Cry series and genes conferring resistance to glyphosate have been commercialized in cotton; highlighting the need for introducing new genes especially derived from plant sources, if possible, would have high chances of acceptance among the community. Thus, the possibilities of exploring new technologies and genes for building more walls of protection to stresses have also been discussed. This gives a birth of new generation of GE cotton. For example, pyramiding of Bt genes with RNA interference (RNAi) has been used to delay the evolution of resistance by chewing insect pests of cotton. Similarly, RNAi method was deployed to develop to generate resilience to whitefly and diseases including Verticillium wilt. Also the role of RNAi in improving agronomic traits and resistance to other stresses has been discussed in this book. Scope of third generation of GE cotton, aiming for producing huge quantity of industrial products cheaply, has also been explained. Acquiring high expression of pharmaceuticals products can be achieved by introducing genes in plastid. Also the use of CRISPR technology for developing third generation GE cotton has several advantages, including development of transgene-free cotton and gossypol-free cottonseed by silencing the genes conferring gossypols in seed. Major advantage of this technology is that the new cultivars can be evolved without introgression of foreign gene; hence, the technology will be acceptable to countries having skeptical views about the GE technology. The success of second and third generation of GE cotton has not yet been demonstrated in the field, however, with the given knowledge and high-tech genomic assays, in near future, when second and third generation of GE cotton will be cultivating at farmer's field. In the second last chapter, skeptical views as well as counter arguments pertaining to GE cotton have been compiled. Discussion on pros and cons of using GE cotton products and by-products will help end-user to get benefit of the modern technologies aiming at improving the quality of life.

Hence, the scope of harvesting sustained cotton production beyond 2030 has been discussed. Many well-known scientists from the cotton research community have shared wisdom, knowledge, views, and application of the genomic knowledge that would lay a firm foundation for initiating precision breeding in cotton. Thus, the book comprehensively covering such kind of information is the need of time.

Faisalabad, Pakistan — Mehboob-ur-Rahman
Islamabad, Pakistan — Yusuf Zafar
Hanzhou, China — Tianzhen Zhang

Acknowledgments

The idea for writing this book was conceived upon a scientific visit to China through a collaborative project “Mining of Elite Genes for High Yield, Super Fiber Qualities and Heat Tolerance and Their Usages in Pakistan and China Cotton Cultivars”—Project No. PSF/BSFC-AGR/P-NIBGE (12).

Contents

Part I
Cotton Breeding and Genomics

Chapter 1
Historical Perspectives: From Conventional to Precision Breeding in Cotton

Mehboob-ur-Rahman, Sana Zulfiqar, Abid Mahmood, Yusuf Zafar, and Tianzhen Zhang

1.1 Historical Perspective

Use of natural fibers for making fabrics is very prehistoric that is dated back to 36,000 BP (Balter 2009; Kvavadze et al. 2009). Natural fibers, made of cellulosic components, can be derived from several plant species including hemp, cotton, jute, flax, sisal, etc., while the animal fibers (wool and silk) are made of proteins (Hansen and Bjorkman 1998). Synthetic fiber, another major type, is made from petroleum products, which contributes around 60% of the total fiber production worldwide. These fibers are rayon, polyester, nylon, acetate, acrylic, spandex, lyocell, etc. (Fig. 1.1). At present, synthetic fibers are considered durable and economically cheaper. Though the debate on the superiority as well as benefit of both types of fibers is going on over the last many years, natural fibers are still preferred by the masses because these are produced naturally and thus considered much safer.

Mehboob-ur-Rahman (✉)
Agricultural Biotechnology Division, National Institute for Biotechnology & Genetic Engineering (NIBGE) College, Faisalabad, Pakistan

Pakistan Institute for Engineering and Applied Science (PIEAS), Islamabad, Pakistan
e-mail: mehboob@nibge.org

S. Zulfiqar
Plant Genomics and Mol. Breeding Lab, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

A. Mahmood
Ayub Agricultural Research Institute (AARI), Faisalabad, Pakistan

Y. Zafar
Pakistan Agriculture Research Council, Islamabad, Pakistan

T. Zhang
College of Agriculture and Biotechnology, Zhejiang University, Hanzhou, Zhejiang Province, China

Mehboob-ur-Rahman et al. (eds.), *Cotton Precision Breeding*,
https://doi.org/10.1007/978-3-030-64504-5_1

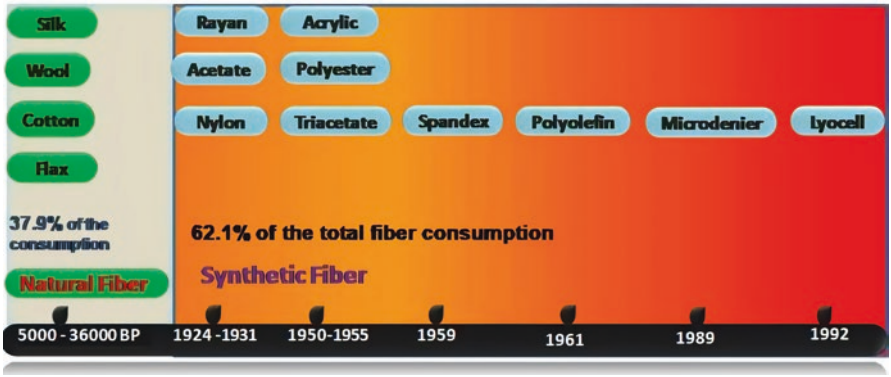

Fig. 1.1 Fibers: an important part of human civilization

Among natural fibers, cotton fiber has been used extensively for making fabrics especially after the onset of industrial revolution in the eighteenth and nineteenth centuries. Later on, semisynthetic fiber, viscose, and synthetic fiber, nylon, were commercialized in 1905 and 1930s, respectively. These fibers particularly nylon captured a major market share of natural fibers. For overcoming this issue, efforts were made for developing cotton varieties with enhanced yield and better lint qualities which can easily be amenable to new textile machinery. All these efforts resulted in the development of improved cotton cultivars by incorporating genes from unadapted cotton germplasm in the mid-twentieth century. Thereafter, discoveries in genomic science paved the way of transferring genes from alien sources into cotton. First genetically engineered (GE) cotton containing *Cry 1Ac* gene was commercialized in 1995 (Rahman et al. 2015). However, over the last many years and even in the present time, the prices of petroleum have been depreciated significantly that may further depress the market share of natural fiber owing to the influx of low-cost synthetic fiber.

1.2 Challenges to Cotton Production

1.2.1 Insect Pest and Diseases

Infestation of insect pests and infection by diseases are the main causes of low yield in cotton. Protection to chewing pests was largely achieved by incorporating *Cry* genes (Rahman et al. 2012). This protection was weakened; resultantly chemicals are sprayed for killing the chewing insect particularly pink bollworm. The potential of minor pests for becoming major pests can be the future challenge, for example, infestation of mealybug and dusky bug in Pakistan. This scenario may also arise in other countries. Similarly, evolution of new strains of pathogens especially new virus strains that cause leaf curl disease, verticillium wilt—also called cotton cancer—and bacterial blight are the major potential threats to cotton production. Cotton leaf curl can spread to cotton-growing countries where whitefly is prevalent

(Rahman et al. 2017). If these two factors are not properly addressed, it may lead to the elimination of cotton cultivation from several cotton-growing regions.

1.2.2 Changing Climate

The gradual increasing trends of heat, drought, unexpected heavy rains and other environmental extremities are negatively impacting the agricultural productivity (Ray et al. 2013; Mills et al. 2018). These fluctuations in environments if occurring together may reduce crop productivity very drastically. Breeding cotton cultivars has been carried out in high-input environments (Rahman 2016); thus the genetic potential of these varieties will not be able to sustain under the adverse environments.

Like many other crops, global warming is a detrimental threat to cotton production worldwide, and its impact partly has been witnessed in the form of high temperature, extensive rainfall, unpredictable climatic adversities, etc. It has been reported that temperature prevailing the crops will continue to increase in the twenty-first century. High temperature retards cotton growth that leads to shedding of small fruiting bodies (square) and reduced boll size; together these factors lessen the lint yields. High temperature also reduces the effectiveness of pesticides. This phenomenon can be more detrimental to those cotton-growing regions where temperature often exceeds 40 °C.

1.2.3 Competitive Ability with Synthetic Fiber

If the present declining trend of petroleum price persists, synthetic fiber can capture the market share of natural fiber in the future. These factors may lead to the depression of the price of natural fiber in the international market; ultimately the profitability of farmers will be reduced. Thus the cotton-growing community can think cultivation of other crops for increasing their farm's profitability. For example, in Pakistan, during the last three years, area under cotton has significantly been reduced by ~30%. The major factor is the high cost of production as well as infestation of pink bollworm.

1.2.4 Yield Losses and Contamination

Yield losses during harvesting time and trash contamination in seed cotton are two major issues which are supposed to be addressed for deriving maximum benefit from cotton. For example, in developing countries cotton is picked manually. In the future, it will be very difficult to harvest cotton crop well in time owing to the mass migration of families from rural to urban areas. Secondly, hesitation of new generation to work in the field is also aggravating the situation. Hence, for harvesting sustained

cotton production, it is important to shift on mechanical picking system like many other advanced countries including USA and Australia which have adopted this system. There is a need to improve the efficiency of mechanical picking system for avoiding the yield losses (at least 10% during picking). Govt department should provide mechanical pickers to farmers at subsidized rates so that harvesting can be carried out well in time. Similarly, contamination issues will be more serious in those cotton-growing countries where varieties have trichomes on various organs including leaves. At maturity when farmers spray defoliants before harvesting, the falling leaves usually stick with seed cotton—thus adding contamination.

1.2.5 Yield Stagnation

Cotton varieties have achieved maximum yield potential by utilizing conventional as well as transgenic approaches. Almost every breeding program is working with very limited genetic diversity (Rahman et al. 2002, 2008; Hu et al. 2019); thus using present breeding approaches, cotton production will not help in breaking yield barriers. Future improvements can be made if some novel genetic resources will be available. These can help in developing cotton varieties with outstanding genetics for combating the challenging environments. Thus improving resilience together with best management practices can ensure harvesting of maximum sustained yield potential with improved lint quality.

1.3 Roadmap for Achieving Sustainable Cotton Production

1.3.1 Genetic Resources

The valuable genetic diversity is present in the available germplasm in various forms including landraces, genotypes, accessions, etc. These germplasm resources can be utilized for tapping genes required for combating abiotic and biotic stresses. Cotton fiber is contributed by four major cultivated cotton species including *G. arboreum*, *G herbaceum*, *G. hirsutum* and *G. barbadense.* One of the major breeding challenges is to broaden the narrow genetic base of the cultivated cotton varieties for adding resilience to environmental factors. For this, new alleles or genes can be introduced from alien species. For the introduction of new alleles or genes from the genus *Gossypium*, it is important to have well-characterized genetic resources. About 70,000 accessions are available in several gene banks of 9 cotton-growing countries (USA, Uzbekistan, India, China, Australia, France, Russia, Brazil, and Pakistan; Fig. 1.2). While in CottonGen database, a total of 20,253 accessions-genotypes-cultivars representing almost all *Gossypium* species have been reported (https://www.cottongen.org/find/germplasm).

The major issue with these gene banks is the preservation of similar accessions. For avoiding these duplicated accessions, it is important to characterize all

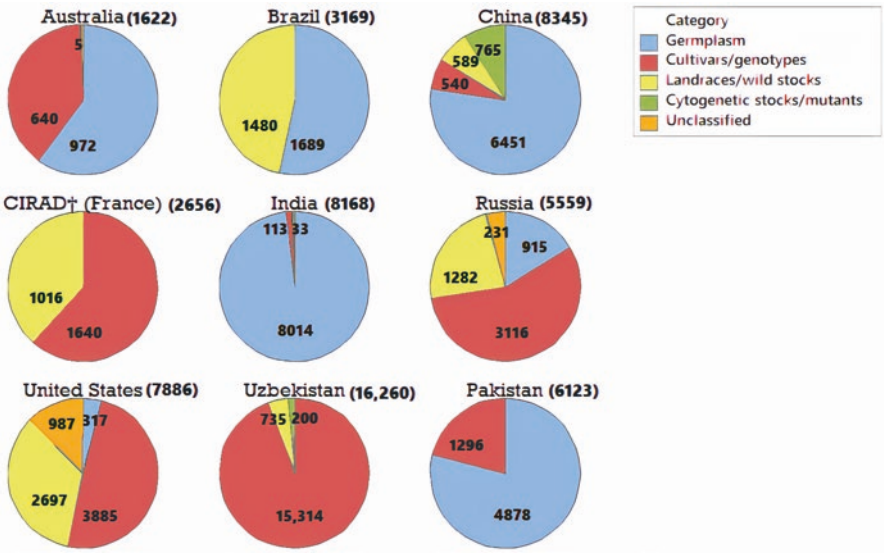

Fig. 1.2 Genetic resources of the genus *Gossypium* in different gene banks across the world

the available germplasm resources phenotypically as well as genotypically (Iqbal et al. 2015). Only the selected accessions (by omitting the duplicated accessions) should be characterized by conducting multilocation trials in several other cotton-growing countries. In this regard, exchange of germplasm among breeders across the cotton-growing countries will remain the most important tool for studying the response of cotton accessions in diverse environments. For example, a big activity of germplasm transfer from the USA to Pakistan (sponsored by the USDA through cotton productivity enhancement project, ID-1198) resulted in the identification of several cotton accessions conferring resistance to cotton leaf curl disease (Rahman and Zafar 2018). Among these, Mac-07 was extensively used in developing cotton cultivars with improved tolerance to the disease. Similarly, another joint activity was initiated [(Pak-China project entitled mining genes for high yield, super fiber, and heat tolerance, PSF/NSFC-AGR/P-NIBGE (12)] for screening germplasm to heat as well as studying fiber traits in various environments. Through this project, heat-tolerant cotton genotypes have been identified for using in future cotton breeding programs. Such activities would add synergy to cotton improvement programs.

For scoring traits of cotton germplasm, it is important to devise a uniform scale for rating or scoring each character so that everyone should speak the same language. It is also important to select leading cultivated cotton varieties and genotypes/accessions with unique traits for doing re-sequencing. It will help in identification of polymorphisms which can be exploited in fabricating unique universal SNP chip as well as issuing passport to each accession and/or cultivar. The SNP chip can be used for initiating targeted breeding, and thus the time for developing new cotton varieties can be reduced. Hence, a substantial amount of funds is

required for undertaking high-throughput phenotyping and genotyping procedures, these will accelerate mapping genes, understanding genetic mechanisms, and downstream breeding (Dempewolf et al. 2014).

1.3.2 Conventional Breeding

Prior to the onset of planned cotton breeding programs, varieties of all the cultivated cotton species were developed through making selections largely by the early farmers. The genetic variations present in these land races or old varieties were used to select superior cotton plants (Rahman et al. 2014). Cotton is largely a self-pollinated crop, and cross pollination by insect was the main source of variations in early bred cotton varieties. Thus the concept of developing homogenous cotton varieties much like wheat and other self-pollinated crops was not possible in those days. At the start of the twentieth century, these variations were exploited through adopting planned hybridization procedures for fixing the desirable traits into one variety. For example, escape from boll weevil infestation in the USA was managed by developing early maturing cotton varieties, thus replaced gradually the cultivated late maturing cotton varieties. The Deltapine cotton varieties were developed by attempting several selections and crossing procedures in 1911 (Poehlman 1987). Wilt resistant cotton plant (cocker 100 wilt), Stoneville type, storm resistant types, etc. were selected from 'Lone Star' type. Similarly, several cotton varieties were developed by selecting plants from the cotton field of a small village 'Acala', Mexico in 1906–1907. Similarly, Pima cotton (*G. barbadense*) varieties were developed in the USA by selecting best cotton plants when grown in Arizona and Southern California in 1903. Heterogeneity was desirable for having persistent hybrid vigor in a variety by bulking seed of cotton progenies showing sufficient uniformity for morphological characteristics, disease and insect resistance, and lint quality (Poehlman 1987).

Like self-pollinated crop varieties, planned hybridization followed by pedigree selection procedure was deployed for developing cotton varieties (Rahman et al. 2014). Similarly, backcross breeding procedure was adopted for bringing new alleles conferring resistance to diseases from untapped genetic resources. Backcrosses were also attempted for bringing male-sterile genes or fertility restoring genes into a variety. Chromosomes can also be transferred into sterile cytoplasms by attempting several rounds of backcrosses.

Other option for enhancing the yield could be the exploitation of hybrid vigor. For instance, hybrid vigor resulted in multifold increase in corn production worldwide. Developing cotton hybrids showing 30% increase in lint production over the standard OP cotton variety may convince farmers to cultivate hybrid cotton. This task is difficult but achievable through exploring the best combiners. Hybrid breeding cotton is handicapped due to the non-availability of reliable genetic resources, lack of mechanical procedures for emasculation and also the efficacy of gametocides. Various genes encoding sterility and restoring fertility were identified, and

efforts were made to produce hybrid cottonseed as successfully demonstrated in corn by adopting the ABR system (Poehlman 1987). However, these genes in cotton were found to be temperature sensitive and thus the potential benefit of this system was not harvested as witnessed in corn. Work on understanding the genetics of hybrid vigor can be the best choice which will help in targeting important genes for amending their expression, thus hybrid vigor can be fixed.

Development of resilience in cotton through conventional breeding can also be helpful for addressing the issues of changing environment. One of the best strategies to combat escalating heat issue is to develop cotton varieties which can tolerate 50 °C day temperature and 30 °C night temperature. In this regard, few traits, for example, boll retention at high temperature is considered as the most appropriate selection criterion. However, such varieties are usually high input demanding. When inputs are scarce, such varieties could not show their yield potential. Thus selection for varieties which can keep on growing even by sacrificing few bolls on first sympodial branches would be the desirable feature for encountering scarce water resources as well as high temperature. Secondly, breeding for small-to-medium sized leaves would allow sun light to reach lower leaves and opened bolls—thus boll rottening can be avoided in rainy season. Canopy of such cotton varieties also facilitates the uniform application of pesticides on even lower leaves of cotton plants.

Sustained cotton production can be achieved by studying the response of genetic resources (parent genotypes and their progenies) under low-input environment. For instance, landraces evolved in low-input regions can be utilized in breeding programs. Deployment of genomic-based selection procedures would be the best choice for transferring genes into the domesticated high-yielding varieties. For this purpose, extensive screening of germplasm would be required for grabbing genotypes exhibiting traits such as resistance to various insect pests and diseases, better nutrient economy, delayed leaf senescence, and yield consistency across the environments. Germplasm originated in different cotton-growing regions can be screened in those countries which are facing the impact of climate change. Segregating populations can also be screened through joint collaboration. Resulting newly developed candidate lines can be tested in different environments for studying their adaptability. This approach is straightforward, and does not require enough resources. Thus, exchange of cotton germplasm will remain the most instrumental tool. DNA markers too can be very instrumental since the trait is not as complex as drought tolerance.

1.3.3 Introgression Breeding

Introgression breeding, transferring of alleles or genes from unadapted genetic resources to adapted genotypes (Hernandez et al. 2020), has played gigantic role in transferring useful genes from different closely related species into the cultivated species including cotton (Rahmat et al. 2019). Before the domestication of tetraploids, exchange of gene transfer between *G. hirsutum* and *G. barbadense* was

reported (Hu et al. 2019). These introgressions facilitated the adoption of a specie in a challenging environment by expanding its genetic diversity. This concept was translated by cotton breeders to transfer genes from other closely related cotton species to the cultivated cotton species at the National Key Laboratory of the Crop Genetic and Germplasm Enhancement, College of Agriculture, Nanjing Agricultural University, Nanjing China under the leadership of Professor Tianzhen Zhang. Pima cotton, *G. barbadense*, is easily crossable with upland cotton *G. hirsutum*. Similarly, diploid cotton species including *G. arboreum*, *G. herbaceum*, *G. raimondii*, etc. have also been used for introducing new genes into the *G. hirsutum* background.

For improving the lint quality, genes related to superior lint quality were introduced in old cotton varieties of *G. hirsutum* through interspecific crosses. The interspecific population may have several transgressive segregants at both extremes for quantitative traits including plant height, leaf shape, number of bolls, boll size, trichome density, lint quality, etc. (Zhang 1993). This phenomenon needs further explorations. There are several stable lines with superior fiber developed after doing extensive backcrossing followed by selections of desirable plants (Ma and Liu 1982; Zhang 1993; Cantrell and Davis 2000; Percy et al. 2005; Gore et al. 2012; Liu et al. 2005).

Some desirable genes from Pima cotton were transferred to Acala cotton for improving its lint quality but success was very limited. Similarly, high-quality lines Pee Dee were developed through interspecific crossing (May 2001). Later on, these lines were used in introducing genes from Acala and Pee Dee lines into cultivated cotton varieties (Bowman and Gutierrez 2003).

Several other genes with qualitative impact were transferred from Pima cotton to upland cotton. These alleles or genes are related to traits such as glandless (Yuan et al. 2000), sub-okra leaf shape (Zhang 1993, 2011) and resistance to bacterial blight (Percy and Kohel 1999), verticillium wilt (Wilhelm et al. 1985; Zhang et al. 2012), thrips (Zhang et al. 2013a, b), and spider mite (Zhang et al. 1992, 1993). Introgressed lines conferring tolerance to drought and salinity were also developed (Zhang and Hughs 2012; Tiwari et al. 2013).

The genetic diversity found in wild tetraploids was also utilized for enriching the genome of cultivated cotton species. Crossing between tetraploid species was easy owing to the equal number of chromosomes. The brown fiber producing tetraploid *G. tomentosum* ($2n = 4\times = 52$) exhibited resistance to leaf hopper and thrips (Jayaraj and Palaniswamy 2005). The tetraploid *G. mustelinum* produced terpenoid aldehydes in leaves that conferred resistance to insect pests (Altaf et al. 1997).

Similarly, genes from diploid cotton species were transferred for improving resilience (Rahmat et al. 2019) as well as fiber traits. For instance, *G. longicalyx* ($2n = 2\times = 26$, F-genome), native to Africa, is resistant to reniform nematode (Yik and Birchfield 1984), and also has genes for fiber quality (Weaver et al. 2013). Similarly, *G. armourianum* ($2n = 2\times = 26$, $D_{2\text{-}1}$ genome) is resistant to whitefly (Jayaraj and Palaniswamy 2005). The *G. herbaceum* and *G. arboreum* (both are diploid cultivated species) are resistant to cotton leaf curl disease (Rahman et al. 2005). Several methods were used to hybridize these two species. In few studies,

chromosomes of these diploid species were doubled ($2n = 4\times$) followed by crossing with *G. hirsutum* to have F_1 for making backcrosses with *G. hirsutum* (Nazeer et al. 2014; Rahmat et al. 2019). Development of synthetic tri-species hybrids have been suggested to introgress genes from diploid species to cultivated tetraploids (Mergeai 2006; Robinson et al. 2007).

The process of introgression breeding has claimed limited success in improving cotton owing to linkage drag of unwanted genes, hybrid breakdown, instability and selective elimination of desirable genes during selfing (Zhang et al. 2014). In the future, DNA markers can help in overcoming the issues of linkage drags as well as lengthy breeding procedures in developing stable lines. Since the genome of many important cotton species have been sequenced, thus designing new DNA markers around the loci involved in conferring high lint quality trait will help in reviving the introgression breeding work.

1.3.4 Mutation Breeding

Use of mutagens for expanding the genetic diversity of cultivated cotton resulted in the development of useful germplasm as well as crop varieties. Mutagens can create new traits or enhance the expression of already present genes. In total, ~3308 mutant crop varieties of 200 crop species have been developed worldwide (https://mvd.iaea.org/). Majority of these mutant varieties were released in China, India, and Japan. First mutant cotton variety was released in the early 1960s, and later on, 48 mutant cotton cultivars were released. Out of these, almost 25% were released in Pakistan by exposing the genetic material with gamma radiations.

By the 1990s, mutation detection assays existed to enable the development of reverse-genetics with point mutations, a process known as targeted induced localized lesion in genomes (TILLING). In the present time, next-generation sequencing assays have made it possible to discover several mutant alleles in one experiment. By developing large M_2 populations, one can identify thousands of mutant alleles in each population. Thereafter, only the selected mutant plants containing maximum number of desirable mutations can be retained for performing phenotypic characterization. Thus the available genetic information can further help in reducing the number of desirable mutants which can be used in breeding programs. The typical mutants can be traced in segregating generations using DNA markers by synthesizing primers around the gene or allele of interest.

Complex traits can also be improved to some extent by exposing the genetic material with chemical mutagens (ethyl-methane sulfonate, *N*-ethyl-*N*-nitrosourea, etc.). However, the role of physical mutagens is marginal for improving the complex traits. For example, in multiple studies, significant improvements in ginning outturn have been reported (Patel et al. 2014, 2016; Aslam et al. 2017; Witt et al. 2018). Key to success is to grow maximum number of mutants in M_2 population for identifying very few mutants containing the desirable traits.

1.3.5 Genomic Resources

The genomic science is heavily dependent upon the availability of genetic resources followed by associating phenotypic diversity with the sequence diversity. Extensive re-sequencing will be required for identifying the natural variations occurred in a particular species. Once the genes are identified, these can be used for introducing in cultivated cotton varieties using transgenic approaches or can be directly mutated for altering their expression using targeted breeding approaches like RNAi and CRISPR. Before deploying these technologies, consensus for cracking the genome of cotton specie was made by a group of scientists for facilitating the process of improving and understanding the genetic circuits of various traits.

The idea for sequencing the most simple cotton genome (*G. raimondii*) was discussed in 2006 by a group of cotton scientists led by Prof. Andrew H Paterson, University of Georgia, USA (Chen et al. 2007). Finally, two papers on D-genome sequencing were published in 2012 by two different research groups (Paterson et al. 2012; Wang et al. 2012). Two years later, one study on sequencing *G. arboreum* (A-genome) was published (Li et al. 2014). Thereafter, two papers on sequencing genome of *G. hirsutum* (known for high yield) were published independently by two different research groups in 2015 (Li et al. 2015; Zhang et al. 2015). In the same year, two studies on sequencing the genome of *G. barbadense* (known for high-quality lint) were published by two different consortia (Liu et al. 2015; Yuan et al. 2016). However, the previous genome assemblies of tetraploid cottons were not contiguous; for example, intergenic regions were represented poorly—rendering a lot of information inaccessible to the researchers. Construction of a contagious genome assembly was really a mammoth task because of their large genome size (~2.6 Gb, near to the human genome size).

More recently, genomes of *G. hirsutum* and *G. barbadense* were sequenced for reducing gaps in the reported genome assemblies by involving several assays including single-molecule real-time sequencing, BioNano optical mapping, and high-throughput chromosome conformation capture. Significant improvement in contiguity was achieved particularly in regions containing a lot of repeats, for example, centromeric regions. Massive genome rearrangements particularly inversions in para-/pericentric regions of 14 chromosomes occurred after polyploidization have been reported. In total, 13 QTLs conferring high fiber quality trait have been identified. These regions can be exploited in future cotton improvement programs (Wang et al. 2019).

Same year, Zhang and colleagues (Hu et al. 2019) adopted a series of procedures and approaches for generating a huge number of sequence reads (800 Gb, >330×)—sufficient to differentiate errors from the actual sequence, which were used to construct assemblies of *G. hirsutum* and *G. barbadense* by merging the corrected scaffolds (long stretch of DNA, often reaching to the length of a chromosome). Order and orientation of the scaffolds were verified by three-dimensional proximity information and ultradense genetic map—constructed by placing ~6.1 million SNPs. Hence, several long-awaited puzzles including evolutionary dynamics of

cotton genome and functional dynamics of genes after combing in one nucleus and getting genomic insight of the traits, were resolved. In total, 2.22 gigabases—91.4% of the total estimated genome size of *G. barbadense* var. Hai7124—was assembled that showed 47- and 90-fold increased contiguity over the two previously poorly assembled genomes (Liu et al. 2015; Yuan et al. 2016), and major chunk (98.2%) was assigned into 26 chromosomes. Similarly, 2.30 Gb of *G. hirsutum* var. TM-1 was assembled, of this 97.4% was organized into 26 chromosomes, and impressively the contiguity of the assembly was increased over 10-fold to one sequence assembly (Zhang et al. 2015) and 20-fold over the second published genome assembly (Li et al. 2015). These assemblies can be used as reference for comparing the sequences of other genotypes of these species as well as with some other species of the genus *Gossypium* or even distantly related plant species (Hu et al. 2019).

Centromere of each chromosome composed of several hundreds of kilobases of repetitive elements—keep on expanding/contracting or giving birth of new sequences at remarkably high rate. This feature makes the mapping of centromere very difficult even if the resolution of genome assembly is very high. This issue was resolved by developing a contiguous assembly of each species, and the distorted position of centromeres on each chromosome reported earlier was painted very precisely. For example, the average length of centromeric regions was narrowed down to 270 Kb in Hai7124 and 385 kb in TM-1. These centromeres happened to be localized in the same corresponding regions of the corresponding chromosomes, further authenticating the high quality of the newly assembled genomes (Hu et al. 2019).

Relatively more number of genes were predicted for *G. barbadense* var. Hai7124 than that of the *G. hirsutum* var. TM-1 (75,071 versus 72,761), and most of these showed a clear demarcation of exon-intron boundaries. Majority of these genes were found in duplicated copy due to their ploidy nature. A huge portion of both sequenced genomes comprised of transposable elements (TE) (1460.46 Mb in TM-1 versus 1374.61 Mb in Hai7124), which is doubled in A subgenomes compared to the D subgenome. These were inserted during the three whole genome duplications occurred each around 15, 26, and 48 MYA, well before the synthesis of allotetraploids (Zhang et al. 2015). Some empirical evidences, for example, high colinearity and conserved gene order between TM-1 and Hai7124, suggested that both allotetraploids diverged from a common parent around 0.4–0.6 MYA. The subgenomes of allotetraploid evolved much faster than their diploid progenitors owing to the frequent exchanges of chromosome blocks occurred between the two subgenomes. Relatively a large number of lost, disrupted, and positively selected genes were found in A subgenome, showing the signatures of strong selection pressure exerted during the evolution of these allotetraploids. Deep sequencing of several accessions of *G. barbadense* and *G. hirsutum* suggested that both species evolved very rapidly, and introgression of large genomic regions was found in all *G. barbadense* accessions particularly expressing extra-long staple, confirming that *G. barbadense* was originated in the Northwestern Peru/SW Ecuador region (Westengen et al. 2005). These introgressions helped in the adaptation of domesticated cottons to different environments.

Several other genomic factors including genes with structure variations (10,633 in number) which are responsible for creating novel traits in a particular species have been characterized, e.g., 2-bp deletion in *WLIM1a* of Hai7124. Then authors probed several presence/absence variations (PAVs) which were found unevenly throughout the genomes of TM-1 and Hai7124. Such variations led to generate several species-specific traits, for example, extra-long staple in *G. barbadense* and high yield potential in *G. hirsutum*; both are remarkable features that facilitated their adaptation and domestication (Hu et al. 2019).

Producing quality lint/fiber together with high productivity has remained the central part of cotton breeding programs; however, both traits are difficult to bring into one cultivar due to the complex genetics of long fiber (ESL). More than 50% of the genes (~45,000) contribute in shaping the fiber phenotype. It has also been shown that sucrose transporter (*GbTST1*), Na^+/H^+ antiporter (*GbNHX1*), aluminum-activated malate transporter (*GbALMT*), and vacuole-localized vacuolar invertase are responsible for ELS in Hai7124. Secondly preferential expansion of few gene families in Hai7124, for example, ADP ribosylation factor (ARF) GTPase family, also plays a role in conferring long but strong fiber in the ELS cotton. Similarly, high resilience and yield potential of TM-1 are the outcomes of divergent evolution (after duplication) experienced by several genes after domestication, thus facilitating the adaptation in new environments (Hu et al. 2019). This information can be used for initiating targeted breeding programs for improving high-quality traits in cotton.

More recently, Chen and colleagues (Chen et al. 2020) reported high-quality genome assemblies of five allotetraploid species including *G. hirsutum* $(AD)_1$, *G. barbadense* $(AD)_2$, *G. tomentosum* $(AD)_3$, *G. mustelinum* $(AD)_4$, and *G. darwinii* $(AD)_5$. Relatively large genome size (2.305 Gb) was reported for *G. hirsutum* than the previous studies. High-quality genome assemblies of *G. tomentosum* $(AD)_3$, *G. mustelinum* $(AD)_4$, and *G. darwinii* $(AD)_5$ with 74,699, 78,338, and 78,303 genes, respectively, were reported. Almost, 99% for *G. mustelinum*, 99.2% for *G. tomentosum*, and 99.1% for *G. darwinii* genome were placed in all chromosomes. However, the genome coverage in chromosomes for *G. hirsutum* (98.9%) and *G. barbadense* (97.0%) was less than that of the three wild tetraploid species (Chen et al. 2020). These findings are useful for undertaking future research on designing new markers as well as editing the genome with upcoming new technologies.

1.3.6 New Breeding Procedures

During the last few years, new breeding procedures have been developed to expedite the breeding process.

1.3.6.1 Marker-Assisted Breeding

Cotton varieties with outstanding genetics together with the application of best management practices can ensure harvesting of maximum yield with improved lint quality. Future yield enhancement with improved lint quality is expected to derive from cultivating the resilient cotton varieties. Thus the role of genetics of a cotton variety will be more important particularly for helping the resource-poor cotton-farming communities.

Several DNA markers including restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs), etc. have been used extensively in cotton (Shaheen et al. 2006, 2016; Rahman et al. 2009). However, SSRs and SNPs are easy to use, abundantly present, amenable to automation, and codominant in expression; all these qualities make them the markers of choice. There are several SNP chips available which can be used for the identification of SNPs in cotton. For instance, a CottonSNP63 (Hulse-Kemp et al. 2015), CottonSNP80K (Cai et al. 2017), etc. have been fabricated for developing high-density genetic maps.

DNA markers have been used to detect QTLs for several agronomic traits (Kalivas et al. 2011; Li et al. 2016), tolerance to drought and salinity (Saeed et al. 2011; Shaheen et al. 2013; Jia et al. 2014; Zhao et al. 2016), and yield and lint quality (Zhang et al. 2011; Cai et al. 2014; Nie et al. 2016; Iqbal and Rahman 2017; Wang et al. 2020). The generated information is being used in the identification of DNA markers and their utility in selection of cotton plants from segregating populations.

The lint-to-seed ratio beyond 50% is very much important to improve per unit yield. Natural genetic variations are present in a few accessions of *G. hirsutum*. A few genotypes exceed 50% ginning outturn which can be used to improve the lint-to-seed ratio. In most countries lint recovery ranges from 30 to 45%. This trait is very complex, but can be tailored using DNA markers otherwise difficult to improve using conventional breeding approaches (Iqbal and Rahman 2017; Feng et al. 2019). These markers can be deployed in interspecific and intraspecific populations for screening plants containing the desirable loci. For developing such populations, one adapted cotton cultivar which can be crossed with the obsolete varieties, accessions, and even wild species.

Deployment of genomic-based selection procedures for making selections (foreground and background) would be the best choice for transferring genes into the domesticated high-yielding varieties. Genomic selection procedure can be opted for improving resistance to insect pests and diseases, better nutrient economy, delayed leaf senescence, and yield consistency across the environments. This approach is straightforward, and does not require enough resources.

1.3.6.2 Development of GM Cotton

Resistance to insect pests in cotton genome has been developed using *cry* genes excised from a soil bacterium—success has been demonstrated (Rahman et al. 2012). Now the resistance conferred by these genes has been broken down in different parts of the world (Seetharaman 2018). The potential of converting minor pests into major pests is another threat to cotton sustainability. For example, before the cultivation of *Bt* cotton in Pakistan, mealybug and dusky bug remained unnoticed on cotton—indirectly controlled by the application of insecticides used to kill lepidopteron pests. Now farmers are supposed to kill these newly emerged insect pests by the application of specific insecticide for this particular pest. Such a scenario can emerge in other cotton-growing countries. Improving genetics of cotton by adding a range of novel genes derived from other alien sources under the tissue-specific promoters would be very handy in combating the insect pests more effectively. For example, efforts toward the identification of new genes from other sources such as *Hvt* gene derived from a spider were tested in cotton for studying its response to chewing insect pests, and were found encouraging (Khan et al. 2006). Similarly, several other genes encoding phytohormones (such as jasmonates) which can add in defense to insect herbivory have recently been characterized, which have the potential for combating bollworms in the future.

The challenge of changing climate can also be addressed using transgenic approaches. For example, tolerance to drought can also be enhanced by deploying genes from other organisms including distantly related plants growing under harsh climatic conditions. Efforts are underway; for example, drought-tolerant genes including *TsVP* and H^+-*PPase*-coding gene derived from *Thellungiella halophile* were introduced in cotton that resulted in improved shoot and root growth than that of their wild type (Lv et al. 2009). Also genes and/or their transcription factors (DREBs, ERFs, ZIP, WRKY, etc.) derived from other plant species (Dou et al. 2014; Zhou et al. 2015; Wang et al. 2016a, b) can be characterized followed by their introduction in cotton for improving resilience to abiotic stresses.

Cotton production can be enhanced by introducing genes using transgenic approaches. For example, *ScALDH21* taken from *Syntrichia caninervis* exhibited greater plant height, larger bolls, and greater fiber yield in cotton (Zhang et al. 2013a, b; Khan et al. 2018). For improving complex traits substantially, for instance cotton yield, it is important to pyramid several such genes (even from other species).

Cotton has been remained the target for improving its resilience against biotic and abiotic stress and also to the herbicide. Efforts are also being made to reduce anti-nutrient contents of cotton such as gossypol. First generation of GM cotton delivered protection to insect pests as well as herbicide; however, the insertion of transgene was random. Secondly, resistance generated by the insertion of *Bt* genes was weakened. Later on, pyramiding of Bt genes with RNA interference (RNAi) against *Helicoverpa armigera* has been used to delay the evolution of resistance by chewing insect pests of cotton (Ni et al. 2017), often referred to as second generation of GM cotton. Resistance to whitefly, another destructive pest in Indian subcontinent, has been demonstrated by targeting specific genes of whitefly using RNAi

assays (Malik et al. 2016; Raza et al. 2016). Similarly, resistance to cotton diseases including verticillium wilt has been expressed using RNAi and virus-induced gene silencing (VIGS) (Wang et al. 2017).

Similarly, RNAi was used to silent genes which controls gossypol formation in seed while retaining these gossypols in all other plant organs for avoiding predation (Wedegaertner and Rathore 2015). For expressing the Bt protein in all plant tissues except seed, green tissue-specific promoter PNZIP was used in specific organs, and it has shown the expression of *Cry9C* gene in all vegetative parts except 100-time lower in reproductive organs including pollen, petals, and developing cotton seed (Wang et al. 2016a, b). RNAi was also used in improving several agronomic traits including fiber as well as resistance to stresses (Wang et al. 2016a, b) and response to stress (Zahid et al. 2016). However, these results are not demonstrated yet at farmer's field. Genes for improving fiber quality can be introduced by excising from *Calotropis procera* that produce hallow trichomes much longer than cotton (Cheema et al. 2010). For using such genes, comprehensive understanding about the development of cotton fiber is required so that the genetic circuits can be changed or new genes can be added which can supplement the existing mechanisms.

The scope of CRISPR technology for developing third generation GM cotton has several advantages including development of transgene-free systems (Zaidi et al. 2019). For example, gossypol-free cottonseed can be produced by silencing the genes conferring gossypols in seed. Major advantages of this assay are that the function of gene can be characterized and new cultivars can be evolved without introgressing foreign gene; hence, the technology will be acceptable to countries having skeptical views about the GM technology.

1.4 Future Prospects

Further enhancement in cotton production is possible by the integration of genetic resources as well as new genetic assays including DNA markers and transformation technologies. For using germplasm resources, a consortium of cotton scientists across all the cotton-growing countries should be involved in phenotypic as well as genotypic characterization. Expression studies at various developmental stages will also add synergy to characterization work. Use of robotic technologies can be very instrumental in collecting the huge quantity of data (particularly physiological) in least possible time from several hundreds of accessions in one experiment with limited hands. Earlier much emphasis was given to explore the upper parts of cotton plant. Diversion of focus is required to study the root traits that will help in identifying genes conferring high root biomass. These studies would also help in studying the interaction of microbes with roots—may lessen the dependency on chemical fertilizer. The generated data can be analyzed quickly using bioinformatic tools. Effective communication among the scientific community is required for dissemination of data for using in breeding programs.

Once these resources have been phenotypically characterized, efforts toward pangenome of the genus *Gossypium* will help in elucidating genetic circuits of important complex traits in cotton as well designing new DNA markers for the traits of interest. This will lead to the initiation of targeted breeding in cotton. Thus a required genetics of a cotton plant can be designed in a computer. For achieving this task, collaboration among cotton experts including molecular biologists, breeders, agronomists, physiologists, and extension workers is the need of time for making the cotton crop more competitive and profitable.

Cotton transformation assay for developing a transgenic plant is very lengthy as well as one or two genotypes respond to regeneration system. In the future, understanding the mechanism for converting callus to plant may lead to the shortening of the current timeline for developing resilience in transgenic cotton.

There are three genomes in a plant cell; however, success story of genetic transformation came from the nuclear genome transformation. The insertion of foreign gene is random. Secondly, the expression is either low or complete silencing of the introduced gene. While insertion of transgene in chloroplast is site specific and also results in very high expression (>70% of total soluble proteins), the genome of chloroplast does not travel through pollen, particularly in field crops, resultantly transgene can be contained. Till now, chloroplast transformation could not bear any fruit owing to laborious and lengthy tissue culture-based protocols for recovering transplastomic plants. Any successful effort will lead to the overexpression of the transgenes; thus protection to insect pest and harvesting high cotton yield with improved quality can be achieved.

Thus, it is summarized that the adoption of high-tech management practices, utilization of untapped genetic resources in breeding, cultivation of cotton varieties with excellent genetics, monitoring of risk and efficacy of transgene in ecosystem, and continued search for new genetic resources would help in sustaining cotton production.

Acknowledgments The idea for writing this chapter and the book entitled Cotton Precision BreedingEditors: Rahman, Mehboob Ur, Zafar, Yusuf, Zhang, Tianzhen (Eds.) was conceived when M.R visited China through a collaborative project "Mining of Elite Genes for High Yield, Super Fiber Qualities and Heat Tolerance and Their Usages in Pakistan and China Cotton Cultivars"—Project No. PSF/BSFC-AGR/P-NIBGE (12). Funds for this visit were provided by the Pakistan Science Foundation.

References

Altaf M, Stewart J, Murphy J (1997) Survey of cotton germplasm for terpenoid aldehydes important in host plant resistance. In: Oosterhuis D, Stewart JM (eds) Special reports—University of Arkansas Agricultural Experiment Station, pp 153–155

Aslam M, Ahmad K, Akhtar MA, Maqbool MA (2017) Salinity stress in crop plants: effects of stress, tolerance mechanisms and breeding strategies for improvement. *J Agric Basic Sci* 2:70–85

Balter M (2009) Clothes make the (hu) man. *Sci* 325(5946):1329

Bowman DT, Gutierrez OA (2003) Sources of fiber strength in the US upland cotton crop from 1980 to 2000. *J Cotton Sci* 7:164–169

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 Integr Plant Biol* 56:51–62

Cai C, Zhu G, Zhang T, Guo W (2017) High-density 80 K SNP array is a powerful tool for genotyping *G. hirsutum* accessions and genome analysis. *BMC Genomics* 18:654

Cantrell RG, Davis DD (2000) Registration of NM24016, an interspecific-derived cotton genetic stock. *Crop Sci* 40:1208

Cheema HMN, Bashir A, Khatoon A, Iqbal N, Zafar Y, Malik KA (2010) Molecular characterization and transcriptome profiling of expansin genes isolated from *Calotropis procera* fibers. *Electron J Biotechnol* 13(6):10–11

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD, Arioli T, Brubaker C, Cantrell RG, Lacape JM, Ulloa M, Chee P, Gingle AR, Haigler CH, Percy R, Saha S, Wilkins T, Wright RJ, Deynze AV, Zhu Y, Yu S, Abdurakhmonov I, Katageri I, Kumar PA, Rahman M, Zafar Y, Yu JZ, Kohel RJ, Wendel JF, Paterson AH (2007) Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol* 145(4):1303–1310

Chen ZJ, Sreedasyam A, Ando A, Song Q, Santiago LMD, Hulse-Kemp AH, Ding M, Ye W, Kirkbride RC, Jenkins J, Plott C, Lovell J, Lin YM, Vaughn R, Liu B, Simpson S, Scheffler BE, Wen L, Saski CA, Grover CE, Hu G, Conover JL, Carlson JW, Shu S, Boston LB, Williams M, Peterson DG, McGee K, Jones DC, Wendel JF, Stelly DM, Grimwood J, Schmutz J (2020) Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nat Genet* 52:525–533

Dempewolf H, Eastwood RJ, Guarino L, Khoury CK, Muller JV, Toll J (2014) Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecol Sustain Food Syst* 38:369–377

Dou L, Zhang X, Pang C, Song M, Wei H, Fan S, Yu S (2014) Genome-wide analysis of the WRKY gene family in cotton. *Mol Genet Genomics* 289:1103–1121

Feng L, Zhang S, Xing L, Yang B, Gao X, Xie X, Zhou B (2019) QTL analysis for yield and fibre quality traits using three sets of introgression lines developed from three *Gossypium hirsutum* race stocks. *Mol Genet Genomics* 294:789

Gore M, Percy RG, Zhang JF, Fang D (2012) Registration of NM24016/TM-1 RIL, a population of recombinant inbred lines of cotton (*Gossypium hirsutum* L.). *J Plant Regist* 6:124–127

Hansen CM, Bjorkman A (1998) The ultrastructure of wood from a solubility parameter point of view. *Holzforschung* 52:335–344

Hernandez J, Meints B, Hayes P (2020) Introgression breeding in barley: perspectives and case studies. *Front Plant Sci* 11:761

Hu Y, Chen J, Fang L, Zhang Z, Ma W, Niu Y, Ju L, Deng J, Zhao T, Lian J, Baruch K, Fang D, Liu X, Ruan Y-L, Rahman M, Han J, Wang K, Wang Q, Wu H, Mei G, Zang Y, Han Z, Xu C, Shen W, Yang D, Si Z, Dai F, Zou L, Huang F, Bai Y, Zhang Y, Brodt A, Ben-Hamo H, Zhu X, Zhou B, Guan X, Zhu S, Chen X, Zhang T (2019) *Gossypium barbadense* and *Gossypium hirsutum* genomes provide insights into the origin and evolution of allotetraploid cotton. *Nat Genet* 51:739–748

Hulse-Kemp AM, Lemm J, Plieske J, Ashrafi H, Buyyarapu R, Fang DD, Frelichowski J (2015) Development of a 63K SNP array for cotton and high density mapping of intraspecific and interspecific populations of *Gossypium* spp. *Genes Genomes Genet* 5(6):1187–1209

Iqbal MA, Rahman M (2017) Identification of marker-trait associations for lint traits in cotton. *Front Plant Sci* 8:86

Iqbal MA, Abbas M, Zafar Y, Rahman M (2015) Characterization of *Gossypium arboreum* L. genotypes for various fiber quality traits. *Pak J Bot* 47(6):2347–2354

Jayaraj S, Palaniswamy P (2005) Host plant resistance in cotton to major insect pests: perspectives and progress. In: Ignacimuthu S, Jayaraj S (eds) Sustainable insect pest management. Alpha Science Int'l Ltd, New Delhi

Jia Y, Sun J, Wang X, Zhou Z, Pan Z, He S, Pang B, Wang L, Du X (2014) Molecular diversity and association analysis of drought and salt tolerance in *G. hirsutum* L. germplasm. *J Integr Agric* 13:1845–1853

Kalivas A, Xanthopoulos F, Kehagia O, Tsaftaris AS (2011) Agronomic characterization, genetic diversity and association analysis of cotton cultivars using simple sequence repeat molecular markers. *Genet Mol Res* 10:208–217

Khan SA, Zafar Y, Briddon RW, Malik KA, Mukhtar Z (2006) Spider venom toxin protects plants from insect attack. *Transgenic Res* 15:349–357

Khan A, Pan X, Najeeb U, Tan DKU, Fahad S, Zahoor R, Luo Z (2018) Coping with drought: stress and adaptive mechanisms, and management through cultural and molecular alternatives in cotton as vital constituents for plant stress resilience and fitness. *Biol Res* 51:47

Kvavadze E, Bar-Yosef O, Belfer-Cohen A, Boaretto E, Jakeli N, Matskevich Z, Meshveliani T (2009) 30,000-year-old wild flax fibers. *Science* 325(5946):1359

Li F, Fan G, Wang K, Sun F, Yuan Y, Song G, Li Q, Ma Z, Lu C, Zou C, Chen W, Liang X, Shang H, Liu W, Shi C, Xiao G, Gou C, Ye W, Xu X, Zhang X, Wei H, Li Z, Zhang G, Wang J, Liu K, Kohel RJ, Percy RG, Yu JZ, Zhu YX, Wang J, Yu S (2014) Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nat Genet* 46:567–572

Li F, Fan G, Lu C, Xiao G, Zou C, Kohel RJ, Ma Z, Shang H, Ma X, Wu J, Liang X, Huang G, Percy RG, Liu K, Yang W, Chen W, Du X, Shi C, Yuan Y, Ye W, Liu X, Zhang X, Liu W, Wei H, Wei S, Huang G, Zhang X, Zhu S, Zhang H, Sun F, Wang X, Liang J, Wang J, He Q, Huang L, Wang J, Cui J, Song G, Wang K, Xu X, Yu JZ, Zhu Y, Yu S (2015) Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insight into genome evolution. *Nat Biotechnol* 33:524–530

Li CQ, Xu XJ, Dong N, Ai NJ, Wang QL, Jenkins J (2016) Association mapping identifies markers related to major early-maturating traits in upland cotton (*Gossypium hirsutum* L.). *Plant Breed* 135:483–491

Liu JS, Yang BX, Gao ZY, Yang SJ, Zheng F, Yan Y, Ban ZJ (2005) High fiber strength medium and long staple cottons developed by hybridization between upland and sea-island cotton. *China Cottons* 32:12–14

Liu X, Zhao B, Zheng HJ, Hu Y, Lu G, Yang CQ, Chen JD, Chen JJ, Chen DY, Zhang L, Zhou Y, Wang LJ, Guo WZ, Bai YL, Ruan JX, Shangguan XX, Mao YB, Shan C, Jiang JP, Zhu YQ, Jin L, Kang H, Chen ST, He X, Wang R, Wang YZ, Chen J, Wang LJ, Yu ST, Wang BY, Wei J, Song SC, Lu XY, Gao ZC, Gu WY, Deng X, Ma D, Wang S, Liang WH, Fang L, Cai CP, Zhu XF, Zhou BL, Chen ZJ, Xu SH, Zhang YG, Wang SY, Zhang TZ, Zhao GP, Chen XY (2015) *Gossypium barbadense* genome sequence provides insight into the evolution of extra-long staple fiber and specialized metabolites. *Sci Rep* 5:14139

Lv SL, Lian LJ, Tao PL, Li ZX, Zhang KW, Zhang JR (2009) Overexpression of *thellungiella halophila* H + -PPase (TsVP) in cotton enhances drought stress resistance of plants. *Planta* 229:899–910

Ma FZ, Liu JS (1982) Genetic advances in economic traits of interspecific progenies between *Gossypium hirsutum* and *G. barbadense*. *J China Agric Univ* 8:23–32

Malik HJ, Raza A, Amin I, Scheffler JA, Scheffler BE, Brown JK, Mansoor S (2016) RNAi-mediated mortality of the whitefly through transgenic expression of double-stranded RNA homologous to acetylcholinesterase and ecdysone receptor in tobacco plants. *Sci Rep* 6:38469

May OL (2001) Registration of PD 94045 germplasm line of upland cotton. *Crop Sci* 41:279–280

Mergeai G (2006) Introgressions interspécifiques chez le cotonnier. *CAH Agric* 15:135–143

Mills G, Sharps K, Simpson D, Pleijel H, Frei M, Burkey K, Emberson L, Uddling J, Broberg M, Feng Z, Kobayashi K, Agrawal M (2018) Closing the global ozone yield gap: quantification and cobenefits for multistress tolerance. *Glob Chang Biol* 24:4869–4893

Nazeer W, Ahmad S, Mahmood K, Tipu A, Mahmood A, Zhou B (2014) Introgression of genes for cotton leaf curl virus resistance and increased fiber strength from *Gossypium stocksii* into upland cotton (*Gossypium hirsutum*). *Genet Mol Res* 13:1133–1143

Ni M, Ma W, Wang X, Gao M, Dai Y, Wei X, Zhang L, Peng Y, Chen S, Ding L, Tian Y, Li J, Wang H, Wang X, Xu G, Guo W, Yang Y, Wu Y, Heuberger S, Tabashnik BE, Zhang T, Zhu Z (2017) Next-generation transgenic cotton: pyramiding RNAi and Bt counters insect resistance. *Plant Biotechnol J* 15:1204–1213

Nie XH, Huang C, You CY, Li W, Zhao WX, Shen C, Zhang B, Wang H, Yan Z, Dai B, Wang M, Zhang X, Lin Z (2016) Genome-wide SSR-based association mapping for fiber quality in nation-wide upland cotton inbreed cultivars in China. *BMC Genomics* 17:352

Patel JD, Wright RJ, Auld D, Chandnani R, Goff VH, Ingles J, Pierce GJ, Torres MJ, Paterson AH (2014) Alleles conferring improved fiber quality from EMS mutagenesis of elite cotton genotypes. *Theor Appl Genet* 127:821–830

Patel JD, Wright RJ, Chandnani R, Goff VH, Ingles J, Paterson AH (2016) EMS-mutated cotton populations suggest overlapping genetic control of trichome and lint fiber variation. *Euphytica* 208:597–608

Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, Llewellyn D, Showmaker KC, Shu S, Udall J (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492:423–427

Percy RG, Kohel RJ (1999) Qualitative genetics. In: Smith CW, Cothren JT (eds) Cotton: origin, history, technology and production. Wiley, New York, pp 319–360

Percy RG, Zhang JF, Cantrell RG (2005) Genetic variation for agronomic and fiber properties in an introgressed recombinant inbred population of cotton. *Crop Sci* 46:1311–1317

Poehlman JM (1987) Breeding field crops, 1st edn. Springer, Netherlands

Rahman M (2016) Cotton improvement for economically stressed economies. ICAC *Recorder* 1:12–13

Rahman M, Zafar Y (2018) Introductory chapter: updates on achieving sustainable cotton production. In: Rahman M, Zafar Y (eds) Past, present and future trends in cotton breeding. Intech Publisher, pp 3–10

Rahman M, Hussain D, Zafar Y (2002) Estimation of genetic divergence among elite cotton (*Gossypium hirsutum* L.) cultivars/genotypes by DNA fingerprinting technology. *Crop Sci* 42:2137–2144

Rahman M, Hussain D, Malik T, Zafar Y (2005) Genetics of resistance to cotton leaf curl disease in *Gossypium hirsutum*. *Plant Pathol* 54:764–772

Rahman M, Ullah I, Ahsraf M, Stewart JM, Zafar Y (2008) Genotypic variation for drought tolerance in cotton. *Agron Sustain Dev* 28:439–447

Rahman M, Zafar Y, Paterson AH (2009) *Gossypium* DNA markers types, number and uses. In: Paterson AH (ed) Genomics of cotton. Springer, Dordrecht

Rahman M, Shaheen T, Tabbasam N, Iqbal MA, Ashraf M, Zafar Y, Paterson AH (2012) Cotton genetic resources: a review. *Agron Sustain Dev* 32:419–432

Rahman M, Rahmat Z, Mahmood A, Abdullah K, Zafar Y (2014) Cotton germplasm of Pakistan world cotton germplasm resources, Ibrokhim Y. Abdurakhmonov, IntechOpen. https://doi.org/10.5772/58620

Rahman M, Zaman M, Shaheen T, Irem S, Zafar Y (2015) Safe use of Cry genes in genetically modified crops. *Environ Chem Lett* 13:239–249

Rahman M, Khan AQ, Rahmat Z, Iqbal MA, Zafar Y (2017) Genetics and genomics of cotton leaf curl disease, its viral causal agents and whitefly vector: a way forward to sustain world cotton fiber security. Frontier in Plant Sciences (Virology). https://doi.org/10.3389/fpls.2017.01157

Rahmat Z, Ahmad S, Zia Z, Rahman M (2019) Genetic monitoring of introgressed alleles from G. arboreum into G. hirsutum using SSR markers: a potential approach for bringing new alleles under cultivation. *Pak J Bot* 51(2):479–486

Ray DK, Mueller ND, West PC, Foley JA (2013) Yield trends are insufficient to double global crop production by 2050. *PLoS One* 8:e66428

Raza A, Malik HJ, Shafiq M, Amin I, Scheffler JA, Scheffler BE, Mansoor S (2016) RNA interference based approach to down regulate osmoregulators of whitefly (*Bemisia tabaci*): potential technology for the control of whitefly. *PLoS One* 11:e0153883